





Automated Chagas Vector Detection Using Artificial Intelligence

Victoria Ludeña, Estudiante¹, Cristian Pinzón, Doctorado², Javier Bajo, Doctorado³, Emilio Serrano, Doctorado⁴
^{1,2}Universidad Tecnológica de Panamá, Panamá; ^{3,4}Universidad Politécnica de Madrid, Madrid,
victoria.ludena@utp.ac.pa¹; cristian.pinzon@utp.ac.pa²; javier.bajo@upm.es³; emilio.serrano@upm.es⁴

Abstract- This scientific article proposes the application of artificial intelligence as an innovative strategy to improve the detection of the triatomine vector of Chagas disease in Panama. An analysis of various methods and algorithms previously used in pathogen detection, including multilayer architectures, Resnet34 and MALDI-TOF, is performed, with a particular focus on the Rhodnius and Triatoma families, which are the main vectors of this disease in the region. Through a local search and literature review, together with the collection of relevant data, a model was designed using artificial intelligence algorithms. This model has as its main objective the identification of vectors through images, offering an assistance tool for specialists in the field. The project culminates with the development of a free software application for specialists, capable of differentiating between two species of triatomines, thus providing an innovative resource for the fight against Chagas disease.

Keywords: Chagas disease, computer vision, neural networks, triatomine bugs, bed bugs, application, identification, image recognition

Deteccción Automatizada del Vector de Chagas Usando Inteligencia Artificial

Victoria Ludeña, Estudiante¹, Cristian Pinzón, Doctorado², Javier Bajo, Doctorado³, Emilio Serrano, Doctorado⁴
^{1,2}Universidad Tecnológica de Panamá, Panamá; ^{3,4}Universidad Politécnica de Madrid, Madrid,
 victoria.ludena@utp.ac.pa¹; cristian.pinzon@utp.ac.pa²; javier.bajo@upm.es³; emilio.serrano@upm.es⁴

Abstract- *This scientific article proposes the application of artificial intelligence as an innovative strategy to improve the detection of the triatomine vector of Chagas disease in Panama. An analysis of various methods and algorithms previously used in pathogen detection, including multilayer architectures, Resnet34 and MALDI-TOF, is performed, with a particular focus on the Rhodnius and Triatoma families, which are the main vectors of this disease in the region. Through a local search and literature review, together with the collection of relevant data, a model was designed using artificial intelligence algorithms. This model has as its main objective the identification of vectors through images, offering an assistance tool for specialists in the field. The project culminates with the development of a free software application for specialists, capable of differentiating between two species of triatomines, thus providing an innovative resource for the fight against Chagas disease.*

Keywords: *Chagas disease, computer vision, neural networks, triatomine bugs, bed bugs, application, identification, image recognition*

* Corresponding author: cristian.pinzon@utp.ac.pa

I. Introducción

En Panamá, el vector del Chagas conocido como Triatomino, es una especie de chinche hematófago, del cual se tiene conocimiento de su primera aparición en el país en la provincia de Colón[1]. El término vector, en el contexto médico, se refiere a un organismo vivo que facilita la transmisión de un agente infeccioso desde un huésped animal, potencialmente infectado, hacia un ser humano o hacia otro animal. Actualmente se han identificado 11 especies transmisoras en todo el país, de las cuales 2 han sido las seleccionadas para la realización de este proyecto debido a su capacidad de adaptación al medio.

Para brindar asistencia a los profesionales de la salud en el ámbito de los vectores, se plantea la implementación de algoritmos de inteligencia artificial para simplificar la clasificación e identificación de estos insectos. Dada la complejidad y el nivel de conocimientos

específicos requeridos para esta tarea, el uso de la visión por computadora junto con técnicas de Aprendizaje Profundo (Deep Learning) presenta una solución innovadora que promete eficiencia y precisión. En este contexto, el objetivo principal se establece en la construcción de un prototipo para la detección automatizada del vector de Chagas. Este prototipo se basará en la identificación y clasificación de imágenes utilizando una red neuronal[2] y un conjunto de datos (dataset)[3] propio de la región a estudiar.

El artículo se encuentra estructurado de la siguiente manera: la primera sección corresponde a la introducción, la segunda sección trata la descripción de la problemática que se plantea resolver, la tercera sección da a conocer la metodología de investigación empleada, en el cuarto apartado se describen los materiales y métodos utilizados para la construcción del prototipo, la quinta sección muestra los resultados, la última sección presenta las conclusiones del trabajo.

II. Marco de referencia

A. Trabajos de referencia

Se llevó a cabo una revisión exhaustiva de diversas propuestas para la identificación de insectos, con el fin de facilitar la identificación de aquellos que son el foco de interés en esta investigación. A continuación, se presentan los enfoques examinados:

- Multilayered architecture: Planteado en Macia De Oliveira, 2017[4], propone el uso de una arquitectura de tres niveles, donde el segundo nivel se usa el lenguaje PHP scripts para el control de la aplicación.
- Resnet34: Propuesto en Cochero, J, 2022[5] usa técnicas de Deep Learning[6] y procesamiento de las imágenes con el objetivo de optimizar el entrenamiento de la red.
- MALDI-TOF: En Merchan, F, 2023[7] se propone la extracción de proteínas con etanol y ácido fórmico de partes específicas de los artrópodos que permite una identificación rápida y precisa.

En base a estos trabajos, se plantea una solución innovadora utilizando algoritmos de Inteligencia Artificial, concretamente algoritmos de aprendizaje automático para identificar los vectores en las regiones de Panamá.

En la fase actual del proyecto, se ha diseñado un estudio de caso enfocado en la identificación de dos géneros pertenecientes a subfamilias de hemípteros presentes en Panamá. Estos géneros son de particular interés para el Ministerio de Salud en Panamá. Uno de ellos es el *Rhodnius*, caracterizado por una cabeza alargada y antenas ubicadas en la parte frontal, cerca del clipeo. El otro es el *Triatoma*, que se distingue por tener una cabeza de longitud media, con antenas situadas entre los ojos y el clipeo [8]

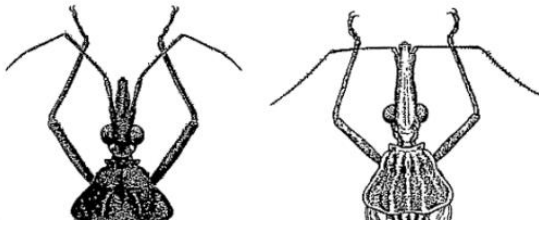


Fig 1. Chinchas en estudio a) *Triatoma Dimidiata* b) *Rhodnius Pallescens*. Fuente: Revista de investigación de la Universidad de Panamá[1]

B. Chinchas en Panamá

En Panamá sobresalen como los principales huéspedes de triatomíneos la Zarigüeya Común (*Didelphis marsupialis*), el Armadillo Blanco (*Dasybus novemcinctus*), el Hormiguero Bandera (*Tamandua mexicana*), el Mono Titi común (*Saguinus oedipus*), la Rata Negra de los Tejados (*Rattus rattus*), el Mocangué (*Proechimys semispinosus*) y el Murciélago Frutero Común (*Artibeus jamaicensis*).



Fig 2. Zarigüeya Común. Fuente: Pixabay[9]

B. Materiales: Tecnologías implicadas

Las tecnologías utilizadas para el desarrollo del prototipo propuesto se mencionan a continuación:

- Python[10]: La herramienta es utilizada para el desarrollo de los algoritmos de creación y

obtención de características, así como pruebas con una red neuronal artificial.

- JavaScript[11]: Esta herramienta se requiere para programar la captura de las imágenes y la interacción con los elementos HTML y CSS.
- Flask [12]: En el marco de nuestro proyecto, esta herramienta facilita la integración de Python dentro de un entorno web, permitiendo una interacción entre diferentes lenguajes de programación para el procesamiento de solicitudes.
- Tensorflow[13]: Esta plataforma de código abierto es indispensable en el proyecto ya que provee los algoritmos de aprendizaje automático necesarios dentro del proyecto.

A continuación, se detalla la metodología de investigación empleada en el desarrollo de este proyecto.

III. Metodología

La investigación propuesta es de naturaleza aplicada y se orienta hacia la resolución de un problema concreto. Adopta inicialmente un enfoque descriptivo, donde se busca comprender y detallar el problema mediante la recopilación exhaustiva de datos y la formulación de una solución potencial.

A continuación, se plantea las metodologías utilizadas.

A. Metodología de investigación

a) Identificación del problema

Se llevó a cabo una búsqueda exhaustiva de herramientas locales orientadas a la identificación de triatomíneos, obteniendo resultados limitados. Las herramientas disponibles se orientan mayoritariamente hacia la identificación de insectos a nivel macro en múltiples países. Además, la herramienta más cercana al enfoque deseado no se especializa en la distinción entre especies de la familia de los triatomíneos, ni está adaptada específicamente a las condiciones de Panamá y sus regiones.

b) Revisión bibliográfica

Se efectuaron búsquedas de información científica del tema objeto de interés. Como herramientas de búsqueda se utilizó Google Scholar[14], UTP-Ridda2[15], Elsevier[16] y otras fuentes de información.

c) Recolección de datos

Se recopiló imágenes del *Triatoma Dimidiata* y *Rhodnius Pallescens* del Instituto Conmemorativo Gorgas de Estudios de la Salud[17] contando también con la colaboración del Smithsonian

Tropical Research Institute[18] en la provincia de Panamá, capital, ver Fig. 3. Este proceso se realiza específicamente en esta zona, ya que el proyecto se centra en Panamá en sus etapas iniciales. Esta decisión se debe a que en esta región se encuentran las bases de datos más completas sobre chinches, lo cual es fundamental para el desarrollo del proyecto.



Fig 3. Área geográfica de captación de los chinches que componen el repositorio. Fuente: AnaMapa [19]

d) Diseño del modelo

Se realizan varios diseños conceptuales del prototipo, y se escogió el diseño adecuado al problema. Luego de este paso se inicia la construcción y prueba del prototipo.

B. Metodología de software

Se plantea una aplicación de software desde la perspectiva de un sistema experto, para esto se requiere un marco estructurado de sistemas basados en el conocimiento.

IV. Prototipo propuesto

A continuación, se presenta el diseño conceptual en la Fig. 4 del prototipo planteado donde se muestran las herramientas de software y hardware a utilizar, las librerías y el flujo de datos que se debe realizar para el funcionamiento del prototipo.

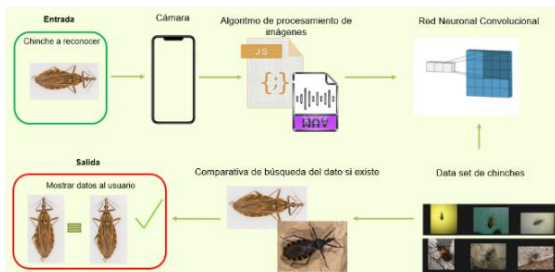


Fig 4. Modelo conceptual del proyecto. Fuente: Propia

A continuación, se explica el funcionamiento del prototipo paso a paso:

A. Captación de imágenes

El proceso inicia mediante la captura de la imagen del chinche que se quiere identificar utilizando un dispositivo móvil o una imagen descargada, ver Fig 5.



Fig 5. Captura de la imagen del chinche. Fuente: Propia

B. Red neuronal convolucional

Se utilizó una red neuronal convolucional [22] (CNN) para encontrar patrones en la imagen. Esta red neuronal artificial realiza el proceso de identificación de la especie de insecto con una buena precisión.

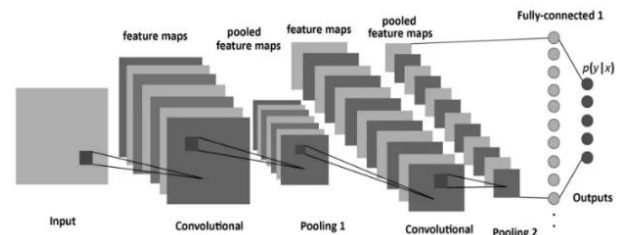


Fig 6. Representación de la CNN. Fuente: IALATAM[20]

C. Conjunto de datos

El conjunto de datos, o dataset, utilizado en este proyecto consiste en una colección cuidadosamente seleccionada y previamente recopilada de datos. Para el caso que nos ocupa, se almacenaron imágenes de chinches que se utilizaron para entrenar la red neuronal convolucional (CNN). Este dataset se creó capturando fotografías de los especímenes seleccionados desde diversos ángulos, aprovechando los recursos disponibles en el Instituto Gorgas y el Smithsonian. Las imágenes se tomaron utilizando la cámara de un teléfono inteligente, contando además con la guía de expertos en la materia para asegurar la captura de los rasgos más significativos. Este enfoque meticuloso tenía el objetivo de conformar un conjunto de datos lo más completo y representativo posible, utilizando los recursos a nuestro alcance.

D. Desarrollo del prototipo

En los siguientes párrafos se explica cómo se desarrolló el prototipo y un resumen de los entrenamientos llevados a cabo por la red CNN para realizar las pruebas del prototipo.

a) Creación del prototipo

El modelo seleccionado de la CNN para la creación y entrenamiento de la red neuronal fue de tipo Secuencial, compuesto por capas convolucionales, capas de agrupación máxima, una capa plana y capas densas. El entrenamiento se realizó bajo las condiciones de 50 épocas, se estructuró con tres capas de convolución, una capa plana para convertir la estructura dimensional en datos, además se hizo uso de técnicas de regulación con capas dropout a fin de evitar el sobre ajuste del modelo.

Se creó una interfaz usando el lenguaje de programación JavaScript, que consta de 3 botones para cargar la imagen desde el dispositivo, otro para identificar la imagen del chinche y un último botón para limpiar la imagen y el nombre generado como salida por el prototipo como se muestra en la Fig 7.



Fig 7. Interfaz del prototipo Fuente: Propia

El primer botón está diseñado para cargar la imagen del chinche. Al utilizarlo, se sube una imagen al prototipo y se almacena en un archivo.

El segundo botón desempeña la función de identificar el espécimen. Una vez activado, el modelo de Red Neuronal Convolucional (CNN) previamente entrenado procede a realizar la predicción. Este modelo y sus respectivos pesos se han convertido y almacenado en archivos con formato .h5 para su utilización y manejo posterior. El siguiente paso consistió en cargar los archivos de imágenes, y luego, se clasifica en base a las categorías de chinches que conoce la red CNN.

V. Resultados

Los resultados obtenidos en esta fase del prototipo fueron satisfactorios. Durante el entrenamiento de

la red neuronal, se logró alcanzar una precisión del 78%. Es importante resaltar que, durante este proceso, se emplearon conjuntos de datos que incluían 160 imágenes de chinches en cada uno, lo que permitió alcanzar un nivel de eficacia aceptable en el modelo.

En situaciones donde se presenten desafíos como la mezcla de dos chinches o la presencia de ruido ambiental, se prioriza la opción que muestre una mayor probabilidad asociada. Sin embargo, es importante reconocer que esta selección no siempre garantiza ser la decisión definitiva o la más acertada.

Se ha desarrollado un prototipo inicial de aplicación que integra una interfaz impulsada por una red neuronal convolucional. Esta red está diseñada para identificar dos familias específicas de chinches y se complementa con un módulo informativo. Dicho módulo, en etapas futuras, se planea ampliar con el objetivo de educar a la población sobre la importancia de reconocer estos insectos, que son vectores de diversas enfermedades, particularmente en regiones donde las condiciones ambientales favorecen su proliferación. Además, el diseño de este prototipo busca facilitar a los investigadores y especialistas la comprensión de sus principios fundamentales, permitiéndoles, en el mejor de los casos, proporcionar retroalimentación que contribuya a la mejora y el refinamiento del producto final. La interfaz del prototipo cuenta con una página de inicio que permite a los usuarios cargar imágenes de chinches para su identificación, así como una sección de ayuda diseñada para garantizar una interacción intuitiva y accesible para los nuevos usuarios.

A. Mejoras futuras del prototipo

El prototipo se encuentra en desarrollo para mejorar los resultados obtenidos. En esta dirección se plantean algunas mejoras:

- Mejorar la interfaz gráfica del prototipo para tener una mejora visual.
- Aumentar el conjunto de datos de cada chinche, además de agregar nuevas especies que representen un riesgo para la salud dentro de las que se encuentran clasificadas.
- Desarrollar una aplicación para dispositivos móviles accesible y de fácil uso para el usuario.

B. Impacto del proyecto

Este proyecto no solo servirá como una herramienta avanzada para la identificación de

vectores por parte de especialistas en salud y enfermedades tropicales, sino que también desempeñará un papel crucial como recurso educativo para la población en general. Al facilitar el acceso a información detallada y fácil de entender sobre estos vectores. La aplicación ayudará a aumentar la conciencia y el conocimiento entre la comunidad sobre la importancia y características de estos organismos. De esta manera, contribuirá significativamente a la educación pública en temas de salud y prevención, potenciando una comprensión más amplia y profunda de los vectores y su impacto en el entorno.

VI. Conclusiones

Se ha desarrollado un prototipo para la identificación de chinches vectores de la enfermedad de Chagas, integrando algoritmos especializados de Inteligencia Artificial. Los aspectos destacables de esta etapa incluyen:

- La creación de una aplicación de software que permite identificar posibles vectores de la enfermedad de Chagas mediante una imagen tomada, por ejemplo, desde un teléfono móvil.
- La creación de un conjunto de datos inicial dedicado al chinche transmisor de la enfermedad de Chagas marca un paso importante en nuestro proyecto. Se llevarán a cabo más esfuerzos para enriquecer la calidad de este conjunto de datos, con el objetivo de aumentar la precisión en la identificación de los chinches.
- El apoyo a la innovación en este campo contribuirá significativamente a la eficacia de las estrategias de lucha contra la enfermedad y fortalecerá la capacidad de los sistemas de salud para abordar este problema de manera integral

VII. Agradecimientos

Al Dr. Edmanuel Cruz, Dr. Carlos Bracho, Doctores Elsy Abreu y Tito Ludeña por el apoyo en la realización de esta investigación, así como también a la Doctora Annette Aiello y al Doctor Lorenzo Cáceres, por su apoyo con la obtención de imágenes de los especímenes

VIII. Referencias

- [1] A. Navarro, Quirós Dora, y G. Contreras, “Revista de investigación de la Universidad de Panamá”.
- [2] “Qué son las Redes Neuronales Convolucionales”. Consultado: el 2 de junio de 2022. [En línea]. Disponible en: <https://www.iebschool.com/blog/redes-neuronales-convolucionales-big-data/>
- [3] “Datasets: Qué son y cómo acceder a ellos | OpenWebinars”. Consultado: el 2 de enero de 2024. [En línea]. Disponible en: <https://openwebinars.net/blog/datasets-que-son-y-como-acceder-a-ellos/>
- [4] L. Márcia de Oliveira *et al.*, “TriatoKey: A web and mobile tool for biodiversity identification of Brazilian triatomine species”, *Database*, vol. 2017, 2017, doi: 10.1093/database/bax033.
- [5] J. Cochero, L. Pattori, A. Balsalobre, S. Ceccarelli, y G. Marti, “A convolutional neural network to recognize Chagas disease vectors using mobile phone images”, *Ecol Inform*, vol. 68, may 2022, doi: 10.1016/j.ecoinf.2022.101587.
- [6] Redacción InnovaciónDigital360, “Aprendizaje profundo o Deep Learning, qué es y qué aplicaciones tiene hoy en día”.
- [7] F. Merchan, K. Contreras, R. A. Gittens, J. R. Loaiza, y J. E. Sanchez-Galan, “Deep metric learning for the classification of MALDI-TOF spectral signatures from multiple species of neotropical disease vectors”, *Artificial Intelligence in the Life Sciences*, vol. 3, p. 100071, dic. 2023, doi: 10.1016/j.aisci.2023.100071.
- [8] C.-P. Dalmiro, “Revisión de los vectores de la enfermedad de Chagas en Venezuela (Hemiptera-Heteroptera, Reduviidae, Triatominae)”.
- [9] “Más de 1 millón de Imágenes Gratis para Descargar - Pixabay - Pixabay”. Consultado: el 1 de febrero de 2024. [En línea]. Disponible en: <https://pixabay.com/es/>
- [10] “Applications for Python | Python.org”. Consultado: el 4 de enero de 2024. [En línea]. Disponible en: <https://www.python.org/about/apps/>

- [11] “¿Qué es JavaScript? - Aprende desarrollo web | MDN”. Consultado: el 30 de enero de 2024. [En línea]. Disponible en: https://developer.mozilla.org/es/docs/Learn/JavaScript/First_steps/What_is_JavaScript
- [12] “¿Qué es Flask? Breve tutorial - IONOS”. Consultado: el 30 de enero de 2024. [En línea]. Disponible en: <https://www.ionos.es/digitalguide/paginas-web/desarrollo-web/flask/>
- [13] “¿Qué es TensorFlow y para qué sirve?” Consultado: el 30 de enero de 2024. [En línea]. Disponible en: <https://www.incentro.com/es-ES/blog/que-es-tensorflow>
- [14] “Google Académico”. Consultado: el 1 de junio de 2022. [En línea]. Disponible en: <https://scholar.google.es/schhp?hl=es>
- [15] “UTP-Ridda2”. Consultado: el 1 de junio de 2022. [En línea]. Disponible en: <https://ridda2.utp.ac.pa/>
- [16] “Elsevier | Un negocio de análisis de información”. Consultado: el 31 de enero de 2024. [En línea]. Disponible en: <https://www.elsevier.com/es-es>
- [17] “Instituto Conmemorativo Gorgas de Estudios de la Salud (ICGES) | Líderes de Investigación, comprometidos con la solución de los problemas de la salud.” Consultado: el 30 de enero de 2024. [En línea]. Disponible en: <https://www.gorgas.gob.pa/>
- [18] “Smithsonian Tropical Research Institute”. Consultado: el 30 de enero de 2024. [En línea]. Disponible en: <https://stri.si.edu/es>
- [19] “Mapa de Panamá - AnnaMapa.com”. Consultado: el 1 de febrero de 2024. [En línea]. Disponible en: <https://annamapa.com/panama/>
- [20] “Entendiendo las redes neuronales: De la neurona a RNN, CNN y Deep Learning - IA Latam”. Consultado: el 1 de febrero de 2024. [En línea]. Disponible en: <https://ia-latam.com/2019/02/06/entendiendo-las-redes-neuronales-de-la-neurona-a-rnn-cnn-y-deep-learning/>