

Universidad del Valle de Guatemala

Facultad de ingeniería

Data Science

Catedrático: Luís Furlán



Proyecto 2.

Análisis Exploratorio

Nelson Eduardo García Bravatti 22434

Joaquín André Puente Grajeda 22296

José Antonio Mérida Castejón 201105

Guatemala, septiembre de 2025

Planteamiento Inicial del Problema

1. Situación Problemática

Las enfermedades transmitidas por mosquitos (p. ej., dengue, zika y chikunguña) requieren sistemas de vigilancia oportunos. Los métodos tradicionales son costosos y lentos; por ello, proyectos de ciencia ciudadana como Mosquito Alert permiten que la población reporte observaciones y fotos de mosquitos, que posteriormente son validadas por especialistas y utilizadas para la gestión de salud pública. Automatizar la validación de imágenes reduciría la carga de trabajo y aceleraría las decisiones de control vectorial.

2. Problema Científico

¿Es posible desarrollar un modelo de visión por computadora que detecte y clasifique mosquitos vectores en imágenes capturadas por ciudadanos, reduciendo la carga de validación manual?

3. Objetivo General

Desarrollar un modelo de aprendizaje automático que detecte y clasifique mosquitos vectores en imágenes capturadas por ciudadanos.

4. Objetivos Específicos

- a. Realizar un análisis exploratorio del dataset de imágenes de mosquitos para identificar características relevantes y desafíos de los datos.
- b. Implementar técnicas de preprocesamiento de imágenes y estrategias para abordar desbalances de clases y variabilidad en calidad de imágenes.
- c. Entrenar y comparar diferentes arquitecturas de detección y clasificación de objetos para evaluar su desempeño en la identificación de mosquitos.
- d. Evaluar los modelos utilizando métricas apropiadas para tareas de detección y clasificación, seleccionando el modelo con mejor rendimiento.

- e. Documentar la solución desarrollada para garantizar reproducibilidad y facilitar su implementación.

Investigación Preliminar

Enfermedades Transmitidas por Vectores

A nivel global, las enfermedades transmitidas por vectores representan más del 17% de todas las enfermedades infecciosas y causan más de 700,000 muertes anuales (WHO, 2024). El dengue se ha convertido en la enfermedad viral transmitida por vectores de mayor propagación, afectando aproximadamente a 3,900 millones de personas en más de 132 países (WHO, 2024). La región de las Américas vivió la mayor epidemia de dengue registrada en 2024, con más de 13 millones de casos reportados, triplicando el récord del año anterior (PAHO, 2025).

Dataset del Mosquito Alert Challenge

El dataset contiene imágenes anotadas de seis clases que incluyen los principales vectores de arbovirosis y malaria: *Aedes aegypti*, *Aedes albopictus*, *Anopheles* (género), *Culex* (género), *Culiseta* (género) y el complejo *Aedes japonicus/koreicus*.

Especies y Géneros del Dataset

Aedes aegypti (especie)

Rasgos morfológicos distintivos: patrón en "lira" blanco característico en el tórax; cría principalmente en recipientes artificiales con agua estancada; actividad de picadura diurna con picos al amanecer y atardecer. Vectores principales: dengue, chikungunya, Zika y fiebre amarilla. Esta especie es responsable de la transmisión primaria del dengue, con casos que aumentaron de 26.45 millones a 58.96 millones entre 1990 y 2021 (Zhang et al., 2025).

Aedes albopictus (especie)

Conocido como "mosquito tigre asiático" por la franja blanca única longitudinal en el centro del tórax; tolera climas más fríos que *Ae. aegypti*, permitiendo su expansión a regiones templadas. Vectores: chikungunya, dengue y otros arbovirus; presenta menor antropofilia que *Ae. aegypti*, lo que puede influir en los patrones de transmisión.

Aedes japonicus/Aedes koreicus (complejo)

Mosquitos invasores de zonas templadas, morfológicamente similares, agrupados como complejo debido a la dificultad de diferenciación. Competencia vectorial demostrada en laboratorio para CHIKV/ZIKV/WNV, aunque su relevancia epidemiológica en campo permanece incierta.

Anopheles (género)

Características únicas: palpos tan largos como la probóscide y postura de reposo característica con el abdomen elevado; actividad crepuscular y nocturna. Importancia epidemiológica: únicos transmisores de malaria en humanos (solo hembras), causando aproximadamente 249 millones de casos y más de 608,000 muertes anuales (WHO, 2024).

Culex (género)

Comportamiento: actividad nocturna predominante; muchas especies prefieren alimentarse de aves (ornitofílicas). Vectores importantes de virus del Nilo Occidental (WNV) y encefalitis según especie y región; *Cx. pipiens*, *Cx. quinquefasciatus* y *Cx. tarsalis* son especies clave en la transmisión de WNV.

Culiseta (género)

Adaptaciones ecológicas: varias especies adaptadas a climas fríos; generalmente de mayor tamaño corporal. Importancia vectorial específica: *Culiseta melanura* mantiene el ciclo enzoótico de encefalitis equina del Este (EEE).

Impacto Económico y Salud Pública

Carga económica

Las especies invasoras *Aedes* (especialmente *aegypti* y *albopictus*) han acumulado costos reportados de al menos US\$ 94.7 mil millones durante 45 años (1975-2020), con costos que se han incrementado 14 veces desde los años 2000 (Roiz et al., 2024). Los gastos en daños y pérdidas superan en diez veces la inversión en prevención y manejo.

Impacto sanitario

En 2024, solo en las Américas se reportaron más de 13 millones de casos de dengue, con 22,684 casos severos (0.17%) y 8,186 muertes (tasa de letalidad 0.063%) (PAHO, 2025). La carga global del dengue se ha duplicado aproximadamente entre 1990 y 2021 (Zhang et al., 2025).

Análisis Inicial

Problema y Dataset

El problema planteado corresponde a una tarea híbrida de detección y clasificación de objetos en imágenes reales capturadas en condiciones no controladas. El Mosquito Alert Challenge Phase 1 proporciona un conjunto de imágenes reales anotadas con bounding boxes y etiquetas de clase, reflejando las condiciones reales de captura por parte de ciudadanos sin entrenamiento especializado. Las seis clases incluyen tanto especies específicas (*Aedes aegypti*, *A. albopictus*)

como géneros completos (*Anopheles*, *Culex*, *Culiseta*) y un complejo de especies (*A. japonicus/koreicus*), lo que introduce variabilidad en el nivel taxonómico de clasificación.

Dificultades Previstas

Las imágenes capturadas por ciudadanos presentan condiciones heterogéneas de iluminación, enfoque, resolución y ángulos de captura, contrastando con datasets controlados de laboratorio. Además, es probable que exista una distribución desigual entre las clases, especialmente considerando que algunas especies son más comunes en áreas urbanas donde se concentra la participación ciudadana. Por último, las diferencias entre géneros relacionados pueden ser sutiles y requerir características específicas que pueden no ser claramente visibles en fotografías de campo.

Referencias

- Pan American Health Organization (PAHO). (2025). *PAHO warns of increased risk of dengue outbreaks due to circulation of DENV-3 in the Americas*. Extraído de <https://www.paho.org/en/news/10-2-2025-paho-warns-increased-risk-dengue-outbreaks-due-circulation-denv-3-americas>
- Roiz, D., Pontifes, P.A., Jourdain, F., et al. (2024). The rising global economic costs of invasive *Aedes* mosquitoes and *Aedes*-borne diseases. *Science of The Total Environment*, 934, 173054.
- World Health Organization (WHO). (2024). Vector-borne diseases. Extraído de <https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases>
- Zhang, W.X., Zhao, T.Y., Wang, C.C., et al. (2025). Assessing the global dengue burden: Incidence, mortality, and disability trends over three decades. *PLoS Neglected Tropical Diseases*, 19(3), e0012932.