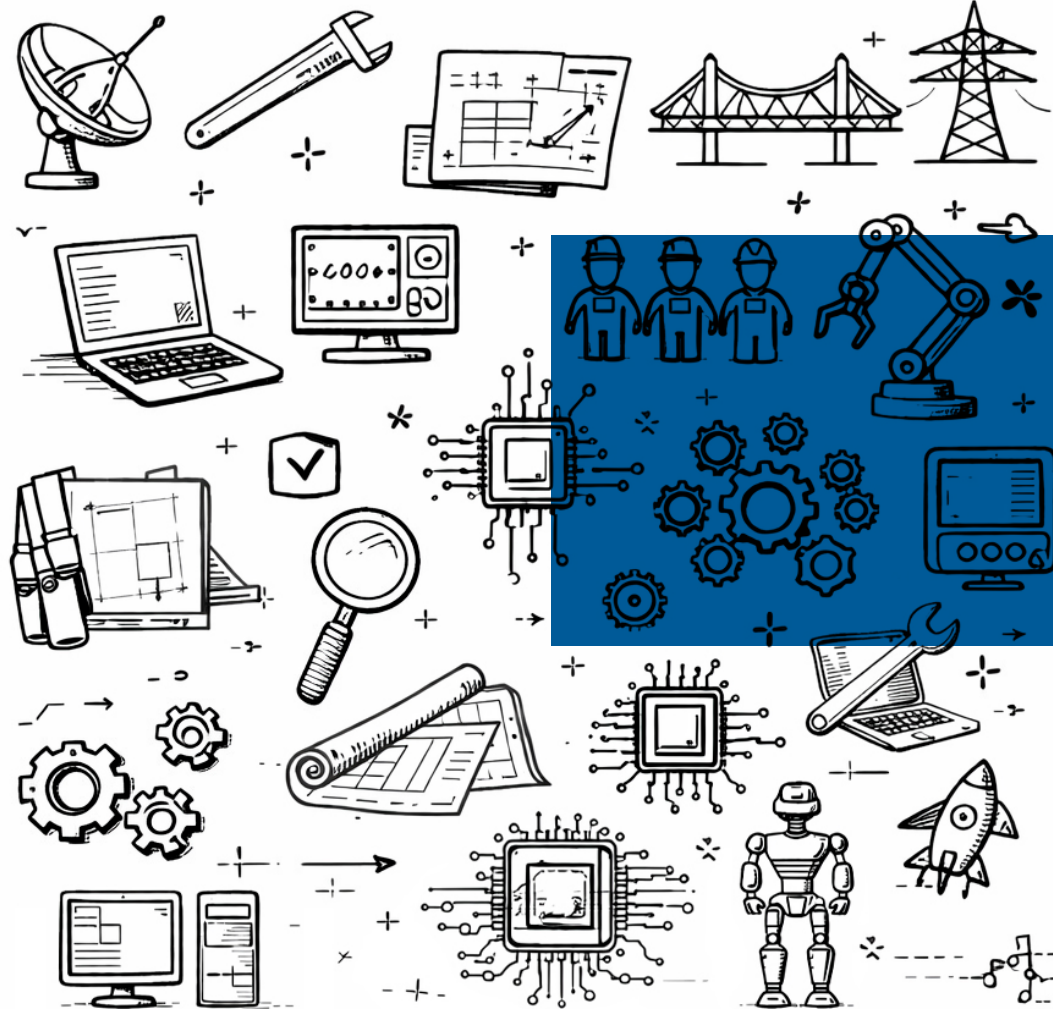




3

DESARROLLO DE ALGORITMO PARA LA CLASIFICACIÓN Y ANÁLISIS DE MOSQUITOS VECTORES DE ENFERMEDADES

DEVELOPMENT OF AN ALGORITHM FOR THE CLASSIFICATION AND ANALYSIS OF DISEASE-CARRYING MOSQUITOES

⁷ Yilber Javier Quinto Díaz

⁸ Janier Emir González Sanclemente

⁹ Haslen Palacios Bonilla

⁷ Doctor en Ciencias Matemáticas, Universidad Tecnológica del Chocó, yilber.quinto@utch.edu.co, <https://orcid.org/0000-0002-8537-0144>, https://scienti.minciencias.gov.co/cvlac/visualizador/generarCurriculoCv.do?cod_rh=0001484023.

⁸ Magister en Didáctica de las Matemáticas, Universidad Tecnológica del Chocó, d-janier.gonzales@utch.edu.co, <https://orcid.org/0009-0007-3434-9437>, https://scienti.minciencias.gov.co/cvlac/visualizador/generarCurriculoCv.do?cod_rh=0001503847.

⁹ Magister en Física Aplicada, Universidad Tecnológica del Chocó, haslen.palacios@utch.edu.co, <https://orcid.org/0000-0002-3405-7502>, https://scienti.minciencias.gov.co/cvlac/visualizador/generarCurriculoCv.do?cod_rh=0001512297.

RESUMEN

El presente trabajo tiene como objetivo desarrollar un algoritmo que clasifique de manera automática la especie y el sexo de mosquitos a partir de la señal acústica generada por el aleteo de sus alas. Para ello, se empleó una base de datos digital con registros de audio asociados a especies y sexos previamente identificados. La metodología utilizada incluyó la aplicación de la transformada de Fourier y la transformada Wavelet para la extracción de características relevantes de las señales, tanto en el dominio temporal como en el frecuencial. Posteriormente, estas características fueron procesadas mediante técnicas de aprendizaje automático, como redes neuronales artificiales, bosques aleatorios, máquinas de vectores soporte (SVM), entre otros, con el fin de realizar la clasificación supervisada de los individuos.

Los resultados, con un 75% de precisión, evidencian que las frecuencias de aleteo presentan patrones distintivos que permiten diferenciar con alta precisión tanto la especie como el sexo de los mosquitos. Esta aproximación constituye una alternativa innovadora frente a métodos tradicionales de identificación, los cuales suelen ser costosos, demandar personal especializado y prolongados tiempos de análisis. En términos de impacto, la propuesta ofrece un aporte significativo en el ámbito de la salud pública, al facilitar el monitoreo de poblaciones de mosquitos en zonas de riesgo y contribuir a la implementación de estrategias más efectivas de control vectorial. Asimismo, el algoritmo desarrollado tiene potencial de escalabilidad hacia la identificación de otros insectos de interés epidemiológico, promoviendo la transferencia tecnológica y abriendo la posibilidad de generar sistemas portátiles de detección y clasificación en campo. En conclusión, este proyecto aporta un modelo computacional innovador que combina el procesamiento de señales con técnicas de inteligencia artificial, orientado a la vigilancia epidemiológica y a la reducción del impacto de enfermedades transmitidas por vectores.

PALABRAS CLAVE: mosquitos, redes neuronales artificiales, transformada de Fourier, transformada Wavelet.

¿Cómo citar este capítulo?

En APA: Quinto, Y., González, J. y Palacios, H. (2026). Desarrollo de algoritmo para la clasificación y análisis de mosquitos vectores de enfermedades. En Bejarano-Díaz, D. y Guzmán-Brand, V. (Ed.). Ingeniería, ciencia y tecnología: perspectivas desde la investigación aplicada (pp. 55-68). Editorial TEINCO.

DOI del capítulo: <https://doi.org/10.64351/978-628-97363-5-9.cap.3>

ABSTRACT

The objective of this study is to develop an algorithm that automatically classifies the species and sex of mosquitoes based on the acoustic signal generated by the flapping of their wings. To this end, a digital database with audio recordings associated with previously identified species and sexes was used. The methodology used included the application of the Fourier transform and the wavelet transform to extract relevant characteristics from the signals, both in the time and frequency domains. These characteristics were then processed using machine learning techniques, such as artificial neural networks, random forests, support vector machines (SVM), among others, in order to perform supervised classification of the individuals.

The results, with 75% accuracy, show that wingbeat frequencies have distinctive patterns that allow for high-precision differentiation of both the species and sex of mosquitoes. This approach is an innovative alternative to traditional identification methods, which are often costly, require specialized personnel, and involve lengthy analysis times. In terms of impact, the proposal offers a significant contribution to public health by facilitating the monitoring of mosquito populations in risk areas.

and contributing to the implementation of more effective vector control strategies. Furthermore, the algorithm developed has the potential to be scaled up to identify other insects of epidemiological interest, promoting technology transfer and opening up the possibility of generating portable detection and classification systems in the field. In conclusion, this project provides an innovative computational model that combines signal processing with artificial intelligence techniques, aimed at epidemiological surveillance and reducing the impact of vector-borne diseases.

KEYWORDS: mosquitos, artificial neural networks, Fourier transform, Wavelet transform.

How to cite this chapter?

In APA: Quinto, Y., González, J. and Palacios, H. (2026). Desarrollo de algoritmo para la clasificación y análisis de mosquitos vectores de enfermedades. In Bejarano-Díaz, D. and Guzmán-Brand, V. (Ed.). Ingeniería, ciencia y tecnología: perspectivas desde la investigación aplicada (pp. 55-68). Editorial TEINCO.

Chapter DOI: <https://doi.org/10.64351/978-628-97363-5-9.cap.3>

INTRODUCCIÓN

Los insectos constituyen el grupo más grande de animales sobre la superficie terrestre, una gran cantidad de estos desempeñan un importante rol para la vida en la tierra. Por su actividad polinizante están relacionados con la producción de muchos cultivos agrícolas como frutas, hortalizas, algodón, etc., estos son los encargados de producir cera, miel, seda entre otros productos de gran valor; además, sirven de alimento a aves, peces y otros animales (Cabezas, 2012). Los mosquitos son los insectos que más enfermedades y muertes causan a los humanos en todo el mundo, estas enfermedades también afectan a otros animales como anfibios, reptiles, aves y pequeños mamíferos. Varios estudios han destacado su gran capacidad de adaptación a nuevos entornos, su rápida reproducción a gran escala, resistencia a determinados insecticidas y alteración de los hábitos alimenticios para sobrevivir a los métodos de control (Qureshi & Saeed, 2019).

Dentro de estas, el dengue es la enfermedad viral transmitida por mosquitos de la especie *Aedes aegypti*, *Aedes albopictus*, entre otros, con la mayor tasa de propagación en todo el mundo, la cual se presenta en ciclos de transmisión endémicos y epidémicos, principalmente en climas tropicales y subtropicales (Bhatt et al., 2013; OMS, 2009). La infección que causa el virus del dengue en humanos puede causar desde cuadros leves de fiebre hasta complicaciones neurológicas como meningitis, encefalitis, retinopatías, y, en ocasiones, la muerte por fuga vascular o el síndrome de shock por dengue (Kapuganti et al., 2023).

La hembra del *Aedes aegypti* es el principal vector del virus del dengue, quienes pican a los humanos ya que necesitan la sangre para madurar sus huevos. Cuando un mosquito sano pica a una persona infectada con dengue, este adquiere el virus que pasa al intestino medio del mosquito donde se replica durante un tiempo aproximado de 3 a 5 días, posteriormente, el virus tarda hasta 14 días en trasladarse hasta la glándula salival. En este punto, es cuando el mosquito se convierte en infeccioso (Guzman et al., 2010; Li et al., 2023). Una vez infectado el mosquito, este permanece con el virus por el resto de su vida, por lo que si este pica a una persona sana, le puede transmitir el virus (Li et al., 2023). Es por esto que la prevención o reducción de la propagación del virus del dengue (y otras enfermedades transmitidas por mosquitos) depende del control que se realice a los mosquitos vectores o de la interrupción del contacto entre humanos y vectores.

Estas acciones para controlar la transmisión deben enfocarse en el *Aedes aegypti* (el principal vector) en aquellos lugares donde se encuentran sus etapas inmaduras y adultas en los hogares y sus alrededores, así como en otros sitios donde se produce el contacto entre humanos y vectores. Entre los principales métodos de control está el manejo ambiental (eliminación de criaderos), el control biológico (introducción de predadores naturales y reemplazo poblacional) y el control químico (uso de insecticidas). Sin embargo, estos métodos pueden ser costosos e ineficientes si no se conoce las especies y la cantidad de mosquitos en un área determinada. Por lo tanto, la presente investigación desarrolla un algoritmo que permite identificar de forma automática la especie el sexo de mosquitos utilizando información de la frecuencia de aleteo de los mosquitos.

OBJETIVOS E HIPÓTESIS

Objetivo general

- Desarrollar un algoritmo que permita identificar de forma automática la especie y el sexo de diversos mosquitos, utilizando información de la frecuencia de aleteo de los mosquitos.

Objetivos específicos

- Buscar una base de datos con la señal digital del aleteo de los mosquitos que contenga información de la especie y el sexo de cada mosquito.
- Analizar, a través de la transformada de Fourier y la transformada wavelet, las señales del aleteo de los mosquitos y extraer sus características principales.
- Clasificar, mediante técnicas de aprendizaje automático (Machine Learning), los mosquitos por especie y sexo, con base en las características mencionadas en el objetivo específico 2.

MARCO CONCEPTUAL O TEÓRICO

Transformada de Fourier

La transformada de Fourier es una técnica ampliamente usada en el análisis y extracción de características de señales de audio. La transformada de Fourier de una función $f \in L^1(R)$ se define por $\hat{f}(w) = \int_{-\infty}^{\infty} f(t) e^{-iwt} dt$.

Esta transformada permite conocer las componentes de frecuencia que tiene una señal dada. La función f se recupera mediante $f(t) = 1/2\pi \int_{-\infty}^{\infty} \hat{f}(w) e^{iwt} dw$.

El siguiente teorema demuestra que la transformada de Fourier preserva la norma de la señal, excepto por un factor de 2π . Teorema (Identidad de Parseval). Si $f, g \in L^1(R) \cap L^2(R)$, entonces $\int_{-\infty}^{\infty} |f(t)|^2 dt = 1/2\pi \int_{-\infty}^{\infty} |\hat{f}(w)|^2 dw$, equivalentemente, $\|f\|^2 = 1/2\pi \|\hat{f}\|^2$.

Transformada wavelet

La transformada wavelet surge en el análisis de señales para superar las limitaciones en resolución tiempo frecuencia de la transformada de Fourier. Dada una función $\psi \in L^2(R)$, la transformada wavelet de una función $f \in L^2(R)$ se define como $(W_{\psi}f)(a,b) = \langle f, \psi_{(a,b)} \rangle = 1/\sqrt{a} \int_{-\infty}^{\infty} f(t) \psi((t-b)/a) dt$, donde $a, b \in R$, $a > 0$, el factor de escala a , el factor de traslación b son variables continuas, la función $\psi_{(a,b)}$ es definida por $\psi_{(a,b)} = 1/\sqrt{a} \psi((t-b)/a)$, y la función ψ es conocida como wavelet madre y debe cumplir las siguientes condiciones:

- I. Promedio cero

$$\int_{-\infty}^{\infty} \psi(t) dt = 0.$$

- II. Condición de admisibilidad

$$C_{\psi} = \int_0^{\infty} \frac{|\hat{\psi}(s)|^2}{s} ds < \infty.$$

Esto significa que $\psi(t)$ es oscilatoria y tiene área cero (Daubechies, 1992).

Máquina de vectores soporte

Las máquinas de vectores soporte (SVM) son un conjunto de algoritmos de aprendizaje supervisado desarrollado por Vladimir Vapnik. Estos métodos son ampliamente utilizados en la identificación de patrones. Así, dado un conjunto de ejemplos de entrenamiento (muestras) se puede etiquetar las clases y entrenar una SVM para construir un modelo que prediga la clase de una nueva muestra (Alba, 2018). De forma general, se considera el problema de clasificación binaria, con datos de entrenamiento dados por $\{(x_i, d_i) : 1 \leq i \leq N\}$, donde x_i es el patrón de entrada, para el ejemplo i -ésimo, y d_i es la respuesta deseada. Si suponemos que el patrón (clase) representado por el subconjunto $d_i = 1$ y el patrón representado por el subconjunto $d_i = -1$ son linealmente separables, entonces la ecuación de una superficie de decisión en forma de un hiperplano que realiza la separación es dada por $w^T x + b = 0$, donde x es un vector de entrada, w es un vector de peso ajustable, y b es un sesgo. Por lo tanto, podemos escribir $w^T x_i + b \geq 0$ para $d_i = 1$ / $w^T x_i + b < 0$ para $d_i = -1$.

Para un vector de peso dado w y sesgo b , la separación entre el hiperplano y el punto de datos más cercano se denomina margen de separación, indicado por ρ . El objetivo de una máquina de vectores de soporte es encontrar el hiperplano particular para el cual se maximiza el margen de separación, ρ . Bajo esta condición, la superficie de decisión se conoce como el hiperplano óptimo (Haykin, 2009).

MATERIALES Y MÉTODOS

Base de datos

La base datos utilizada en esta investigación cuenta con información de la especie, sexo y señal de aleteo de cuatro especies de mosquitos y dos especies de moscas. Las señales de aleteo se captaron con un sensor láser de baja potencia que apunta hacia una placa de fototransistores, luego se almacenaron en formato de audio.

Tabla 4. Clases de las señales de aleteo

Clase	Abreviatura	Muestras
Aedes aegypti (female)	Ae. aegypti.f	5.000
Aedes aegypti (male)	Ae. aegypti.m	5.000
Culex quinquefasciatus (female)	Cx. quinq.f	5.000
Culex quinquefasciatus (male)	Cx. quinq.m	5.000
Culex stigmatosoma (female)	Cx. stigma.f	5.000
Culex stigmatosoma (male)	Cx. stigma.m	5.000
Culex tarsalis (female)	Cx. tarsalis.f	5.000
Culex tarsalis (male)	Cx. tarsalis.m	5.000
Drosophila simulans	D. simulans	5.000
Musca domestica	M. doméstica	5.000
Total		50.000

Nota. Elaboración propia.

Técnicas de extracción de características

Las técnicas de extracción de características que se usaron en esta investigación se basan en la teoría de la transformada de Fourier y la transformada wavelet; estos métodos han sido ampliamente aplicados en tareas de reconocimiento de señales de audio y voz.

Características en el dominio de frecuencias. La extracción de características en el dominio de frecuencias de una señal se basa en el análisis de la frecuencia fundamental y de las componentes de frecuencia llamadas armónicos. Sea f la señal de aleteo de los insectos, $\hat{f}$ la transformada de Fourier de f , w_0 la frecuencia fundamental y w_k la frecuencia del k -ésimo armónico. A continuación, describimos algunas de las características extraídas del espectro de la señal.

Inarmonicidad. Suma de errores relativos en la estimación de las frecuencias de cada uno de los armónicos de una señal

$$Inharm = \sum_{k=1}^{P-1} \frac{|w_k - kw_0|}{kw_0}$$

Irregularidad espectral. Representa la irregularidad en la forma del espectro de una señal

$$SI = \sum_{k=1}^{P-1} \left| \tilde{f}(k) - \frac{\tilde{f}(k-1) + \tilde{f}(k) + \tilde{f}(k+1)}{3} \right|,$$

donde

$$\tilde{f}(k) = 20 \log \log (|\tilde{f}(w_k)|).$$

Triestímulo. El método es capaz de distinguir entre notas de diferentes instrumentos musicales utilizando características relacionadas con la magnitud de los armónicos

$$ts1 = \frac{\hat{f}(w_0)}{\sum_{k=0}^{P-1} |\hat{f}(w_k)|},$$

$$ts2 = \frac{\hat{f}(w_1) + \hat{f}(w_2) + \hat{f}(w_3)}{\sum_{k=0}^{P-1} |\hat{f}(w_k)|},$$

$$ts3 = \frac{\sum_{k=4}^{P-1} |\hat{f}(w_k)|}{\sum_{k=0}^{P-1} |\hat{f}(w_k)|}.$$

Punto espectral roll-off. Se define como la mínima frecuencia en la que se encuentra el 95% de la energía total del espectro de potencia $rolloff = r$, siempre que

$$\sum_{k=0}^r |\hat{f}(k)|^2 = 0.95 \sum_{k=1}^{N-1} |\hat{f}(k)|^2.$$

Otras características del dominio de la frecuencia que se utilizaron se mencionan en la Tabla 5. Además de un conjunto de características clásicas del dominio del tiempo.

Tabla 5. Características en el dominio del tiempo y de la frecuencia

Dominio	Característica
Tiempo	<i>Root mean square; short-time energy; zero-crossing rate; excursión de pico; complejidad estimada; centroide temporal; media; desviación estándar; varianza; asimetría (skewness) y kurtosis.</i>
Frecuencia	Frecuencia fundamental; inarmonicidad; triestímulo; irregularidad espectral; flujo espectral; <i>roll-off flatness</i> ; centroide espectral; media; desviación estándar; varianza; asimetría (<i>skewness</i>) y kurtosis.

Nota. Elaboración propia.

Características en representación de wavelets ortogonales. Sea $f \in L^2(R)$ una función suave, y sea ψ una wavelet de valor real, ya que ψ tiene promedio cero, la integral

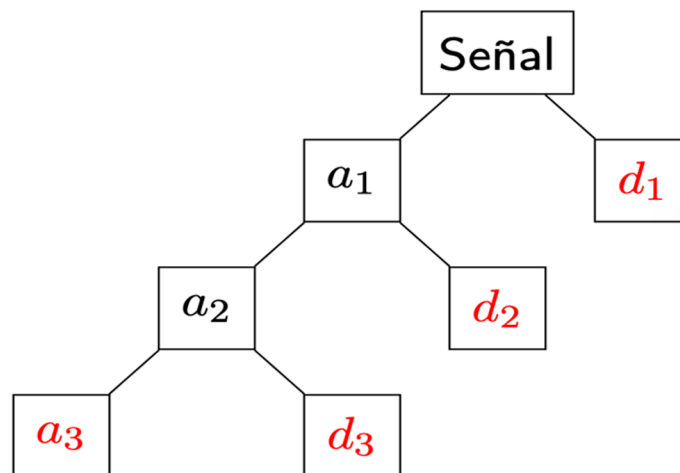
$$d_j(n) = \langle f, \psi_{j,n} \rangle = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-2^j n}{2^j}\right) dt,$$

mide la variación de f en una vecindad de $2^j n$ proporcional a la escala 2^j . Cuando la escala tiende a cero, el decaimiento de los coeficientes wavelet d_j caracteriza la regularidad de f en la vecindad de $2^j n$ (Daubechies, 1990).

Esto se utiliza a menudo para detectar cambios bruscos en las señales (Zhang & Yan, 2001), eliminación de ruido (Baldazzi et al., 2020), y en tareas de reconocimiento del habla (Koduru et al., 2020). El método de extracción de características utilizado en esta investigación se basa en el método propuesto por (Mini et al., 2021) para el reconocimiento de señales de voz, que consiste en descomponer la señal original en coeficientes de aproximación a_j y detalle d_j mediante el algoritmo de la DWT y dos versiones de este denominados WPD y DWPD.

El algoritmo de descomposición DWT descompone la señal en diferentes niveles llamados niveles de descomposición. En el primer nivel, se obtienen los coeficientes de aproximación y detalle calculando la convolución de la señal f y dos filtros paso-bajos $\tilde{h}$ y paso-altos $\tilde{g}$, respectivamente. El proceso se repite de manera sucesiva tomando la convolución de los coeficientes de aproximación a_j y los filtros hasta llegar al nivel de descomposición deseado. Con este algoritmo, se construyó un vector de características concatenando la sucesión de coeficientes de detalle d_j de los primeros seis niveles de descomposición y el coeficientes de aproximación a_6 . En las Figuras 6 y 4 se muestra un ejemplo para el caso de tres niveles de descomposición.

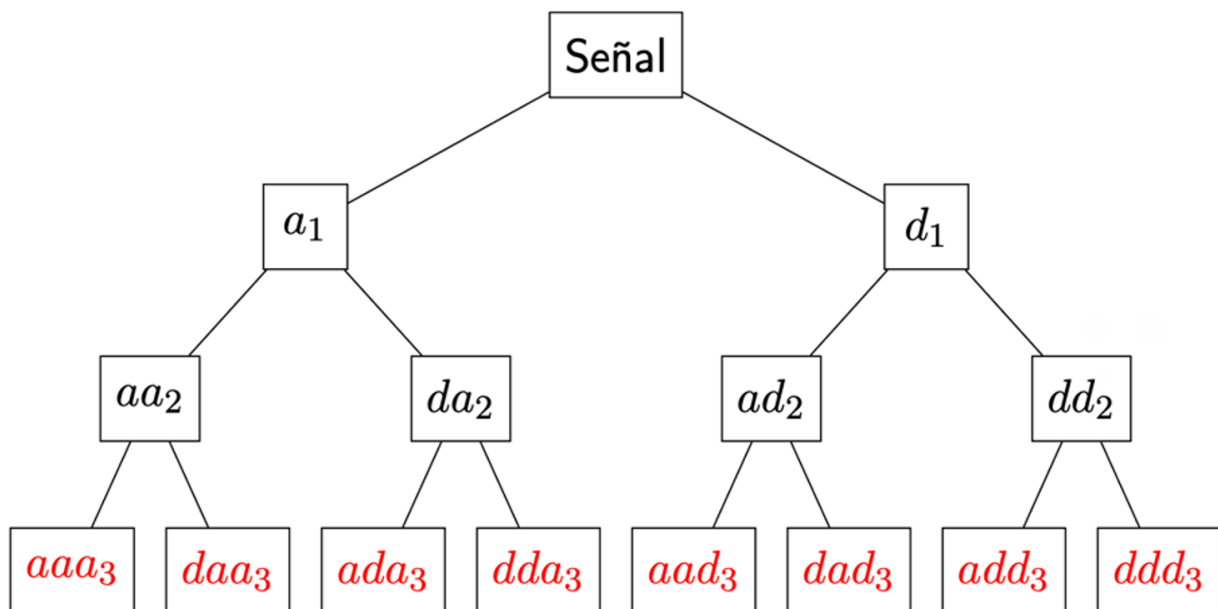
Figura 19. Ejemplo de método DWT para el caso de 3 niveles de descomposición



Nota. Tomado de Mini et al., 2021.

La WPD (*Wavelet Packets Decomposition*) es una generalización del algoritmo DWT, que no solo descompone los coeficientes de aproximación, sino que en cada nivel de descomposición aplica el algoritmo DWT a los coeficientes de detalle (Coifman & Wickerhauser, 1992), lo cual ofrece un análisis más robusto del contenido de frecuencia de la señal. El vector de características se obtuvo concatenando todas las sucesiones finales de coeficientes de aproximación y detalle en el sexto nivel de descomposición (ver Figura 20).

Figura 20. Ejemplo del método WPD para el caso de 3 niveles de descomposición

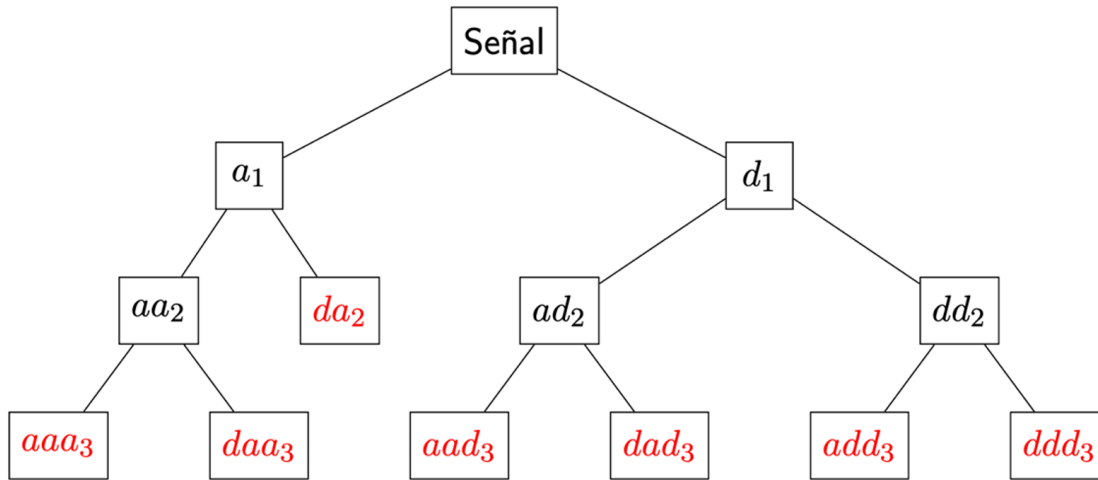


Nota. Tomado de Mini et al., 2021.

El método DWPD es un algoritmo mixto que descompone los coeficientes a_1 mediante el método DWT, mientras que los coeficientes d_1 se descomponen a través del método WPD. El vector de características resultante es una concatenación de

los coeficientes detalle de cada nivel de descomposición y el vector de coeficientes de aproximación final de la rama de descomposición DWT de a_1 , junto con los coeficientes finales de la rama de descomposición WPD de d_1 (ver Figura 21).

Figura 21. Ejemplo del método DWPD para el caso de 3 niveles de descomposición



Nota. Tomado de Mini et al., 2021.

Modelos de clasificación

Existe una gran variedad de modelos de aprendizaje automático que se utilizan en la clasificación de patrones, la mayoría de estos modelos cuentan con hiper parámetros que deben ser ajustados para garantizar la óptima generalización del modelo. Los modelos que se usaron en esta investigación se describen en la Tabla 6, los cuales se eligieron debido a que son los que mejor rendimiento promedio tuvieron, sin ningún ajuste de hiper parámetros, entre los 30 modelos entrenados con el módulo de Python Lazy Predict.

La fase de entrenamiento y selección de hiper parámetros se realizó mediante validación cruzada k-Fold, con $K=5$, y una división de datos de 80% para entrenamiento y 20% para prueba. Los modelos MLP y SVM suelen funcionar mejor con datos estandarizados, lo cual requiere estandarizar cada característica restando el promedio a cada entrada del vector de característica y escalando a la varianza unitaria.

Tabla 6. Modelos de aprendizaje automático y su rango de hiper parámetros

Modelo	Hiper parámetros
MLP	Capas ocultas: 2 Número de neuronas: 64 32,128 64 y 256 128
LGBMClassifier	Número de árboles $N = 10, 20, \dots, 400$
XGBClassifier	Número de árboles $N = 10, 20, \dots, 400$

SVM RBF	$C = 10^i; i = -6, -5, \dots, 5$ $\gamma = 10^k; k = -4, -3, \dots, 0$
Random Forest	Número de árboles $N = 5, 10, \dots, 80$

Nota. Elaboración propia.

HALLAZGOS O RESULTADOS

La extracción de características se realizó en MATLAB R2022a, ejecutado en un MacBook Pro, 13-inch, Mid 2012, el análisis exploratorio se hizo en JupyterLab 4.0.11 y el entrenamiento de los modelos de clasificación en Google Colab Pro. Los resultados de precisión de los modelos de clasificación para cada conjunto de características se muestran en la Tabla 7. En general, el MLP obtuvo un rendimiento superior a los otros modelos en cada conjunto de características; el mejor resultado (71,68%) se obtuvo sobre las características MFCC. Los conjuntos de características que se extraen de la señal de aleteo aplicando la transformada wavelet (DWT, WPD y DWPD) no superaron el 47,53% de precisión. Las mayores tasas de confusión en la clasificación se presentan entre las especies de moscas y las especies de mosquitos, la tasa de confusión por sexo entre individuos de la misma especie de mosquitos es relativamente baja a excepción del *Culex tarsalis* (ver Figura 22).

Tabla 7. Resultados de precisión por conjunto de característica, modelo e hiper parámetro

Característica	Modelo	Hiper parámetros	Precisión (%)
Temporal	MLP	128/64	42,17
	LGBMClassifier	N = 90	36,26
	XGBClassifier	N = 150	34,67
	SVM RBF	$C = 10^5, \gamma = 10^{-3}$	36,81
	Random Forest	N = 15	29,58
Espectral	MLP	256/128	65,85
	LGBMClassifier	N = 90	64,58
	XGBClassifier	N = 150	62,96
	SVM RBF	$C = 10, \gamma = 10^{-2}$	62,96
	Random Forest	N = 30	57,45
DWT	MLP	256/128	47,53
	LGBMClassifier	N = 50	33,05
	XGBClassifier	N = 90	36,52
	SVM RBF	$C = 10^2, \gamma = 10^{-2}$	43,59
	Random Forest	N = 30	33,22
WPD	MLP	128/64	37,18
	LGBMClassifier	N = 70	33,82
	XGBClassifier	N = 170	33,55
	SVM RBF	$C = 10^3, \gamma = 10^{-2}$	36,13
	Random Forest	N = 20	29,97

DWPD	MLP	128/64	47,02
	LGBMClassifier	N = 50	38,99
	XGBClassifier	N = 110	37,70
	SVM RBF	$C = 10^2, \gamma = 10^{-2}$	43,42
	Random Forest	N = 30	32,70
MFCC	MLP	256/128	71,68
	LGBMClassifier	N = 50	66,44
	XGBClassifier	N = 90	64,53
	SVM RBF	$C = 1, \gamma = 10^{-3}$	67,85
	Random Forest	N = 30	58,81

Nota. Elaboración propia.

Figura 22. Matriz de confusión (en porcentaje) obtenida del modelo MLP y el conjunto de características MFCC

Clase Verdadera	Ae. aegypti.f	63.40	0.30	20.10	0.30	9.50	2.80	0.40	0.70	2.00	0.50
	Ae. aegypti.m	0.00	87.50	0.10	8.10	0.10	0.10	0.40	1.70	1.30	0.70
	D. simulans	16.80	0.60	67.50	0.10	12.00	0.10	0.50	0.10	1.60	0.70
	M. domestica	0.00	7.00	0.30	60.30	0.70	3.80	0.20	23.30	2.90	1.50
	Cx. quinq.f	6.60	0.50	12.30	0.80	64.40	0.20	6.80	0.70	4.00	3.70
	Cx. quinq.m	1.40	0.70	0.10	2.40	0.20	83.20	0.10	9.00	1.40	1.50
	Cx. stigma.f	0.90	0.50	0.80	0.60	8.90	0.60	78.70	0.60	4.50	3.90
	Cx. stigma.m	0.30	4.20	0.10	21.52	0.50	11.11	0.50	57.86	2.00	1.90
	Cx. tarsalis.f	1.20	0.30	0.90	1.10	3.10	0.70	2.70	0.90	76.30	12.80
	Cx. tarsalis.m	0.20	0.50	0.10	0.80	3.30	0.20	3.30	1.00	13.00	77.60
		Ae. aegypti.f	Ae. aegypti.m	D. simulans	M. domestica	Cx. quinq.f	Cx. quinq.m	Cx. stigma.f	Cx. stigma.m	Cx. tarsalis.f	Cx. tarsalis.m
		Predicción									

Nota. Elaboración propia.

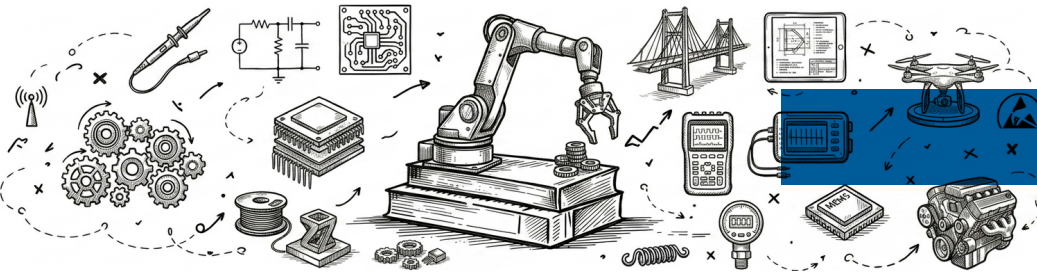
CONCLUSIONES

En la presente investigación se diseñaron diversos algoritmos para la clasificación automática de mosquitos, según la especie y el sexo, con base en el análisis de características extraídas del aleteo de los mosquitos. Con base en estos datos se estudiaron técnicas clásicas de extracción de características basadas en la transformada de Fourier y se compararon

con técnicas más recientes basadas en la Transformada wavelet. Se utilizó un enfoque de representación de características como vectores de R^n , unidas a diferentes técnicas de clasificación automática.

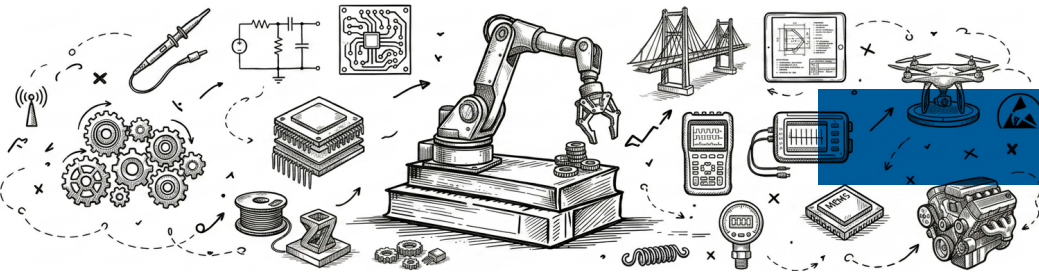
Se extrajeron seis conjuntos de características en R^n : uno en el dominio del tiempo, un conjunto en el dominio de la frecuencia, tres conjuntos utilizando la transformada Wavelet y un conjunto en representación cepstrum. Se utilizaron cinco modelos de clasificación para las características en R^n . El mejor resultado en la clasificación (71,68% de precisión) se obtuvo con MFCC y MLP.

Los resultados obtenidos en esta investigación destacan la importancia de utilizar técnicas tradicionales de extracción de características como la transformada de Fourier para mejorar la precisión en la identificación de la especie y sexo de mosquitos. En trabajos futuros, se podría entrenar modelos de redes neuronales convolucionales y características en imágenes, como el espectrograma y el escalograma para la clasificación de mosquitos, haciendo uso de información adicional de los mosquitos como la actividad de ritmos circadianos, para diseñar y desarrollar un dispositivo de bajo costo que implemente el algoritmo propuesto para la clasificación automática de mosquitos.



REFERENCIAS

- Alba Castro, J. L. (2018). *Máquinas de vectores de soporte (SVM)*. <http://www.gts.tsc.uvigo.es/~jalba/doctorado/SVM.pdf>
- Baldazzi, G., Solinas, G., Del Valle, J., Barbaro, M., Micera, S., Raffo, L., & Pani, D. (2020). Systematic analysis of wavelet denoising methods for neural signal processing. *Journal of Neural Engineering*, 17(6). <https://doi.org/10.1088/1741-2552/abc741>
- Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., Drake, J. M., Brownstein, J. S., Hoen, A. G., Sankoh, O., Myers, M. F., George, D. B., Jaenisch, T., William Wint, G. R., Simmons, C. P., Scott, T. W., Farrar, J. J., & Hay, S. I. (2013). The global distribution and burden of dengue. *Nature*, 496(7446), 504–507. <https://doi.org/10.1038/nature12060>
- Cabezas, F. (2012). *Introducción a la entomología* (1.ª ed.). Trillas.
- Coifman, R. R., & Wickerhauser, M. V. (1992). Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2), 713–718. <https://doi.org/10.1109/18.119732>
- Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36(5), 961–1005. <https://doi.org/10.1109/18.57199>
- Daubechies, I. (1992). *Ten lectures on wavelets*. SIAM.
- Guzman, M. G., Halstead, S. B., Artsob, H., Buchy, P., Farrar, J., Gubler, D. J., Hunsperger, E., Kroeger, A., Margolis, H. S., Martínez, E., Nathan, M. B., Pelegrino, J. L., Simmons, C., Yoksan, S., & Peeling, R. W. (2010). Dengue: A continuing global threat. *Nature Reviews Microbiology*, 8(12 Suppl.), S7–S16. <https://doi.org/10.1038/nrmicro2460>
- Haykin, S. (2009). *Neural networks and learning machines* (3rd ed.). Pearson Education.
- Kapuganti, S. K., Saumya, K. U., Verma, D., & Giri, R. (2023). Investigating the aggregation perspective of Dengue virus proteome. *Virology*, 586, 109–118. <https://doi.org/10.1016/j.virol.2023.07.010>
- Koduru, A., Valiveti, H. B., & Budati, A. K. (2020). Feature extraction algorithms to improve the speech emotion recognition rate. *International Journal of Speech Technology*, 23(1), 45–55. <https://doi.org/10.1007/s10772-020-09672-4>



REFERENCIAS

- Li, H. H., Su, M. P., Wu, S. C., Tsou, H. H., Chang, M. C., Cheng, Y. C., Tsai, K. N., Wang, H. W., Chen, G. H., Tang, C. K., Chung, P. J., Tsai, W. T., Huang, L. R., Yueh, Y. A., Chen, H. W., Pan, C. Y., Akbari, O. S., Chang, H. H., Yu, G. Y., ... Chen, C. H. (2023). Mechanical transmission of dengue virus by *Aedes aegypti* may influence disease transmission dynamics during outbreaks. *EBioMedicine*, 94, 104723. <https://doi.org/10.1016/j.ebiom.2023.104723>
- Mini, P. P., Thomas, T., & Gopikakumari, R. (2021). Wavelet feature selection of audio and imagined/vocalized EEG signals for ANN based multimodal ASR system. *Biomedical Signal Processing and Control*, 63, 102218. <https://doi.org/10.1016/j.bspc.2020.102218>
- Organización Mundial de la Salud, & TDR para la Investigación sobre Enfermedades Tropicales. (2009). *Dengue: Guías para el diagnóstico, tratamiento, prevención y control*. World Health Organization. https://www.who.int/denguecontrol/resources/dengue_guidelines_2009/es/
- Qureshi, A. I., & Saeed, O. (2019). *Dengue virus disease: From origin to outbreak*. Academic Press. <https://doi.org/10.1016/C2018-0-01619-9>

