

Modelo híbrido de programación genética y redes neuronales para el reconocimiento de emociones

Alvaro A Colunga-Rodriguez¹, Alicia Martínez-Rebollar^{1,*}, Hugo Estrada-Esquivel¹,
Eddie Clemente¹

¹ Tecnológico Nacional de México / CENIDET, Morelos,
México

{d21ce089, alicia.mr, hugo.ee, eddie.ct}@cenidet.tecnm.mx

Resumen. El reconocimiento automático de emociones es un área clave de la computación afectiva, con aplicaciones en interacción humano-computadora, salud mental y personalización de sistemas inteligentes. En este contexto, las señales fisiológicas, como el ritmo cardíaco, han cobrado relevancia debido a su relación con el estado emocional y la posibilidad de ser registradas de manera no invasiva mediante dispositivos portátiles. Este estudio propone un modelo híbrido basado en Programación Genética y Redes Neuronales para el reconocimiento de emociones a partir del ritmo cardíaco. Se llevó a cabo un experimento controlado para recopilar datos de cuatro emociones: calma, enojo, felicidad y tristeza. Tras el preprocesamiento con un filtro de media móvil y la extracción de características, la Programación Genética transformó los datos en un espacio más adecuado para la clasificación mediante una red neuronal multicapa. El modelo alcanzó una precisión del 95% en entrenamiento y 94% en prueba, aunque se observaron dificultades en la diferenciación entre enojo y tristeza. Estos resultados demuestran la viabilidad del uso de dispositivos portátiles y metodologías híbridas para la detección de emociones, contribuyendo al desarrollo de la computación afectiva.

Palabras clave. Reconocimiento de emociones, Ritmo Cardíaco, Programación Genética, Redes Neuronales, Dispositivos inteligentes, Clasificación de emociones.

Hybrid Model of Genetic Programming and Neural Networks for Emotion Recognition

Abstract. Automatic emotion recognition is a key area of affective computing, with applications in human-computer interaction, mental health, and intelligent system personalization. In this context, physiological

signals such as heart rate have gained relevance due to their relationship with emotional states and the possibility of being recorded non-invasively through wearable devices. This study proposes a hybrid model based on Genetic Programming and Neural Networks for emotion recognition from heart rate data. A controlled experiment was conducted to collect data corresponding to four emotions: calmness, anger, happiness, and sadness. After preprocessing with a moving average filter and feature extraction, Genetic Programming transformed the data into a more suitable space for classification through a multilayer neural network. The model achieved an accuracy of 95% in training and 94% in testing, although difficulties were observed in distinguishing between anger and sadness. These results demonstrate the feasibility of using wearable devices and hybrid methodologies for emotion detection, contributing to the advancement of affective computing.

Keywords. Emotion recognition, heart rate, genetic programming, neural networks, smart devices, emotion classification.

1. Introducción

Las emociones desempeñan un papel fundamental en la vida cotidiana, ya que cumplen una función adaptativa que permite a las personas responder de manera adecuada a diversas situaciones. Sin embargo, emociones negativas como el enojo, el miedo o la depresión pueden afectar significativamente la salud física y mental de un individuo [8]. En este contexto, el reconocimiento automático de emociones ha cobrado gran relevancia en diversas áreas, como la salud mental, la interacción humano-computadora y la personalización de sistemas inteligentes.

La computación afectiva [9] es un campo de estudio que busca dotar a los sistemas computacionales de la capacidad de reconocer, interpretar y responder a las emociones humanas. Para ello, se utilizan diferentes fuentes de información, como expresiones faciales, señales de electroencefalografía (EEG), electrocardiografía (ECG) y respuesta galvánica de la piel (GSR), entre otras [14]. No obstante, la adquisición de estos datos suele requerir equipos especializados y procedimientos invasivos, lo que limita su aplicación en entornos cotidianos. El ritmo cardíaco es una señal fisiológica estrechamente relacionada con el estado emocional de una persona [6].

A diferencia de otras métricas fisiológicas, el ritmo cardíaco puede obtenerse de manera no invasiva mediante dispositivos como relojes inteligentes o sensores, los cuales permiten registrar de forma continua esta variable. Dispositivos como el Polar OH1 registran la frecuencia cardíaca con una precisión muy alta, mostrando una correlación de 0.99 respecto a un electrocardiograma (ECG) estándar [18]. Esto abre la posibilidad de desarrollar modelos de detección de emociones basados en datos fisiológicos accesibles y de fácil adquisición.

En este trabajo, se presenta un modelo híbrido basado en Programación Genética y Redes Neuronales para el reconocimiento de emociones. La metodología propuesta combina Programación Genética para la generación de funciones no lineales que optimizan la representación de los datos, y Redes Neuronales para la clasificación de emociones. Para evaluar el modelo, se llevó a cabo un experimento controlado en el que se recopilaban datos de ritmo cardíaco asociados a cuatro emociones: calma, enojo, felicidad y tristeza. Los resultados obtenidos demuestran que el enfoque híbrido propuesto permite identificar estos estados emocionales con una efectividad del 95% con datos de entrenamiento y 94% prueba.

El resto del artículo se organiza de la siguiente manera: en la Sección 2 se presentan los trabajos relacionados con la detección de emociones mediante el ritmo cardíaco. La Sección 3 describe la metodología utilizada para la construcción del modelo híbrido basado en Programación Genética y Redes Neuronales. En la Sección 4 se exponen los resultados experimentales obtenidos.

Finalmente, la Sección 5 presenta las conclusiones y las posibles líneas de investigación futura.

2. Trabajos relacionados

El uso de relojes inteligentes para la detección de emociones ha crecido significativamente en los últimos años debido a su capacidad para recopilar datos fisiológicos de manera continua y no invasiva. Entre las variables más utilizadas en estos estudios destacan el ritmo cardíaco (HR), la respuesta galvánica de la piel (GSR) y la temperatura de la piel (SKT), las cuales han sido ampliamente exploradas en el contexto del reconocimiento de emociones [15]. En esta sección se presentan algunos trabajos de investigación relevantes en la detección de emociones.

El estudio de Pepa [7] emplea un reloj inteligente para monitorear las emociones en pacientes con enfermedad de Parkinson, utilizando datos obtenidos de una banda inteligente en un entorno controlado. Las emociones se clasifican en términos de valencia y activación, siguiendo el modelo de Russell [10], alcanzando una precisión del 93.4% en valencia y 78.6% en activación, una desventaja de este trabajo es que se aplica a un grupo específico de pacientes que por sus características puede no generalizarse. Por otro lado, Francisti [2] analiza el impacto de las emociones en el rendimiento académico durante la pandemia de COVID-19.

Aunque su trabajo no realiza una clasificación explícita de emociones, sus hallazgos sugieren que los cambios en la activación emocional, medidos a través del ritmo cardíaco, están relacionados con el desempeño de los estudiantes. Este estudio refuerza la viabilidad del uso de dispositivos inteligentes para el monitoreo emocional, sin embargo, no realiza una clasificación de emociones.

Otras propuestas de investigación han explorado la clasificación de emociones empleando datos obtenidos de sensores de relojes inteligentes. Wang [13] propone un sistema adaptativo para la detección de felicidad, tristeza, enojo, miedo y estado neutral, utilizando volumen sanguíneo, actividad electrodermal y temperatura

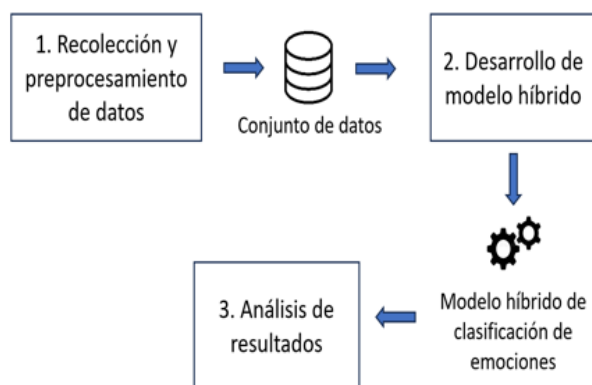


Fig. 1. Metodología de desarrollo de modelo híbrido

de la piel, con una efectividad del 74.3%. La principal fortaleza de este estudio es su enfoque multimodal, que combina varias señales fisiológicas y permite abordar una gama amplia de emociones básicas. Sin embargo, presenta una efectividad moderada y utiliza sensores que pueden ser invasivos.

En otro trabajo, Shu [11] desarrolla un modelo basado en el ritmo cardíaco, logrando una precisión del 84% en la clasificación de neutral, felicidad y tristeza. El aporte de este trabajo es su eficiencia en la clasificación de emociones a pesar de limitarse a solo tres.

Takeshita [12] enfoca su investigación en la detección de miedo y no miedo a partir de datos obtenidos de un reloj inteligente durante la visualización de videos de diferentes géneros, alcanzando una efectividad del 90%. Este trabajo destaca su alta efectividad en la detección de una emoción específica y por su diseño experimental que simula de manera efectiva reacciones emocionales reales mediante estímulos audiovisuales. Sin embargo, se limita a solo una emoción.

Estos estudios destacan la importancia de los dispositivos inteligentes en la detección de emociones y subrayan las diferencias en las métricas de precisión según las emociones evaluadas y las metodologías utilizadas. El modelo propuesto en esta investigación busca optimizar el proceso de reconocimiento de emociones, combinando la capacidad adaptativa de la Programación Genética con la robustez en el aprendizaje de patrones de las Redes Neuronales

en la detección de emociones a partir de señales fisiológicas.

3. Desarrollo del modelo híbrido

El desarrollo del modelo de detección de emociones de una persona utilizando programación genética y redes neuronales se muestra en esta sección como se puede ver en la metodología de la Figura 1. La primera parte llamada recolección y preprocesamiento de datos describe cómo se obtuvieron los datos para el modelo mediante un experimento controlado. La segunda parte denominada desarrollo del modelo híbrido muestra cómo se utilizó la técnica de programación genética y la red neuronal para desarrollar el modelo de clasificación de emociones, esta parte se desarrolla en la sección 4. La tercera parte, análisis de resultados muestra los resultados obtenidos por el modelo y se presenta en la sección 5.

3.1. Recolección y preprocesamiento de datos

La recolección de datos para el desarrollo del modelo de clasificación de emociones se realizó a través de un experimento controlado con estudiantes del Tecnológico Nacional de México, campus Ensenada. La muestra estuvo conformada por 11 participantes, 8 hombres y 3 mujeres, con edades entre 21 y 39 años ($M = 24.55$, $SD = 5.76$). La participación fue voluntaria, y antes del experimento, se les explicó el objetivo del estudio y su estructura. La limitación en el tamaño de la muestra se debió principalmente a restricciones de tiempo disponibles para la implementación del experimento.

A pesar de la muestra reducida, los datos obtenidos permiten una evaluación inicial del desempeño del modelo y sientan las bases para estudios posteriores con poblaciones más amplias. No se aplicó ningún test médico o psicológico previo para descartar condiciones específicas como enfermedades cardíacas, trastornos relacionados con el estrés u otras variables fisiológicas o emocionales que pudieran influir en los resultados. Debido al carácter exploratorio del estudio y a las limitaciones de tiempo disponibles para la recolección de datos, la

En general la emoción que percibí durante el proceso fué:

☐ Calma ☐ Felicidad ☐ Tristeza ☐ Enojo

La intensidad de la emoción que percibí fué:

☐ Baja ☐ Media ☐ Alta

En estos momentos la emoción que siento es:

☐ Calma ☐ Felicidad ☐ Tristeza ☐ Enojo

La intensidad de la emoción que estoy percibiendo es:

☐ Baja ☐ Media ☐ Alta

Fig. 2. Auto reporte de emociones

participación fue abierta a estudiantes voluntarios sin un proceso de selección adicional.

El experimento se llevó a cabo en un entorno privado y sin distracciones para garantizar la validez de los datos. El diseño experimental contempló cuatro etapas de inducción emocional: (i) Calma, (ii) Enojo, (iii) Tristeza y (iv) Felicidad. Los procesos tuvieron un espacio de 5 minutos entre cada uno. Para inducir las emociones, se utilizaron estímulos audiovisuales, incluyendo imágenes del conjunto de datos IAPS [4], sonidos del conjunto IADS [1] y música seleccionada [3]. Después de cada proceso los participantes contestaron el auto reporte mostrado en la Figura 2 para indicar la emoción y su intensidad. Esto nos permitió validar el estado emocional del participante.

Cada etapa del experimento se realizó en dos fases, con la siguiente secuencia:

- En la primera fase, los participantes visualizaron un video diseñado para inducir un estado de calma, con el objetivo de neutralizar cualquier emoción previa y estandarizar el punto de partida.
- En la segunda fase, se les presentó un video correspondiente a la emoción específica que se deseaba inducir, siguiendo el protocolo definido para cada etapa.

Los videos fueron elaborados utilizando imágenes del conjunto IAPS [4], sonidos del conjunto IADS [1] y música [3]. Estas imágenes nos fueron proporcionadas exclusivamente para fines de investigación y no tenemos permitido difundirlas en algún medio. Por esta razón, solo se indican los identificadores de cada imagen para su referencia. Los videos de calma contienen

imágenes de paisajes, naturaleza y objetos. El video de enojo contiene imágenes de violencia, maltrato animal entre otros. El video de felicidad contiene imágenes de personas sonriendo y situaciones agradables. El video de tristeza contiene imágenes de desastres naturales, injusticias y duelo.

Video Calma 1

Imágenes IAPS:

5001, 5020, 5030, 5199, 5410, 7490, 2620, 2580, 5720, 1450, 5726, 5530, 5390, 7100, 2320, 7053, 5010, 2870, 5711, 2980.

Audio: Jean Michael Jarre – Oxygene, Pt. 4

Video Calma 2

Imágenes IAPS:

2515, 7283, 2380, 5781, 7495, 2791, 5201, 7242, 7285, 2594, 4605, 7190, 2392, 2579, 2374, 5631, 7237, 2384, 2488, 5220.

Audio: Dave Brubeck – Take Five.

Video Enojo

Imágenes IAPS:

6000, 9415, 4621, 2220, 6350, 6315, 6360, 6840, 9040, 6838, 9150, 2751, 2745, 9295, 9402, 9427, 9560, 9520, 6825, 9163, 2688, 2691.

Sonidos IADS:

319, 910, 730, 710, 502, 380, 116, 261, 310, 719, 714, 715, 115, 420, 624, 711, 709, 712.

Video Felicidad

Imágenes IAPS:

7200, 1908, 1659, 4603, 5849, 2340, 7405, 8206, 8163, 8170, 8470, 8300, 8350, 4628, 5600, 1710.

Audio: Dolly Parton – 9 to 5.

Video Tristeza

Imágenes IAPS:

2811, 9926, 9070, 9561, 2700, 9171, 2795, 9041, 2900, 7520, 2456, 3300, 9002, 2205, 9471, 9290, 9265, 2752, 9280, 2520, 2230.

Audio: Louis Armstrong - Summertime

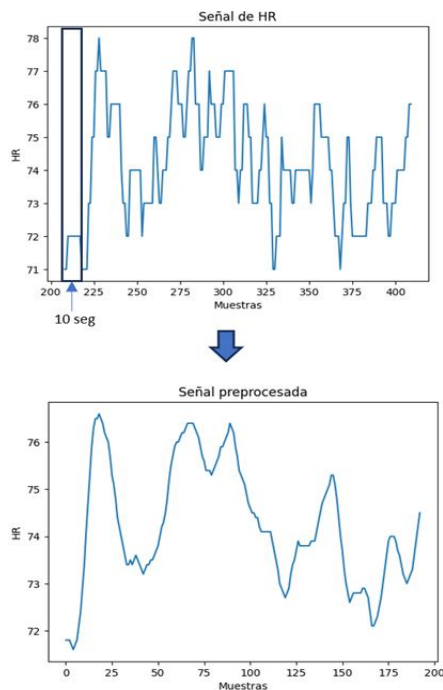
La secuencia de videos utilizados en cada proceso se detalla en la Tabla 1. El primer video de cada proceso correspondía a la emoción de calma, con el objetivo de relajar a los participantes y estandarizar su estado emocional antes de la inducción. Posteriormente, se presentaba el video

Tabla 1. Secuencia de videos de experimento

Video 1	Video 2
Calma 1	Calma 2
Calma 1	Enojo
Calma 1	Felicidad
Calma 1	Tristeza

Tabla 2. Videos utilizados para provocar emociones

Video	Duración (s)
Calma 1	3:27
Calma 2	3:21
Enojo	3:40
Felicidad	2:33
Tristeza	3:31

**Fig. 3.** Señal preprocesada mediante filtro de media móvil

diseñado para provocar la emoción específica correspondiente a la etapa del experimento. Para la formación del conjunto de datos, se utilizaron únicamente los registros de la segunda fase de cada proceso, ya que en esta etapa los participantes estaban expuestos directamente al estímulo destinado a inducir la emoción objetivo. El ritmo cardíaco se extrajo considerando la

duración de cada canción utilizada en los estímulos, como se detalla en la Tabla 2. El equipo utilizado para registrar el ritmo cardíaco fue un sensor óptico de la marca Polar modelo OH1 que registra el ritmo cardíaco en latidos por minuto (bpm).

Durante la segunda fase, se solicitó a los participantes que presionaran determinadas teclas después de observar cada imagen en los videos, con el fin de registrar la intensidad de la emoción percibida. El análisis de los datos recopilados reveló inconsistencias en algunos registros: algunos participantes no reportaron las emociones inducidas de la manera esperada. Otros no registraron datos debido a una falla técnica. A pesar de los inconvenientes en la recolección de datos, fue posible contar con registros completos de todas las emociones inducidas, por lo que se extrajo la información del ritmo cardíaco en cada uno de los procesos del experimento.

El preprocesamiento de datos es una etapa fundamental en el desarrollo del modelo de clasificación de emociones, ya que permite transformar los datos brutos en un formato óptimo para su análisis y procesamiento. En esta investigación, las señales de ritmo cardíaco obtenidas del experimento fueron preprocesadas con el objetivo de eliminar el ruido presente en los registros. Para este proceso, se aplicó la técnica de filtro de media móvil de 10 segundos, la cual suaviza la señal y reduce fluctuaciones no deseadas. Esta técnica se define mediante la siguiente fórmula:

$$y[n] = \frac{1}{M} \sum_{k=0}^{M-1} x[n-k], \quad (1)$$

donde M es el tamaño de la ventana, $x[n]$ es la señal de entrada y $y[N]$ la señal suavizada. La técnica de filtro de media móvil nos permitió suavizar la señal eliminando el ruido que presentaba inicialmente. La señal suavizada se puede ver en el ejemplo de la Figura 3. Esta técnica se aplicó a los datos de ritmo cardíaco, imagen y características de audio.

Extracción de características

La extracción de características es una etapa crucial en el proceso de reconocimiento de emociones, ya que permite obtener los atributos más representativos de las señales capturadas,

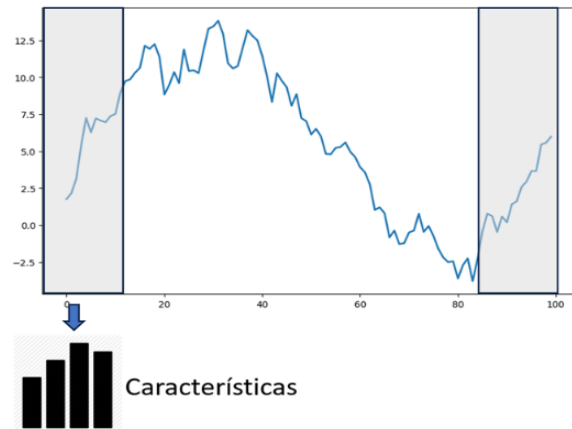


Fig. 4. Extracción de características con Ventana deslizante

facilitando así el entrenamiento de los modelos de aprendizaje automático.

En este trabajo, se empleó una técnica de ventana deslizante para segmentar las señales en fragmentos de tamaño predefinido [16]. Este enfoque permite analizar la variabilidad temporal de las señales y capturar patrones dinámicos en la evolución del ritmo cardíaco. La ventana deslizante facilita la detección de transiciones emocionales al proporcionar información en intervalos sucesivos, lo que resulta esencial para el reconocimiento preciso de estados afectivos.

Para cada fragmento obtenido mediante la ventana deslizante de 10 segundos, se calcularon un conjunto de características. Las características calculadas para cada fragmento del ritmo cardíaco corresponden principalmente a medidas estadísticas clásicas (media, varianza, percentiles, entre otras), comúnmente empleadas en estudios previos relacionados con la clasificación de emociones a partir de señales fisiológicas [17]. Estas características incluyen métricas estadísticas, frecuencia y dominio temporal, entre otras. La Figura 4 ilustra el proceso de segmentación y la extracción de características a partir de los fragmentos generados.

Las características del ritmo cardíaco fueron:

Valor mínimo: Representa el valor más bajo dentro de ese rango de ritmo cardíaco.

$$\min(x) = \min(x_1, x_2, \dots, x_n). \quad (2)$$

Rango (máximo – mínimo): Representa la amplitud de los valores del ritmo cardíaco.

$$\text{rango} = \max(x) - \min(x). \quad (3)$$

Media: Indica el valor general del ritmo cardíaco dentro de ese rango.

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i. \quad (4)$$

Desviación estándar: Indica que tanto fluctúa el ritmo cardíaco dentro de ese rango de valores.

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}. \quad (5)$$

Varianza: Representa otra medida de variabilidad de los valores.

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2. \quad (6)$$

Energía: Es una medida de la actividad de la señal acumulada en un periodo de tiempo.

$$E = \sum_{i=1}^n x_i^2. \quad (7)$$

Potencia: Permite observar cómo se mantiene la energía de la señal a lo largo del tiempo.

$$P = \frac{1}{n} \sum_{i=1}^n x_i^2. \quad (8)$$

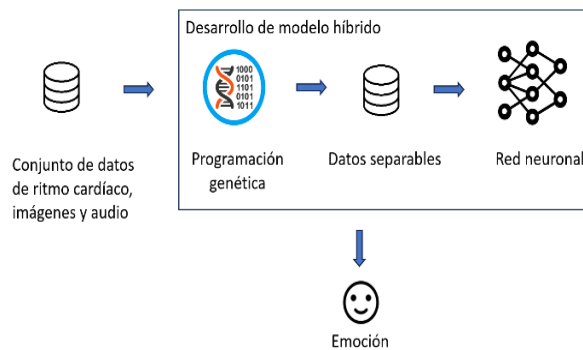
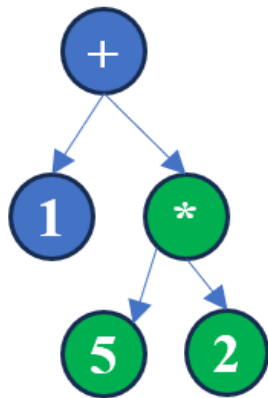
Los archivos de la música utilizada en los estímulos fueron procesados mediante el paquete de lenguaje Python pyAudioAnalysis versión 0.3.14. La función ShortTermFeatures extrae las características por cada segundo de cada archivo

Tabla 3. Ejemplo de archivo con datos de audio

zcr	energy	energy_entropy
0.020182321	4.96E06	2.69062057
0.027574947	5.01E05	3.137089981
0.046940904	0.000177845	3.294968515

Tabla 4. Ejemplo de archivo con datos de imágenes de la emoción de calma

Valencia	Activación	Emoción
6.09	3.8	Calma
6.09	3.8	Calma
6.09	3.8	Calma

**Fig. 5.** Desarrollo de modelo híbrido de clasificación de emociones**Fig. 6.** Ejemplo de individuo de programación genética

de música utilizado en el experimento. El total de las características de audio es de 68 y son: zcr,

energy, energy entropy, spectral centroid, spectral spread, spectral entropy, spectral flux, spectral rolloff, mfcc 1 a mfcc 13, chroma 1 a chroma 12, chroma std. Además, las primeras derivadas delta zcr, delta energy, delta energy entropy, delta spectral centroid, delta spectral spread, delta spectral entropy, delta spectral flux, delta spectral rolloff, delta mfcc 1 a delta mfcc 13, delta chroma 1 a delta chroma 12, delta chroma std. La Tabla 3 muestra un ejemplo de los datos.

Los datos de las imágenes utilizadas en cada estímulo también se ordenaron de acuerdo al orden en que se fueron presentando, formando así un archivo por cada emoción. Los datos corresponden a los valores de valencia y activación de cada imagen que fueron tomados como sus características y se muestran en la Tabla 4.

Los archivos del participante seleccionado para el desarrollo del modelo de clasificación de emociones que contienen los datos de ritmo cardíaco, imágenes y características de audio forman un total de 12; tres por cada proceso.

Las características del ritmo cardíaco, imágenes y audio se etiquetaron con las emociones de calma, enojo, felicidad y tristeza.

El conjunto de datos utilizado para el entrenamiento del modelo de clasificación de emociones se construyó a partir de la unión de los datos de características de ritmo cardíaco calculadas por las ecuaciones 2 a la ecuación 8, los datos de audio de la Tabla 3 y los datos de las imágenes de la Tabla 4 con un total de 78 columnas y 749 renglones. El conjunto se dividió en dos partes la primera de tamaño 78x599 para entrenamiento y la segunda de tamaño 78x150 para prueba.

3.2. Desarrollo del modelo híbrido de clasificación

El desarrollo del modelo híbrido de clasificación de emociones utilizando la técnica de programación genética y redes neuronales se observa en la Figura 5.

La primera parte utiliza el conjunto de datos de ritmo cardíaco, imágenes y audio como entrada para la técnica de programación genética y se obtienen datos en un espacio de valores que

Tabla 5. Definición de variables utilizadas por los individuos

Variable	Característica	Variable	Característica
x0	Zcr	x39	delta spectral entropy
x1	Energy	x40	delta spectral flux
x2	energy entropy	x41	delta spectral rolloff
x3	spectral centroid	x42	delta mfcc 1
x4	spectral spread	x43	delta mfcc 2
x5	spectral entropy	x44	delta mfcc 3
x6	spectral flux	x46	delta mfcc 5
x7	spectral rolloff	x47	delta mfcc 6
x8	mfcc 1	x48	delta mfcc 7
x9	mfcc 2	x49	delta mfcc 8
x10	mfcc 3	x50	delta mfcc 9
x11	mfcc 4	x51	delta mfcc 10
x12	mfcc 5	x52	delta mfcc 11
x13	mfcc 6	x53	delta mfcc 12
x14	mfcc 7	x54	delta mfcc 13
x15	mfcc 8	x55	delta chroma 1
x16	mfcc 9	x56	delta chroma 2
x17	mfcc 10	x57	delta chroma 3
x18	mfcc 11	x58	delta chroma 4
x19	mfcc 12	x59	delta chroma 5
x20	mfcc 13	x60	delta chroma 6
x21	chroma 1	x61	delta chroma 7
x22	chroma 2	x62	delta chroma 8
x23	chroma 3	x63	delta chroma 9
x24	chroma 4	x64	delta chroma 10
x25	chroma 5	x65	delta chroma 11
x26	chroma 6	x66	delta chroma 12
x27	chroma 7	x67	delta chroma std
x28	chroma 8	x68	HR
x29	chroma 9	x69	VAL
x30	chroma 10	x70	ACT
x31	chroma 11	x71	HR max
x32	chroma 12	x72	HR min
x33	chroma std	x73	HR std
x34	delta zcr	x74	HR var
x35	delta energy	x75	HR range
x36	delta energy entropy	x76	HR power
x37	delta spectral centroid	x77	HR energy
x38	delta spectral spread		

puedan ser separables. La siguiente parte entrena una red neuronal con los datos transformados para realizar la clasificación de las emociones de calma, enojo, felicidad y tristeza.

Programación genética

Las técnicas de cómputo evolutivo se basan en la teoría de la evolución de Darwin. Los algoritmos genéticos, optimización por enjambre de partículas, programación genética son algunas de las técnicas utilizadas en computo evolutivo. La

programación genética tiene la ventaja de construir soluciones que tengan una representación flexible usualmente en forma de árbol como se muestra en la Figura 5.

Estos árboles son modificados mediante operadores, como la mutación (cambios aleatorios en nodos o ramas del árbol) y el cruce (intercambio de subárboles entre soluciones candidatas). De esta manera, el árbol sintáctico de la Figura 6 es el resultado de un proceso evolutivo que busca optimizar la función de clasificación.

Tabla 6. Definición de funciones utilizadas por los individuos

Función	Nombre
Sum	Suma
Sub	Resta
Mul	multiplicación
pDiv	división
Sqrt	raíz cuadrada
Sin	Seno
Cos	Coseno
Tan	tangente

La programación genética comienza con una población generada aleatoriamente compuesta por individuos o programas. Los nodos terminales son seleccionados aleatoriamente de un conjunto de terminales que contiene variables o constantes. Los nodos internos son seleccionados de un conjunto de funciones que contiene operaciones aritméticas y operadores lógicos.

Los individuos son expresiones que incluyen a las características. Los individuos son evaluados por una función de aptitud y los que tengan mayor aptitud son seleccionados para la siguiente generación. Los operadores genéticos de cruce y mutación son aplicados para producir nuevos individuos. El proceso de programación genética continúa seleccionando individuos para la siguiente generación y generando individuos mediante los operadores genéticos hasta que se cumple un criterio de paro.

El proceso de programación genética tiene como salida al mejor individuo de la población.

Terminales: Las soluciones que se desarrollaron con la técnica de programación genética se representan en forma de árbol sintáctico de operaciones. Las terminales se definen en este caso como características con las que se harán las operaciones. Las variables y su correspondiente característica se muestran en la Tabla 5.

Funciones: Las soluciones desarrolladas por la técnica de programación genética se definen como un árbol sintáctico en el que los nodos intermedios son funciones matemáticas y los nodos terminales son las características antes mencionadas.

Las funciones utilizadas (Tabla 6) por las soluciones son operaciones aritméticas básicas:

suma, resta, multiplicación, división y raíz cuadrada; funciones aritméticas: seno, coseno, tangente. La división se encuentra protegida para que en el caso de que el denominador sea cero devuelva el valor de 1. La raíz cuadrada se redefine también como la raíz cuadrada del valor absoluto pasado como parámetro para proteger la función en caso de error.

Operación de cruce: La operación de cruce que se realiza sobre los árboles consiste en seleccionar un punto en 2 árboles e intercambiar sus ramas para así generar dos nuevas soluciones.

Operación de mutación: La operación de mutación se realiza seleccionando un nodo al azar en donde se cambia la función generando así una nueva solución.

Individuos: Los individuos generados con programación genética se representan mediante un árbol sintáctico en el que las hojas son las variables o características y los nodos intermedios son funciones aritméticas. El individuo al representar una función transforma las características de entrada en un valor de salida. La función a pesar de ser no lineal y al representar un solo valor puede no ser suficiente para ser utilizada como un clasificador. El enfoque multi árbol en programación genética permite que cada individuo pueda contener varias funciones.

Un individuo multi árbol puede generar más de una función para que de esa forma se obtenga como salida varios valores. Los valores de salida de las funciones del individuo pueden utilizarse como entrada en una red neuronal para realizar la separación de las instancias y crear un clasificador de las emociones. Un ejemplo de los individuos generados con la técnica de programación genética y el enfoque multi árbol se muestra en la Figura 6.

Operación de mutación: La operación de mutación se realiza seleccionando un nodo al azar en donde se cambia la función generando así una nueva solución.

Individuos: Los individuos generados con programación genética se representan mediante un árbol sintáctico en el que las hojas son las

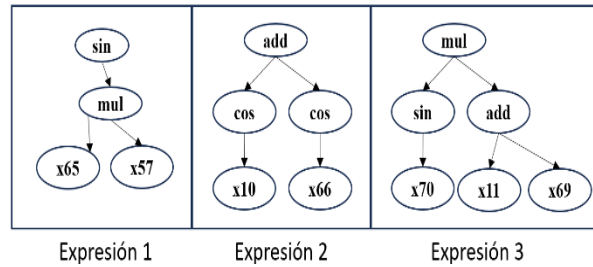


Fig. 7. Ejemplo de individuo multi árbol

Tabla 7. Parámetros utilizados en el proceso evolutivo

Parámetro	Valor
Probabilidad de cruce	70%
Probabilidad de mutación	30%
Tamaño de la población	100
Número de generaciones	20

variables o características y los nodos intermedios son funciones aritméticas.

El individuo al representar una función transforma las características de entrada en un valor de salida. La función a pesar de ser no lineal y al representar un solo valor puede no ser suficiente para ser utilizada como un clasificador. El enfoque multi árbol en programación genética permite que cada individuo pueda contener varias funciones. Un individuo multi árbol puede generar más de una función para que de esa forma se obtenga como salida varios valores. Los valores de salida de las funciones del individuo pueden utilizarse como entrada en una red neuronal para realizar la separación de las instancias y crear un clasificador de las emociones. Un ejemplo de los individuos generados con la técnica de programación genética y el enfoque multi árbol se muestra en la Figura 7.

Función de aptitud: Los individuos con mayor aptitud son seleccionados para aplicar las operaciones de cruce y mutación. Esta selección se basa en el valor asignado a cada individuo. La función de aptitud debe elegirse cuidadosamente según las características del problema que se busca resolver. El desempeño del modelo de clasificación de emociones depende de su capacidad para cumplir con la tarea asignada. En

este contexto, la métrica de evaluación F1 resulta especialmente útil, ya que permite medir el equilibrio entre precisión y exhaustividad. Esta métrica proporciona una visión clara sobre la cantidad de errores cometidos y la calidad de la clasificación obtenida. Durante el proceso evolutivo, la métrica F1 se empleó para evaluar el rendimiento de los individuos. A continuación, se presentan las métricas utilizadas en el proceso de evaluación.

F1-score: Combina las métricas de precisión y recall para definir la capacidad de clasificación de un modelo.

$$F1 = 2 * \frac{precision * recall}{precision + recall} \quad (9)$$

La métrica F1-score permite encontrar soluciones en las que la clasificación se realiza de manera balanceada entre la precisión y el recall.

Precisión: Indica qué proporción de los ejemplos predichos como positivos son realmente positivos.

$$precision = \frac{TP}{TP + FP} \quad (10)$$

En donde TP son los verdaderos positivos y FP los falsos positivos

Recall o sensibilidad: Es la capacidad del modelo para identificar correctamente los casos positivos.

$$recall = \frac{TP}{TP + FN} \quad (11)$$

En donde TP son los verdaderos positivos y FN los falsos negativos.

Parámetros del proceso evolutivo: Los parámetros del proceso evolutivo utilizados permiten controlar la búsqueda de soluciones mediante programación genética y se definen como: las probabilidades de que los individuos de la población se crucen y muten, el tamaño de la población y el número de generaciones. Los valores específicos de estos parámetros fueron

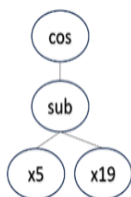


Fig. 8. Parte 1 de solución de programación genética

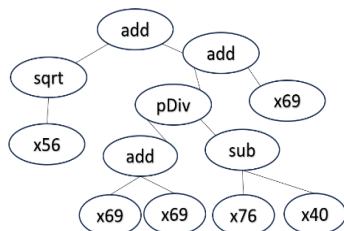


Fig. 9. Parte 2 de solución de programación genética

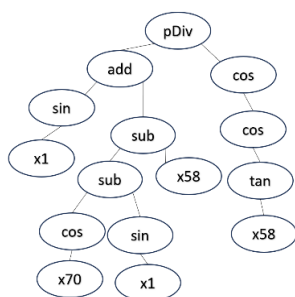


Fig. 10. Parte 3 de solución de programación genética

Tabla 8. Características utilizadas en la solución encontrada con programación genética

Variable	Característica
x1	Energy
x5	spectral entropy
x19	mfcc 12
x40	delta spectral flux
x56	delta chroma 2
x58	delta chroma 4
x69	VAL
x70	ACT
x76	HR power

seleccionados de manera empírica, a partir de pruebas preliminares en las que se buscó un balance entre la exploración del espacio de soluciones y la eficiencia computacional del modelo. Esta elección buscó evitar tanto la

convergencia prematura como la prolongación innecesaria del proceso evolutivo. Los parámetros finales utilizados se resumen en la Tabla 7.

La técnica de programación genética realiza una búsqueda de funciones que transforman el conjunto de datos de entrada en valores que puedan ser utilizados por la red neuronal para ser clasificados dentro de las emociones de calma, enojo, tristeza y felicidad. La solución encontrada se muestra en tres partes en la Figura 8, Figura 9 y Figura 10.

La solución forma un sistema de ecuaciones en el que los datos son transformados en 3 valores. Los valores obtenidos de las ecuaciones son clasificados por la red neuronal. Las ecuaciones de la solución son las siguientes:

$$y = \cos(x_5 - 19), \quad (12)$$

$$y = \sqrt{x_{56}} + \frac{2x_{69}}{x_{76} - x_{40}} + x_{69}, \quad (13)$$

$$y = \frac{\sin(x_1) + (\cos(x_{70}) - \sin(x_1) - x_{58})}{\cos(\cos(\tan(x_{58})))}, \quad (14)$$

Las ecuaciones 12, 13 y 14 muestran las variables utilizadas que corresponden a las características de las imágenes, audio y ritmo cardíaco del conjunto de datos. Las características incluidas en el conjunto de datos como se mencionó anteriormente forman un total de 78 pero como se muestra en la Tabla 8 sólo se utilizan 9 en la solución.

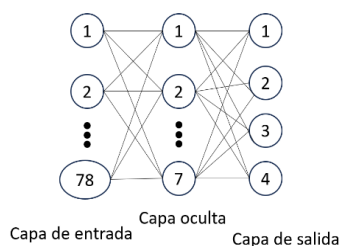
La técnica de programación genética además de encontrar una solución en la que la red neuronal pudiera clasificar de manera correcta los datos también encontró una solución que utiliza menos características.

Red neuronal

Una red neuronal artificial (RNA) es un modelo computacional inspirado en la estructura y funcionamiento del cerebro humano. Consiste en un conjunto de nodos o neuronas artificiales, organizadas en capas que procesan información mediante conexiones ponderadas. Estas redes

Tabla 9. Parámetros utilizados en la red neuronal

Parámetro	Valor
activation	relu
solver	adam
hidden layer sizes	7
alpha	0.0001
batch size	16
learning rate init	0.001
learning rate	Adaptive
max iter	100000
early stopping	True
Verbose	False

**Fig. 11.** Arquitectura de la red neuronal

son utilizadas para resolver problemas complejos de aprendizaje automático, como clasificación, regresión, reconocimiento de patrones y toma de decisiones. Las redes neuronales artificiales están compuestas por tres tipos principales de capas:

- Capa de entrada: recibe los datos de entrada y los transmite a las capas ocultas.
- Capas ocultas: realizan cálculos y transformaciones mediante funciones de activación, ajustando los pesos de las conexiones en función de un algoritmo de aprendizaje.
- Capa de salida: genera la salida final del modelo, que puede ser una predicción o clasificación.

La red neuronal utilizada para el entrenamiento del modelo se implementó utilizando la clase `MLPClassifier` integrada en paquete de Python Scikit-learn versión 1.3.2 utilizando los parámetros que se muestran en la Tabla 9.

Los parámetros de la Tabla 9 definen la arquitectura de la red neuronal que se muestra en la Figura 11 integrada por tres capas: i) La capa de

entrada con 78 valores correspondientes a las características del conjunto de datos; ii) La capa oculta con 7 neuronas, y iii) La capa de salida del modelo de clasificación de emociones está compuesta por 4 neuronas. Cada capa corresponde a una de las emociones: calma, enojo, tristeza y felicidad. Esta capa emplea la función de activación ReLU, que contribuye a que la red aprenda a distinguir entre las emociones.

3.3. Análisis de resultados

Los resultados obtenidos tras aplicar la técnica de programación genética al conjunto de datos y entrenar la red neuronal para clasificar las emociones de una persona se presentan en esta sección. La programación genética también actúa como un mecanismo implícito de selección de características, ya que su proceso evolutivo tiende a identificar subconjuntos de variables que maximizan la función de aptitud.

El modelo desarrollado logró una precisión del 95% en los datos de entrenamiento y 94% en los datos de prueba para la clasificación de las emociones de calma, enojo, felicidad y tristeza.

La cantidad de instancias correctamente clasificadas por el modelo puede observarse en las Figuras 12 y 13, lo que permite evaluar su desempeño en términos de exactitud y fiabilidad. Las emociones de enojo y tristeza resultaron con algunos fallos en su clasificación como se muestra en las Figuras 12 y 13.

Los datos que se utilizan para el cálculo de las funciones utilizan características de audio, imagen y ritmo cardíaco.

El clasificador construido a partir de estos datos nos indica que para el caso de las emociones de enojo y tristeza los valores obtenidos son similares después de realizar los cálculos. Las imágenes y audio pueden tener valores similares entre estas dos emociones.

El ritmo cardíaco de la persona utilizado en el conjunto de datos también puede tener valores similares entre las emociones de enojo y tristeza. El resultado obtenido por el clasificador nos da un indicio de que para construir un modelo que pueda clasificar más emociones sea necesario involucrar más datos.

Test Confusion Matrix

True Label \ Predicted Label	calma	enojo	felicidad	tristeza
calma	38	0	0	0
enojo	0	37	1	4
felicidad	0	0	29	0
tristeza	3	1	0	37

Fig. 12. Matriz de confusión con datos de prueba

Train Confusion Matrix

True Label \ Predicted Label	calma	enojo	felicidad	tristeza
calma	154	0	0	0
enojo	0	159	8	2
felicidad	0	0	115	0
tristeza	11	11	0	139

Fig. 13. Matriz de confusión con datos de entrenamiento

5 Conclusiones y trabajos futuros

Este estudio demuestra la eficacia de combinar programación genética y redes neuronales para clasificar emociones a partir del ritmo cardíaco captado por dispositivos inteligentes. El experimento controlado permitió inducir emociones mediante estímulos audiovisuales, preprocesar la señal con un filtro de media móvil y extraer características relevantes mediante ventanas deslizantes. La programación genética generó un sistema de ecuaciones no lineales que facilitó la clasificación, alcanzando un 95% de precisión en entrenamiento y 94% en prueba.

Sin embargo, las emociones de enojo y tristeza presentaron una mayor tasa de error,

evidenciando la necesidad de explorar atributos adicionales para diferenciarlas mejor. Aunque los resultados se obtuvieron con un único participante, la viabilidad del enfoque queda demostrada, reconociendo la necesidad de validarlo con una muestra más amplia.

Este trabajo forma parte de un proyecto que busca, en el futuro, desarrollar modelos capaces de inducir cambios emocionales. Se están evaluando técnicas de selección de características más rigurosas y la integración de señales fisiológicas adicionales como GSR, HRV y temperatura corporal. También se contempla la aplicación del modelo en contextos reales, como entornos laborales o educativos, para evaluar su desempeño en condiciones más dinámicas y con poblaciones diversas.

Referencias

1. Bradley, M., Lang, P. (2007). The International Affective Digitized Sounds (2nd Edition; IADS-2).
2. Francisti, J., Balogh, Z., Reichel, J., Benko, L., Fodor, K., Turcani, M. (2023). Identification of heart rate change during the teaching process. *Scientific Reports*, Vol. 13, No. 1, pp. 16674.
3. Griffiths, D., Cunningham, S., Weinle, J., Picking, R. (2021). A multi-genre model for music emotion recognition using linear regressors. *Journal of New Music Research*, Vol. 50, No. 4, pp. 355–372.
4. Lang, P., Bradley, M., Cuthbert, B. (1997). International affective picture system (IAPS): Affective ratings of pictures and instruction manual.
5. Limbani, D., Yatin, D., Chaturvedi, N., Moorthy, V., ..., Kumar, S. (2023). WEARS: Wearable Emotion AI with Real-time Sensor data.
6. Nishibe, T., Uchiya, T., Ichi, T., Funase, A. (2023). Analysis of the Relationship between Various Heart Rate Indices and Emotion during Emotion Induction by Image Stimulation. 2023 IEEE 12th Global

- Conference on Consumer Electronics (GCCE), IEEE, Nara, Japan, pp. 1166–1167.
7. **Pepa, L., Capecci, M., Ceravolo, M. G. (2019).** Smartwatch based emotion recognition in Parkinson's disease. 2019 IEEE 23rd International Symposium on Consumer Technologies (ISCT), IEEE, Ancona, Italy, pp. 23–24.
 8. **Pham, M., Do, H. M., Su, Z., Bishop, A., Sheng, W. (2021).** Negative Emotion Management Using a Smart Shirt and a Robot Assistant. IEEE Robotics and Automation Letters, Vol. 6, No. 2, pp. 4040–4047.
 9. **Picard, R. (1995).** Affective Computing. M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 321.
 10. **Russell, J. A. (1980).** A circumplex model of affect. Journal of Personality and Social Psychology, Vol. 39, No. 6, pp. 1161–1178.
 11. **Shu, L., Yu, Y., Chen, W., Hua, H., Li, Q., Jin, J., Xu, X. (2020).** Wearable Emotion Recognition Using Heart Rate Data from a Smart Bracelet. Sensors, Vol. 20, No. 3, pp. 718.
 12. **Takeshita, R., Shoji, A., Hossain, T., Yokokubo, A., Lopez, G. (2021).** Emotion Recognition from Heart Rate Variability Data of Smartwatch While Watching a Video. 2021 Thirteenth International Conference on Mobile Computing and Ubiquitous Network (ICMU), IEEE, Tokyo, Japan, pp. 1–6.
 13. **Wang, Z., Yu, Z., Zhao, B., Guo, B., Chen, C., Yu, Z. (2020).** EmotionSense: An Adaptive Emotion Recognition System Based on Wearable Smart Devices. ACM Transactions on Computing for Healthcare, Vol. 1, No. 4, pp. 1–17.
 14. **Younis, E. M. G., Mohsen, S., Houssein, E. H., Ibrahim, O. A. S. (2024).** Machine learning for human emotion recognition: A comprehensive review. Neural Computing and Applications, Vol. 36, No. 16, pp. 8901–8947.
 15. **P. J. Bota, C. Wang, A. L. N. Fred and H. Plácido Da Silva (2019),** "A Review, Current Challenges, and Future Possibilities on Emotion Recognition Using Machine Learning and Physiological Signals," in IEEE Access, vol. 7, pp. 140990-141020, 2019.
 16. **Oyeleye, M., Chen, T., Titarenko, S., Antoniou, G. (2022).** A Predictive Analysis of Heart Rates Using Machine Learning Techniques. International Journal of Environmental Research and Public Health, 19(4), 2417.
 17. **Pamungkas, Y., Wibawa, A. D., Rais, Y. (2022).** Classification of Emotions (Positive-Negative) Based on EEG Statistical Features using RNN, LSTM, and Bi-LSTM Algorithms. 2nd International Seminar on Machine Learning, Optimization, and Data Science (ISMODE), Jakarta, Indonesia, 2022, pp. 275–280.
 18. **Hettiarachchi, I., Hanoun, S., Nahavandi, D., Nahavandi, S. (2019).** Validation of Polar OH1 optical heart rate sensor for moderate and high intensity physical activities. PLOS ONE. 14.

Article received on 09/01/2025; accepted on 24/07/2025.

**Corresponding author is Alicia Martínez-Rebollar.*