

Modelo para la Detección de Estados Emocionales Basado en el Enfoque Cognitivo, usando máquinas de aprendizaje

Marco Antonio Mojica Martinez

División de Estudios de Posgrado e Investigación
Tecnológico Nacional de México - Campus Acapulco
Acapulco, Guerrero.
mm20320011@acapulco.tecnm.mx

José Luis Hernández Hernández

Depto. Desarrollo de sistemas inteligentes
Tecnológico Nacional de México - Campus Chilpancingo
Chilpancingo, Guerrero.
joseluis.hh@itchilpancingo.edu.mx

Miriam Martinez Arroyo

División de Estudios de Posgrado e Investigación
Tecnológico Nacional de México - Campus Acapulco
Acapulco, Guerrero.
miriam.ma@acapulco.tecnm.mx

José Antonio Montero Valverde

Depto. Desarrollo de sistemas inteligentes
Tecnológico Nacional de México - Campus Acapulco
Acapulco, Guerrero.
jose.mv@acapulco.tecnm.mx

Resumen

En el campo de la inteligencia artificial existen diversos trabajos relacionados con el estudio de las emociones, el objetivo de esta propuesta de investigación es presentar un prototipo sobre un modelo para percibir el estado emocional de las personas.

Para poder llevar a cabo este estudio de investigación se propone un modelo fundamentado en el enfoque cognitivo tanto para la detección de emociones y la intensidad niveles de emoción de un persona; desde una perspectiva teórica de acuerdo con el estado del arte se enfoca en dos metodologías en expresiones faciales y reacciones fisiológicas.

El modelo que se pretende obtener conjunta la visión artificial y el uso de señales continuas, mediante el uso de una cámara y un sensor que permite medir la sudoración de la piel conocido como GSR.

El trabajo propuesto se encuentra en fase de desarrollo y actualmente con avance en marcado de zonas de interés de rasgos de la cara.

Keywords Emociones, Cognitiva, Expresiones Faciales, GSR

I. INTRODUCCIÓN

De acuerdo con Russek, S.[1] las emociones son instintos de supervivencia y nos sirven para adaptarnos al medio ambiente y son: Alegría, tristeza, rabia, miedo, amor y enojo.

Los enfoques que permiten el estudio de las emociones en un ser humano son: Teoría de James-Lange basada en la Percepción → Reacción Motora → Reacción Visceral → Emoción o la Teoría de Cannon-Bart se apoya en Percepción → Reacción talámica → Sensación emocional[2].

Como objetivo principal de este trabajo de investigación es la presentación de un modelo que permita la percepción de las emociones utilizando:

1. La visión artificial para la detección de cara y expresiones faciales de una persona.
2. La detección de reacciones fisiológicas que son captadas utilizando distintos sensores de acuerdo a la reacción que se desea estudiar. En este caso la sudoración de la piel utilizando un sensor GSR.
3. Máquinas de aprendizaje, el cual permite la construcción de aplicaciones inteligentes usando la inteligencia artificial.

4. El uso del enfoque cognitivo para el desarrollo de un modelo para la detección de emociones combinado con los puntos mencionados anteriormente.

La estructura del trabajo esta compuesto por: Enfoque cognitivo Base, Trabajos Relacionados, Metodología, Aplicación, Resultados Preliminares, Discusión y Conclusiones.

II. ENFONQUE COGNITIVO BASE

De acuerdo con Legg, T.[2] el enfoque de Cannon-Bard sobre el estado de la emociones nos dice que las emociones son eventos simultáneos disparados por respuestas de sentimientos y emociones, así como de reacciones físicas que ocurren al mismo tiempo, es decir, una reacción física no depende de una reacción emocional y viceversa.

El enfoque fue propuesto por Walter B. Cannon y su estudiante graduado Philp Bard , los cuales proponen que las reacciones emocionales se originan en el tálamo que responde al recibir información sensorial y la transmite para poder ser procesada por el cerebro. Así que cuando los eventos se disparan el tálamo manda una señal a la amígdala que procesa emociones fuertes como enojo, angustia o miedo, también a la corteza cerebral donde se encuentra los pensamientos consientes y finalmente al sistema nervioso como al esquelético encargados de controlar las reacciones física. Este enfoque es conocido como la teoría talámica de la emoción,(Fig. 1)

Una reacción física puede ser la sudoración o tensión muscular.



Fig. 1. Teoría talámica de Canno-Brad.[3]

El enfoque cognitivo hace posible percibir las emociones usando la inteligencia artificial por ejemplo se puede usar la visión artificial en las expresiones faciales y el uso de

sensores para captar las señales de reacciones corporales, en la siguiente capítulo se da detalles sobre su aplicación.

Las emociones se producen por la activación corporal con unión de la interpretación cognitiva de la persona

Este enfoque permite conocer los cimientos psicológicos para la creación de un modelo que nos permita el estudio de los estados emocionales.

III. TRABAJOS RELACIONADOS

El estudio de las emociones usando las ciencias computaciones fue propuesto por Picard,R.[4] que planteo la combinación de la neurociencia, la psicología y la inteligencia artificial como base para el nacimiento de la computación afectiva.

Este campo permite el estudio y el desarrollo de sistemas capaces de reconocer e interpretar las emociones humanas. Actualmente existen diversos trabajos aplicados a esta disciplina.

El modelo propuesto para el estudio de las emociones y los niveles de emoción se basa en dos trabajos:

- 1) Propuesto por Hassouneh , A.[5] cuyo objetivo es la detección de emociones en tiempo real, usando expresiones faciales y ondas cerebrales a través conocidas como EEG (Electroencefalograma); ahora bien para el caso de la expresiones faciales usa redes neuronales convolucionales y para el caso de las señales EEG usa redes neuronales recurrentes para la clasificación de la emoción detectada.

- 2) El segundo trabajo planteado por Memar, M.[6] detecta los niveles de estrés de una persona usando un reacción corporal del ser humano conocida como sudoración la cual es cuantificado por un sensor GSR (Galvanic Skin Response) el cual mide los cambios eléctricos que presenta la piel por la actividad de las glándulas sudoríparas de acuerdo al estado emocional; para la determinación de los niveles de la emoción con el uso del sensor GSR se usa un clasificador ANOVA o análisis de varianza.

IV. METODOLOGÍA

Existen diversos métodos para percibir las emociones de las personas, de acuerdo a un estudio llevado por Saxena, A.[7] (Fig. 2) son cuatro tipos:

- 1) Texto que usa la escritura para la detección del estado emocional de un individuo.

- 2) Voz que dispone del reconocimiento vocal.

3) Expresiones faciales que se basa en la interpretación de los gestos de la cara.

4) Señales fisiológicas como la sudoración de la piel (GSR), EEG (Electroencefalograma) o ECG(ritmo cardíaco).

El modelo propuestos como prototipo hace uso de los trabajos relacionados expuestos en el punto anterior los cuales son el uso de expresiones faciales y el que se basa en la señales fisiológicas, en este caso las que son causadas por las glándulas sudoríparas, que son un reflejo de la intensidad de nuestro estado emocional.

Por lo que se propone el uso de cámara para captar las expresiones faciales y un sensor GSR para medir la sudoración de la piel.

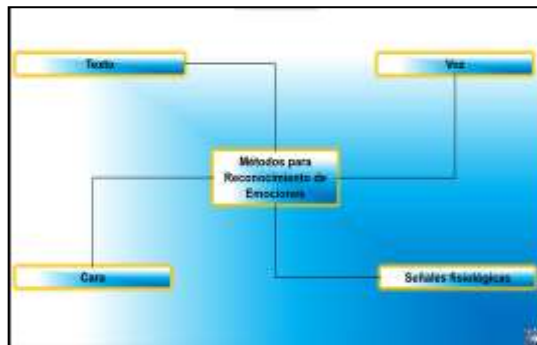


Fig. 2. Métodos plateados por Saxena, A.

4.1. Modelo para detección de emociones mediante expresiones faciales.

Este modelo se estructura, de acuerdo con la imagen (Fig. 3) la detección de emociones tiene los siguientes pasos:

1) Comienza con la adquisición de un imagen del entorno utilizando una cámara digital, la cual es procesada mediante algoritmos de visión artificial: como cambio a escala de grises, redimensionamiento la imagen y segmentación.

2) Detección de Cara en este paso se puede utilizar el método de HOG conocido como Histograma de Extracción de Característica propuesto por Dalal y Triggs[8] el cual divide la imagen en pequeñas regiones llamadas celdas y en cada una de ellas se calcula su histograma de dirección de gradiente o bordes de orientación, como siguiente paso discreta cada celda dándole un peso al contenedor angular y finalmente cada contenedor es agrupado y normalizado o también se puede usar el método Haar-Cascade que para el caso de captación de la cara tiene 32 clasificadores en una estructura de cascada que exponen Viola y Jones[9] en su investigación sobre Detección de Cara Robusta en tiempo

real en la cual integra un concepto llamado imagen integral que permite la evaluación de las características más rápido y son computadas en pocas operaciones por pixel.

3) Extracción de características, en este paso se extrae los puntos clave de la cara como ojos, boca, nariz, mentón y cejas usando marcadores faciales conocidos en inglés como

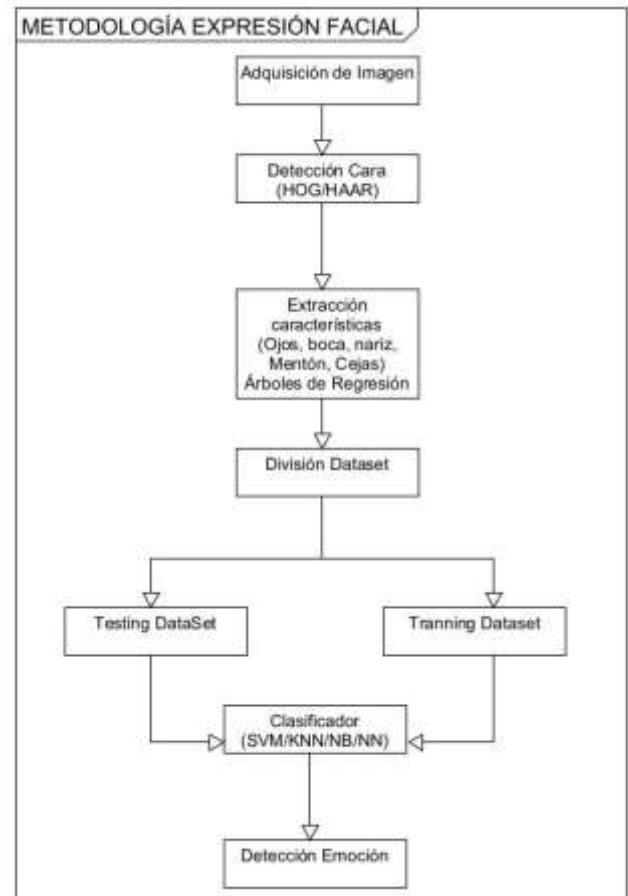


Fig. 3. Metodología de expresiones faciales [5].

landmarks que usan arboles de regresión propuesto por Kazemi, V, & Sullivan, J. [10].

4) En este paso se procede a la creación del dataset el cual guarda las distancias de cada una de las características del punto anterior; donde cada registro representa una emoción a clasificar. Ya con el número de muestras organizadas en el dataset se procede a la división del archivo en dos partes una para un conjunto de pruebas conocido como Testing DataSet y el otro para entrar el modelo a usar denominado Tranning DataSet.

5) Clasificador, en esta etapa se procede a la clasificación de la imagen a analizar; los clasificadores a utilizar son: máquinas de vectores de soporte o máquinas de vector soporte (del inglés Support Vector Machines, SVM), el método de los k vecinos más cercanos (en inglés, k-nearest

neighbors, KNN), clasificador bayesiano ingenuo (del inglés, Naive Bayes, NB) y redes neuronales (en inglés, Neuronal Network, NN).

Es necesario evaluar cada clasificador de acuerdo con la matriz de confusión el cual nos permite evaluar el desempeño de cada algoritmo así como la validación de precisión del mismo.

6) El resultado del clasificador revelará en el paso final, así el modelo identificará la emoción que se está analizando en una imagen.

Estas seis fases son necesarias para el uso del método de expresiones faciales para la detección de emociones usando la inteligencia artificial.

4.2. Modelo para detección de Estado Emocional usando Señales Fisiológicas.

A continuación se procede con el desarrollo de la segunda parte del modelo, el cual se centra en la detección de estados emocional usando la señal fisiológica producida para las glándulas sudoríparas.

Las emociones fuertes pueden provocar un estímulo en el sistema nervioso simpático, lo que hace que las glándulas sudoríparas secreten más sudor. GSR (Fig. 4) le permite detectar emociones tan fuertes simplemente colocando dos electrodos en dos dedos de una mano de acuerdo con la página seestudio[11].



Fig. 4. Sensor GSR, Imagen recuperada [11].

El modelo para medir los niveles de emoción, se representa tal como se muestra en la imagen (Fig. 7), el cual toma de base el método de reacciones fisiológicas utilizando un sensor (GSR) que permite medir la sudoración de la piel.

Los pasos a seguir son los siguientes:

1) En el primer paso del modelo es la adquisición de la señal continua la cual es a través de un sensor GSR, usando IDE de Arduino para su programación.

2) Después de obtener la señal es necesario filtrarla con el fin de eliminar ruido y atenuarla con un filtro de bandpass, obtenido una mejor calidad para su procesamiento.

3) La extracción de características de la señal es un paso importante para la determinación de los niveles de emoción de una persona, de acuerdo a Memar, M.[6] hay dos componentes vitales: lo niveles de conductancia de la piel (en inglés, skin conductance level, SCL) conocido como nivel tónico y la respuesta a la conductancia de la piel (en inglés, skin conductance response , SCR) denominada fases de respuesta. La SCL no contiene suficiente información, pero la SCR si la hay para identificar la emoción estimulada por eventos; mediante las siguientes particularidades:

- a) La lactancia de respuesta.
- b) La amplitud.
- c) El tiempo de subida.
- d) La mitad del valor del tiempo en recuperar la señal.

4) La siguiente etapa es la creación de dataset, con las características extraídas en el paso con el objetivo de poder clasificar la señal y obtener los niveles emocionales.

5) Continuando con el proceso se procede con la división de dataset en dos partes: una muestra del conjunto original para las pruebas conocido como Testing Dataset y otra para el entrenamiento del modelo denominado Training Dataset.

6) En la etapa de clasificación se agrupan la señal dando como resultado un nivel emocional; mediante el uso de los siguientes método ANOVA o análisis de varianza, es un modelo estadístico que proporciona una prueba si los medios de varios grupos son iguales. O también se pueden utilizar otros clasificadores como KNN, SVM, y NN mencionados en el punto anterior

4.3. Combinación de los modelos de expresión facial y reacciones fisiológicas

En la imagen de la (Fig. 5) se observa que el modelo de gestos faciales da la emoción de una persona y por el otro lado los niveles de emoción, la conjunción de los dos modelos permitirá obtener el estado emocional de una persona de acuerdo con los estímulos de su entorno. Con el estado emocional obtenido, permitirá el estudio del comportamiento del ser humano expuesto a diversas situación como por ejemplo, ver la reacción de un individuo

cuando una persona se cambia de hogar, el cual experimenta una sensación de felicidad si el cambio es positivo y en un caso contrario tristeza.

En la actualidad el campo de aplicación en la detección ha crecido en la siguiente sección se da una panorámica de su uso en la vida real.

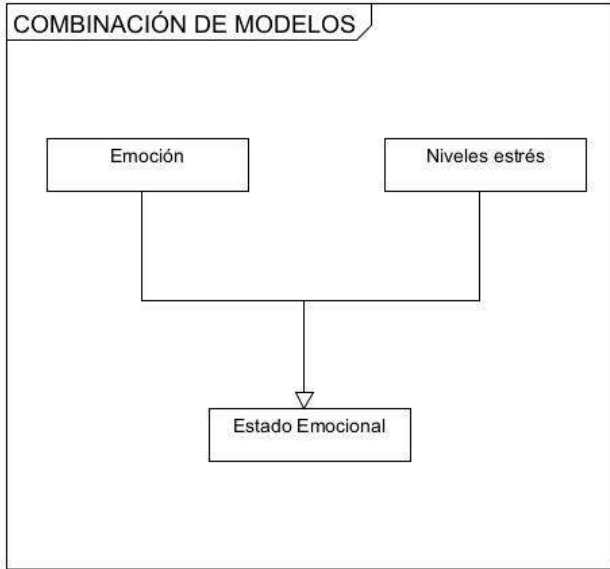


Fig. 5. Combinación de Modelos.

V. RESULTADOS PRELIMINARES

Actualmente el trabajo de investigación está en la etapa de diseño por lo que se presentan los modelos para determinar el estado emocional, como resultado se presentan dos etapas del modelo de expresiones faciales los resultados son los siguientes.

5.1. Detección de Rostros.

Mediante el uso de visión artificial y utilizando una cámara digital se procede con la captura de la imagen y aplicando HOG para le detección de objetos (cara), en este caso se tiene ya identificado el rostro de un persona (Fig. 6).



Fig. 6. Detección de cara Marco A. Mojica.

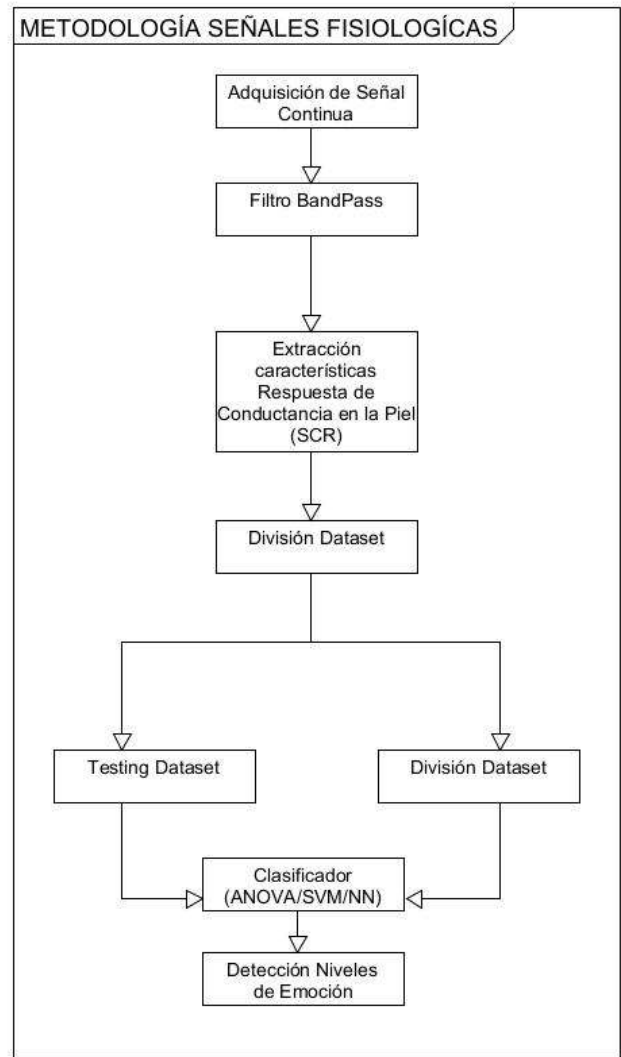


Fig. 7. Metodología de señales fisiológicas [6].

5.2. Extracción de Características.

Se procede a marcar zonas de interés, por medio de árboles de regresión se procede a marcar zonas de interés del rostros de una persona con el fin de determinar la expresiones faciales en este caso ya es posible la detección de ojos, cejas, mentón y boca de un individuo (Fig. 8).



Fig. 8. Zonas de interes del rostro.

Con este paso es posible la extracción de los rasgos importantes para la detección de emociones.

Por lo que actualmente se está trabajando en el paso 3 para expresiones faciales.

VI. APLICACIONES

La detección de emociones tiene las siguientes aplicaciones:

1) Como auxiliar de recursos humanos: En este caso la detección de emociones permite evaluar si el candidato es honesto y está interesado en la posición vacante con el fin de crear un reporte final sobre su reclutamiento.

2) Medir la seguridad: La percepción de emociones es útil en escuelas en la prevención de violencia, mejorando la seguridad del lugar en general.

3) Servicios al cliente: Como herramienta para determinar la satisfacción de un comprador del servicio que recibió, si hay una calificación baja el sistema puede aconsejar al empleado para mejorar la calidad del servicio.

4) Testing de video Juegos. Con el fin de recibir retroalimentación el reconocimiento de emociones es usado en diferentes etapas del desarrollo del video juego.

5) Salud: Estudio del comportamiento humano, es útil en la psicología para determinar si necesita ayuda especializada.

6) Tutores Inteligentes. Con el fin de medir el estado emocional de una persona durante su aprendizaje en línea que emplea un agente para la enseñanza.

Este tipo de desarrollos tienen entornos que poseen fuertes estímulos para los individuos, lo cual permite ser analizados con las propuestas anteriormente y tener herramientas de apoyo.

VII. DISCUSIÓN

De acuerdo con lo planteado anteriormente la metodología utilizada es más económica en comparación con el uso de un sensor EEG, ya que lo que se busca es la facilidad de adquisición por parte del usuario.

Aun que todo va a depender del escenario donde se utilice, por ejemplo el uso de un sensor EEG permite evaluar el estado emocional de una persona con parálisis facial, donde el método de expresiones faciales no aplicaría en este caso.

Así que es necesario ver el contexto para la aplicación de las metodologías a aplicar, en este caso se va a dirigir a una solución informática que nos permita evaluar las emociones con tutores inteligentes.

Dado al incremento de herramientas educativas en línea que usan tutores inteligentes para el aprendizaje, este estudio se enfocará para evaluar los estados emocionales de los estudiantes.

VIII. CONCLUSIONES

El estudio de las emociones permite entender el comportamiento del ser humano desde un punto de vista en el que las emociones son estimuladas por el entorno que los rodea.

Tener modelos y aplicarlos ayudarán a entender este tipo de conductas con el fin de ser una herramienta de apoyo en el campo de la psicología o la educación.

Actualmente la salud mental es un problema en crecimiento debido a la pandemia y problemas sociales, así que contar con instrumentos de apoyo es una necesidad para su estudio.

IX. AGRADECIMIENTOS

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) y al Posgrado en Sistemas Computacionales de Tecnológico Nacional de México – Campus Acapulco (TecNM - ITA) por proporcionarme su apoyo y la beca otorgada para poder realizar la investigación de este artículo y seguir desarrollando este trabajo durante mi estancia en el posgrado.

REFERENCIAS

- [1] Russek, S. El origen de nuestras emociones, Recuperado: <https://www.crecimiento-y-bienestar-emocional.com/origen-emociones.html>, 2007.
- [2] Legg, T. What Is the Cannon-Bard Theory of Emotion?, Recuperado <https://www.healthline.com/health/cannon-bard>, on September 18, 2018.
- [3] Imagen de uso libre, recuperada en: <https://unsplash.com/>.
- [4] Picard, R. Affective Computing, 1995.
- [5] Hassouneh, A., Mutawa A. & Murugappan M., Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods, 2020.
- [6] Memar, M. & Mocaribolhassan, A., Stress level classification using statistical analysis of skin conductance signal while driving, 2021.

- [7] Saxena, A., Khanna, A. & Gupta D., Emotion Recognition and Detection Methods: A Comprehensive Survey, 2020.
- [8] Dalal, N y Triggs, B., Histograms of Oriented Gradients for Human Detection, 2010.
- [9] Viola,P y Jones, M., Robust Real-time Face Detection, 2015.
- [10] Kazemi, V. & Sullivan, J., One Millisecond Face Alignment with an Ensemble of Regression Trees. 2014.
- [11] Grove - GSR Sensor, Recuperado:
https://wiki.seeedstudio.com/Grove-GSR_Sensor/.