

AVANCES EN DESARROLLO DE UN SISTEMA PARA EL ESTUDIO DE LA RESPUESTA EMOCIONAL MEDIANTE EL ANÁLISIS DE LAS EXPRESIONES FACIALES

Nelson Dugarte¹; Antonio Alvarez¹; Buccheri Matías¹; Negman Alvarado¹; Marcelo Gomez².

¹ Centro Regional de Investigación y Desarrollo en Computación y Neuroingeniería (CeReCoN), Facultad Regional Mendoza (FRM) de la Universidad Tecnológica Nacional (UTN), Mendoza, Argentina / matiasbuccheri@gmail.com

² Cátedra BIOELECTRÓNICA, Departamento INGENIERÍA ELECTRÓNICA Facultad Regional La Rioja (FRLR) de la Universidad Tecnológica Nacional (UTN), La Rioja, Argentina / mgomez_ar@hotmail.com

Resumen: La capacidad para medir y comprender las emociones humanas de manera precisa y no invasiva es crucial en múltiples campos, desde la salud mental hasta el marketing y la interacción humano-computadora. En este artículo se presentan los avances en el desarrollo de un sistema para el estudio de expresiones faciales, empleando un método de análisis vectorial sobre los cambios que presenta el rostro del paciente ante una evaluación médica psicológica. Los vectores se obtienen midiendo la distancia entre un grupo definido de puntos virtuales. Estos puntos virtuales se capturan en cada fotograma del video que registra el rostro del paciente durante la consulta. La variación de distancia entre cada par de estos puntos en la secuencia de los fotogramas del video, captura un indicador dinámico de los cambios de plegado de cada sección del rostro. El método de análisis vectorial, se desarrolla con el propósito de generar índices que permitan al médico especialista contar con un método de aproximación que mejore la evaluación del estado emocional e incluso la respuesta psicosomática de las personas que se sometan a este estudio. Todo el sistema se integra en un formato de historia médica electrónica, diseñado bajo los estándares HL7, desde donde se puede adquirir el video del paciente, definir el número de puntos virtuales para realizar el análisis y generar los índices de resultado en cada consulta médica.

Palabras claves: estudio de la respuesta emocional – análisis de las expresiones faciales – Procesamiento con inteligencia artificial.

INTRODUCCIÓN

El análisis de expresiones faciales mediante inteligencia artificial (IA) ha emergido como un campo de estudio fascinante y prometedor en la intersección de la tecnología y las ciencias médicas (Chen, 2019). La IA, un campo de la informática que busca crear sistemas capaces de

realizar tareas que normalmente requieren inteligencia humana, tiene sus raíces en los años 1950, pero ha experimentado un crecimiento exponencial en las últimas décadas gracias a avances en poder computacional y disponibilidad de datos. En el contexto del análisis facial, la IA combina técnicas de visión por computadora y algoritmos de aprendizaje profundo, permitiendo la detección y medición precisa de los movimientos faciales y ofreciendo una ventana única a las emociones humanas (Merritt, 2023).

Históricamente, el análisis facial automatizado surgió en el ámbito del marketing y la publicidad, donde se utilizaba para evaluar las respuestas emocionales de los consumidores. Sin embargo, su potencial en aplicaciones médicas rápidamente captó la atención de investigadores y profesionales de la salud. La capacidad de cuantificar objetivamente las expresiones faciales ha abierto nuevas vías para el diagnóstico y seguimiento de condiciones neurológicas, psiquiátricas y del desarrollo.

En el contexto médico, la medición precisa de los movimientos faciales mediante IA ofrece ventajas significativas. Permite la detección temprana de anomalías sutiles, facilita el seguimiento longitudinal de pacientes, y proporciona datos objetivos para la evaluación de tratamientos (Chen, 2019). Esta tecnología no invasiva tiene el potencial de revolucionar campos como la neurología, la psiquiatría y la rehabilitación, ofreciendo una comprensión más profunda de la expresión emocional humana y sus alteraciones en diversos estados de salud. La aplicación de la IA en este campo demuestra cómo los avances tecnológicos pueden traducirse en herramientas prácticas y valiosas para la atención médica, mejorando nuestra capacidad de entender y tratar una amplia gama de condiciones relacionadas con la expresión facial y emocional.

METODOLOGÍA

Aun cuando los sistemas modernos con IA logran una buena detección de la emoción reflejada en el rostro, no siempre detectan el estado emocional verdadero (Chen, 2019; Pascual, 2022). Esto genera un importante nivel de error en la detección de patologías compulsivas psíquicas, que en algunos casos puede llegar a extremos como el suicidio o el asesinato.

En este proyecto, el sistema se diseñó para identificar la respuesta del rostro del paciente mientras experimenta un estímulo emocional. La idea

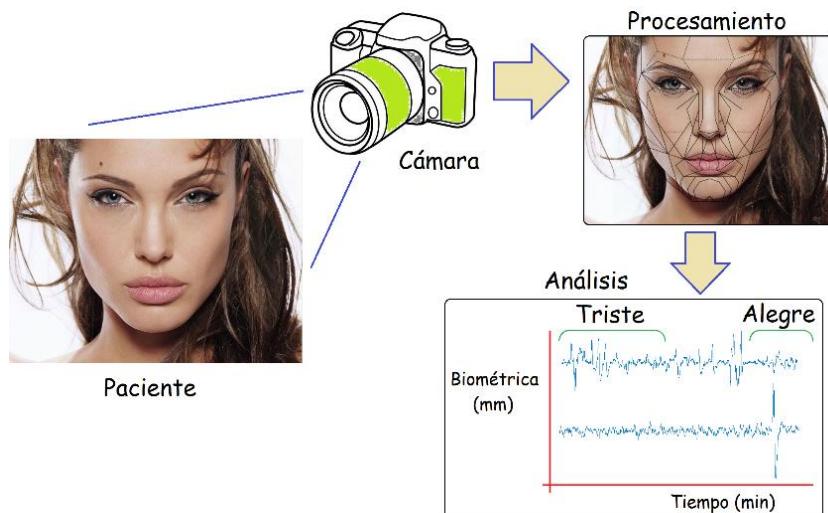
es explorar el estado emocional como respuesta a la empatía ante un video diseñado para generar cambios extremos entre emociones contrastantes. El procesamiento consiste en obtener valores métricos cuantitativos en formas de respuestas vectoriales, que sirvan como indicadores de la respuesta de empatía (Moya Albiol, 2018) reflejado por la emocionalidad del paciente y por consiguiente sirvan de referente al médico especialista para la detección temprana de patologías que pueden pasar desapercibidas.

El objetivo de este proyecto es desarrollar un sistema que permita desencadenar reacciones emocionales a partir de estímulos controlados. Los estímulos consisten en un video con secciones elaboradas con un guion que puedan cambiar el estado emocional del paciente. La idea, es analizar cada cuadro de imagen del video adquirido del paciente en relación sincrónica en el tiempo en el que visualiza los estímulos, generando múltiples vectores de parametrización que muestran los cambios de fisonomía causados por la empatía hacia los sectores del video estímulo.

El sistema está compuesto por una etapa de hardware y una de software. La Figura 1, presenta un esquema simplificado del diseño.

Figura. 1

Esquema simplificado del diseño del sistema.



El hardware utiliza una cámara de vídeo que toma la imagen del rostro del paciente mientras se somete a un estímulo visual. La etapa de software del sistema se desarrolla en plataforma libre, utilizando el software de programación Python (González Duque, 2011). Consta de tres partes: identificación de puntos característicos del rostro, medición de distancias entre los puntos seleccionados y análisis de los vectores que reflejan las facciones emocionales.

Estructura del estímulo visual

El video estímulo consiste en una película, cuya duración es de ciento ochenta y siete segundos. La película contiene una secuencia de 8 segmentos cortos independientes, que no guardan ninguna relación unos con otros. El objetivo individual de cada segmento es presentar una serie de estímulos emocionales definidos por la empatía a la persona que lo está viendo.

Para este proyecto, el video estímulo sólo toma en cuenta dos emociones: alegría y tristeza. La película alterna segmentos que pueden producir alegría con segmentos que podrían generar tristeza, intercalados por segmentos cortos que muestran una serie de imágenes dinámicas no biológicas. Denominamos a estas últimas 'imágenes neutras' porque no debieran aportar estímulos emocionales adicionales. El objetivo que se persigue es captar la respuesta emocional del paciente mediante la grabación de un video de su rostro mientras visualiza el video estímulo.

Procesamiento del video de respuesta

El video de respuesta emocional (VRE) se toma del rostro del paciente mientras visualiza el video estímulo. Para el almacenamiento de la información, se creó una base de datos diseñada para que funcione como historia médica electrónica (HME). Esta base de datos, permite relacionar el paciente que se somete al estudio con cada una de las adquisiciones que se realicen.

El procesamiento de la información, consiste en un algoritmo que realiza una secuencia automatizada en el computador, sobre cada cuadro de imagen captado en el VRE del paciente.

El software detecta la forma del rostro dentro del marco del video, para posteriormente identificar la posición donde se ubica cada rasgo de dicho rostro. Para esto, el software detecta: los bordes de la cara, las

cejas, los ojos, la nariz y la boca, asignando a cada una de estas zonas mencionadas un número de puntos específicos que resultan clave para el estudio de los gestos en el rostro. La detección de puntos característicos consiste en ubicar y fijar 68 marcas virtuales sobre el rostro del paciente en una matriz de coordenadas. La Figura 2, muestra un fotograma donde se visualiza el efecto que fija los puntos virtuales sobre el rostro del paciente.

Lo que se pretende es que cada punto esté fijo a una sección del rostro y por consiguiente se desplace en el espacio siguiendo los movimientos de cada sección de forma específica. Esto permite medir los cambios de contorno de la piel en el rostro entre un fotograma y el siguiente. Para ubicar automáticamente los puntos deseados en cada fotograma, se utilizó un método proporcionado con licencia libre, que utiliza técnicas de machine learning (IA) programadas en lenguaje Python (Google AI for Developers, 2024).

Por sí solos, los puntos detectados sobre la imagen del rostro, no representan utilidad en la investigación. La información médica relevante se determina con el análisis de cómo cambian las distancias entre los puntos con respecto al tiempo.

Figura 2

Captura automática de puntos virtuales sobre el rostro del paciente.



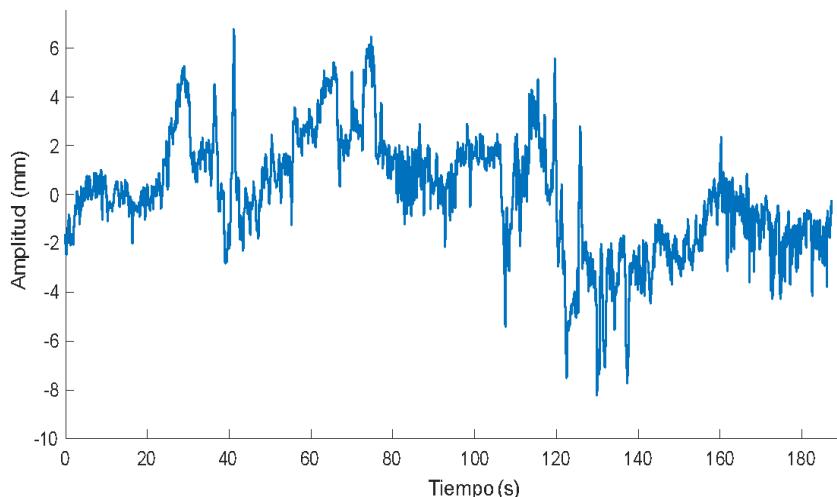
Análisis de la respuesta emocional del paciente

Los cambios del rostro en la secuencia de imágenes, se obtienen con la medición de distancias entre los puntos característicos. La secuencia de mediciones en la sucesión de imágenes del video entre dos puntos específicos determina un vector representativo del movimiento en la sección del rostro cubierta entre estos puntos, y por consiguiente puede mostrar el estado del paciente en función del mapa emocional de Ekman (Ekman and Oster, 1981).

En esta investigación se han elegido 27 pares de puntos para estudiar los movimientos del rostro, específicamente para identificar la respuesta de los estados emocionales de felicidad y de tristeza. Las mediciones entre estos pares de puntos representan 27 vectores correspondientes. La Figura 3, presenta la gráfica de uno de los vectores adquiridos. Este vector muestra la respuesta entre los puntos adyacentes en la comisura de los labios y el mentón del paciente.

Figura 3

Respuesta entre la comisura de los labios y el mentón del paciente.



RESULTADOS

Para obtener resultados cuantitativos del estado emocional del paciente es necesario comparar cada vector adquirido del vídeo estímulo con algún tipo de patrón de referencia. Dado que no existe un patrón de referencia de uso estándar de este tipo, fue necesario generar un patrón aproximado de forma práctica.

Para encontrar un patrón indicativo en cada vector, se realizó el análisis de la adquisición de 50 pacientes control. El patrón de referencia se establece promediando los intervalos correspondientes a los sectores donde los pacientes visualizaron cada estímulo emocional. Esto genera una señal patrón que sirve de referencia para estimar el estado emocional de felicidad y otro para estimar el estado de tristeza.

Los vectores patrón son más cortos que los vectores originales porque corresponden exclusivamente a los segmentos donde se evalúa un estado emocional específico.

Cuando se analiza un vector cualquiera correlacionándolo con uno de los vectores patrón se nota una elevada correspondencia con los segmentos donde el paciente refleja un estado emocional acorde con el patrón y viceversa donde el estado emocional es opuesto.

La correlación total entre el vector patrón y el vector en estudio se logra por medio de un procedimiento que desplaza el patrón sobre toda la extensión de la señal en estudio realizando la comparación entre ambas señales. El procedimiento inicia alineando el patrón con una sección del mismo tamaño al inicio del vector en estudio, la comparación entre señales se realiza calculando el coeficiente de Pearson en esa alineación (Fiallos, 2021). Luego, se desplaza el patrón en una muestra sobre el vector en estudio. Esto se repite hasta completar todas las muestras del vector en estudio, con lo cual se genera un nuevo vector con los resultados de cada comparación. El coeficiente de Pearson (r_{xy}) se puede calcular por medio de la relación dada por la Ecuación (1).

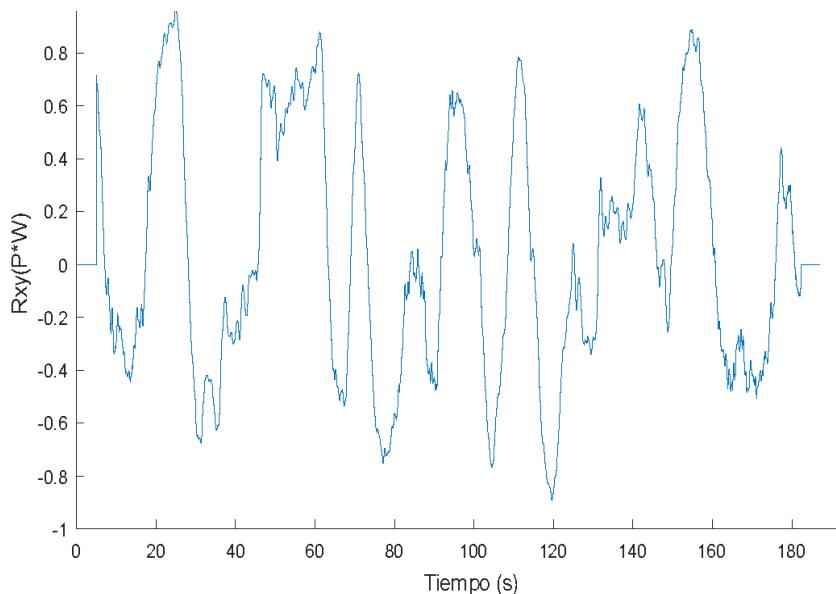
$$r_{xy}(i) = \frac{cov(X,Y)}{\sqrt{\sigma_x^2 * \sigma_y^2}} \quad \text{donde: } -1 \leq r_{xy}(i) \leq +1 \quad (1)$$

La ecuación 1, establece la comparación entre las señales X e Y. Donde X está determinada por el vector patrón e Y representa el vector en estudio. La Figura 4, muestra la gráfica de los valores $r_{xy}(i)$ que se

obtienen al comparar el patrón de felicidad con la señal obtenida como respuesta de la comisura de los labios al mentón en el estudio de un paciente en particular.

Figura 4

Gráfica de $r_{xy}(i)$ entre el patrón de felicidad y la señal obtenida como respuesta de la comisura de los labios al mentón.



Los resultados de $r_{xy}(i)$ muestran con valores más cercanos a 1 donde la señal que se analiza es más parecida al patrón de felicidad y con valores aproximados a -1 donde posee mayor diferencia. La gráfica de la Figura 4 evidencia con valores más positivos la correspondencia de la señal del paciente con los sectores de felicidad del vídeo estímulo, y denota con valores negativos la correspondencia con los sectores donde el paciente visualizaba estímulos que le causaron tristeza.

Actualmente se trabaja en un software de análisis que interprete los 27 vectores seleccionados para generar un índice más eficiente al momento de evaluar el estado de emocionalidad del paciente.

CONCLUSIONES

La implementación de un sistema de análisis vectorial para el estudio de expresiones faciales representa un avance significativo en la intersección de la tecnología y la medicina. Este enfoque innovador proporciona a los profesionales de la salud una herramienta objetiva y cuantificable para evaluar las respuestas emocionales de los pacientes, lo que podría revolucionar campos como la psicología, la neurología y la psiquiatría.

Las afirmaciones sobre la eficacia del método de análisis vectorial en la captura de indicadores dinámicos de cambios faciales son prometedoras. La capacidad de medir con precisión las variaciones en las distancias entre puntos faciales específicos a lo largo del tiempo ofrece una nueva dimensión en la comprensión de las expresiones emocionales. Este enfoque no solo permite una evaluación más detallada y objetiva de las respuestas emocionales, sino que también podría facilitar la detección temprana de trastornos psicológicos y neurológicos.

AGRADECIMIENTOS

Los autores de este trabajo desean agradecer al grupo GENESIS y al Instituto Regional de Bioingeniería (IRB) suscritos del Centro Regional de Investigación y Desarrollo en Computación y Neuroingeniería (CeReCoN), de la Universidad Tecnológica Nacional de la República Argentina, por su colaboración científica y tecnológica en el desarrollo del proyecto. Así mismo, se hace extensivo el agradecimiento a todas las personas e instituciones que han permitido que este proyecto sea factible.

REFERENCIAS

- Chen, A. (2019, julio). La IA es incapaz de identificar emociones a partir de la expresión facial. *MIT Technology Review*. <https://www.technologyreview.es/s/11336/la-ia-es-incapaz-de-identificar-emociones-partir-de-la-expresion-facial>
- Ekman, P. & Oster, H. (1981). Expresiones faciales de la emoción. *Studies in Psychology: Estudios de Psicología*, 2(7), 115-144. <https://doi.org/10.1080/02109395.1981.10821273>
- Fiallos, G. (2021). La Correlación de Pearson y el proceso de regresión por el Método de Mínimos Cuadrados. *Ciencia Latina Revista Científica Multidisciplinaria*, 5(3), 2491-2509, https://doi.org/10.37811/cl_rcm.v5i3.466.

- González Duque, R. (2011). *Python para todos*. <https://persoal.citius.usc.es/eva.cernadas/informaticaparacientificos/material/libros/Python%20para%20todos.pdf>.
- Google AI for Developers (21 de junio de 2024). *Guía de detección de puntos de referencia faciales*. Google AI Edge - Soluciones https://ai.google.dev/edge/mediapipe/solutions/vision/face_landmarker?hl=es-419
- Merritt, R. (2023, febrero 3). *¿Qué Es la Computación de IA?*. Publicaciones NVIDIA. <https://la.blogs.nvidia.com/blog/que-es-la-computacion-de-ia/>
- Moya Albiol, L. (2018). *La empatía: Entenderla para entender a los demás* (1ra ed.). Plataforma Editorial S.L.
- Pascual, M. G. (14 de enero de 2022). El algoritmo se equivoca. La cara no siempre dice lo que uno siente. *El País*. <https://elpais.com/tecnologia/2022-01-15/el-algoritmo-se-equivoca-la-cara-no-siempre-dice-lo-que-uno-siente.html>

* * * * *