



La ingeniería de las emociones humanas



Franyelit Suárez
Luis Rosales
Omar Flor

Autores:

Franyelit Suárez

Luis Rosales

Omar Flor



La Ingeniería de las Emociones Humanas



La ingeniería de las emociones humanas

© Franyelit María Suárez Carreño, Luis Dionisio Rosales Romero, Omar Cristóbal Flor Unda ©

Universidad de Las Américas Facultad de Ingeniería y Ciencias Aplicadas, Ingeniería Industrial. Campus Queri José Queri www.udla.edu.ec Quito, Ecuador

Universidad Nacional Experimental Politécnica Antonio José de Sucre, Vice Rectorado Puerto Ordaz. Venezuela.

Primera edición: Noviembre, 2020

Cuidado de la edición; AutanaBooks del Ecuador

Corrección y estilo: Erka Nessi

Diseño de cubierta: Manuelis Salazar

Diseño gráfico y diagramación: Manuelis Salazar

Editorial: AutanaBooks

ISBN: 978-9942-38-662-5

Publicación en Línea

ISBN: 978-9942-38-662-5



Reservados todos los derechos. Ninguna parte de este libro puede ser reproducida de cualquier forma o por cualquier medio, incluidos los sistemas de almacenamiento, recuperación electrónica, excepto con la autorización y permiso explícito, previo y por escrito del editor, para parafraseo, extractos y fines académicos. Esta obra cumplió con el proceso de revisión por pares anónimos, durante los meses de mayo de 2019 a junio de 2020.



Los Autores



Franyelit Suárez, Doctora en Ciencias de la Ingeniería. Universidad de las Américas, Quito- Ecuador. Facultad de Ingeniería y Ciencias Aplicadas, Ingeniería Industrial.

Investigadora en el área de Ciencias de la Ingeniería, Física computacional, inteligencia artificial.



Luis Rosales, Doctor en Física. Universidad Nacional Experimental Politécnica Antonio José de Sucre, Puerto Ordaz, Venezuela.

Investigador en el área de Ciencias de la Ingeniería, Física computacional, simulación y modelación. Coordinador de Doctorado en Ciencias de la Ingeniería.



Omar Flor, Master en automática, robótica y telemática. Universidad De las Américas, Quito-Ecuador. Facultad de Ingeniería y Ciencias Aplicadas, Ingeniería Industrial.

Investigador en el área de Automatización, Diseño de prototipos y robótica. Especialista



Tabla de contenido

Capítulo 1: Las emociones y el estrés	10
Capítulo 2: La ingeniería en el campo de las emociones	20
Fundamentación Metodológica	20
Enfoque Epistemológico	20
Diseño de Experimento	22
Diseño de la Investigación	22
¿Qué aporta este trabajo al conocimiento y a la investigación sobre el tema?	22
¿Qué tipo de conocimiento y de qué naturaleza se ha generado con esta investigación?	23
Fundamentación Matemática	23
Análisis de comportamiento	26
Modelado matemático	27
Diseño del algoritmo	31
El espectro de voz	32
La voz y los estados emocionales	33
El estrés y la voz	34
El estrés y la expresión facial	35
Capítulo 3: Algoritmos de evaluación y clasificación	36
Las redes neuronales artificiales	36
Redes neuronales convolucionales	38
Redes de memoria a corto plazo	38
Algoritmo K-Nearest Neighbors	40
Algoritmo SVM	41
Algoritmo HOG	43
Algoritmo LBP	44



Fases para el desarrollo del algoritmo	45
Capítulo 4: Pruebas computacionales	46
Estimación de emociones mediante análisis de imagen facial	49
Selección de la base de datos de imágenes faciales	50
Estimación de emociones mediante análisis de voz	56
Pruebas y funcionamiento	60
Capítulo 5: Hallazgos de la ingeniería en la detección de emociones.	63
Computación inteligente en el análisis de emociones humanas	63
Análisis de emociones a través de la voz	65
Análisis de fisionomía del rostro	66
Identificación del rostro dependiendo de la distancia	69
Iluminación del ambiente en la captura de la imagen del rostro.	69
Análisis del espectro de voz	69
Optimización de los algoritmos	70
Validación	74
Pruebas con especialista médico	74
Entorno de trabajo	74
Análisis Estadístico	77
Otros resultados de medición de estrés	79
Delimitación	82
Bibliografía	84



Prólogo

Quizás todos tenemos una idea de lo que significa escribir un libro, anécdotas sobre algunos autores cuando se enfrentaron a la escritura de alguna de sus obras son bien conocidas; por ejemplo, Ray Bradbury decidió aislarse en el sótano de la biblioteca de la Universidad de California y utilizar una máquina de escribir alquilada por 10 centavos la media hora para escribir su novela Fahrenheit 451. Otra anécdota interesante es la escritura de Cien Años de Soledad de Gabriel García Márquez, quien lo dejó todo a cargo de su esposa incluyendo el sustento económico de su familia para encerrarse con sus pensamientos para la escritura de esta obra que ahora reconocemos como impercedera.

En ambos casos estos autores tuvieron la fortuna y el privilegio de aislarse voluntariamente para escribir sus obras; en este caso y a diferencia de los autores antes mencionados, nuestra autora tiene el enorme mérito de, incluso sin tener a disposición ese invaluable recurso, brindarnos una obra del mayor rigor científico, librando el día a día de las profesionales actuales y todo lo que ello implica, siendo madre, hija amiga y además, inmigrante. Una labor como la que implica un libro de esta envergadura nos da la dimensión del esfuerzo inquebrantable de su autora, de las horas a destiempo de investigación, de contraste de resultados, de evaluaciones técnicas y de simulaciones informáticas, de escribir y reescribir párrafos, de las luchas contra el sueño y la fatiga. Franyelit Suárez es una representación de la lucha de las circunstancias humanas contra una voluntad inquebrantable, donde el resultado es la victoria del espíritu, del intelecto y del amor a la ciencia. He sido un testigo privilegiado de la tenacidad irreductible de la autora para brindarnos esta obra que tiene todas las características de convertirse en referencia obligada en el estudio de la ingeniería aplicada a las emociones humanas y la manera en que la ingeniería puede ser un apoyo a los tratamientos médicos enfocados en el estrés.

Te invito a descubrir este libro como lo que es, un aporte a la salud del presente y del futuro.

Ángel Lezama



Capítulo 1: Las emociones y el estrés.

En este libro se presenta el desarrollo de un algoritmo para el reconocimiento de emociones a partir de las variables de rostro y voz. Las emociones son parte esencial de las personas, sin embargo estas comenzaron a estudiarse en la segunda mitad del siglo XX [1]. Actualmente las emociones humanas son consideradas por un número importante de investigadores de diferentes disciplinas, convirtiéndolo en un activo prolífico en las áreas médicas, neurociencias, biológicas, sociológicas y económicas. Las emociones constituyen una respuesta compleja del organismo donde convergen aspectos mentales, fisiológicos y motrices [1], [2].

Diferentes desarrollos se han planteado en torno al reconocimiento de emociones [3], [4], [5], [6], en los que una amplia variedad de herramientas software han destacado para la consolidación de propuestas que permiten estimar las reacciones emotivas del organismo. Estos desarrollos han mostrado importantes conclusiones en torno al uso de la tecnología en situaciones de salud, así como la valoración de las principales variables que influyen en la detección de emociones.

Las emociones son una reacción natural del organismo, en 1872, Charles Darwin [7] publicó los resultados de uno de los trabajos más controversiales de los últimos cien años, en el que expuso la recopilación de información de sus treinta años de estudios sobre las emociones. En estos resultados Darwin manifestó dos ideas fundamentales; las emociones son innatas y universales, y que las emociones son producto de la evolución. Darwin postuló que las expresiones faciales son un reflejo de las emociones y que estas no son exclusivas de los humanos [8]. Sin embargo, en las referencias [9], [10] se afirma que la sonrisa y el llanto son exclusivas de los humanos, como manifestación de las emociones como mecanismo de adaptación de la especie [11].

Las emociones pueden generar diferentes respuestas en el organismo [2], tanto favorables como adversas, entre los que se cuentan la alegría, tristeza, enfado, desagrado, entre otros. Estas emociones están sujetas a las condiciones sociales y culturales, que afectan la conducta de las personas y por ende, influyen en la percepción de la realidad que desencadenan dichas emociones [2], [12], [13].

Tomando en cuenta lo expuesto en [1] y [14], en este trabajo se esta-



blecen las emociones básicas, que a partir de sus compuestos dan origen a emociones secundarias, así entonces se lista la ira, asco, miedo, alegría, tristeza, sorpresa y emoción nula. De modo que una emoción secundaria podría ser por ejemplo el espanto, producto del miedo con la sorpresa.

La detección de emociones de manera no invasiva, es considerada en este trabajo como eje fundamental para el desarrollo del algoritmo. Por tanto, se ha considerado la variable de fisionomía facial para evaluar las características asociadas a estados emocionales que se reflejan en las líneas básicas de expresión y que pueden ser analizadas por un sistema computacional inteligente, que se soporta en una base de datos con más de 200 imágenes para estimar y categorizar las emociones. Así mismo se ha considerado el análisis del espectro de voz, cuyos autores manifiestan es una de las variables del organismo más difíciles de evadir ante la presencia de emociones [15].

Las emociones que alcanzan niveles intensos y se presentan con mucha frecuencia tienden a producir cambios en la conducta, de manera que se dejan a un lado los aspectos saludables y se empiezan a desarrollar conductas adictivas que afectan el estado de salud, desencadenando reacciones fisiológicas que determinan el estrés [16].

Las emociones están relacionadas al estrés, y por ende los efectos del estrés afectan tanto en la salud física y mental como en el rendimiento laboral y desempeño social de las personas [17]. El problema del estrés está asociado a los requerimientos de la modernidad, centrada en la búsqueda de alternativas para una mejor calidad de vida, y por ende no se evalúan las consecuencias de la salud física y mental de las personas afectadas [17], [18]. Los estímulos estresores son factores que producen una reacción fisiológica en el organismo, y como respuesta, se activa el eje hipofisoadrenal y el sistema nervioso vegetativo [19], [20], produciendo la liberación de hormonas responsables de los cambios en el organismo [21], [22].

Esta liberación de hormonas, solo puede detectarse a través de mecanismos invasivos, sin embargo producen reacciones en la fisionomía del rostro que pueden ser detectadas por algoritmos inteligentes capaces de procesarlos para lograr una estimación aproximada del nivel de estrés. Estos sistemas se basan principalmente en Redes Neuronales Artificiales, de distintas categorías, que pueden aprender a través de patrones y ofrecer resultados satisfactorios. Se presenta entonces un desarrollo basado en una Máquina de Soporte Vectorial, que emplea para su funcionamiento Redes Neuronales Artificiales, y que permite la estimación de resultados confiables [23].

Los algoritmos inteligentes ofrecen la posibilidad de trabajar con variables no lineales, que en el caso de la medicina resulta común. Además es posible combinarlos para lograr resultados más óptimos, así como es via-



ble la utilización de bases de datos confiables para lograr entrenamientos y aprendizajes apropiados que ofrezcan mejores resultados. En este trabajo se han utilizado también redes Adams para la optimización de datos.

Definir las características del estrés desde el punto de vista de ingeniería, resulta de gran relevancia para asumir posibles estrategias de aplicación de un algoritmo de estimación de emociones. Esto conduce a la generación de una línea de investigación que pudiera ser considerada para diversas áreas profesionales. La definición de modelos de ingeniería en la estimación de emociones se llevó a cabo a partir del estudio de las características clínicas de estrés, y su efecto sobre la salud de las personas.

El estrés ha sido un tema de interés y de inquietud para diversos científicos de la conducta humana [24], por cuanto sus efectos inciden tanto en la salud física y mental, como en el desempeño laboral y académico del individuo. Puede provocar preocupación y ansiedad conduciendo a trastornos personales, desórdenes familiares e incluso sociales [24], [25].

La situación actual de las organizaciones ha conducido a situaciones de estrés en el colectivo de los trabajadores de las empresas de producción y servicios [26], siendo ésta la principal causa en la proliferación de enfermedades a nivel del sistema cardiovascular, sistema gastrointestinal y sistema dérmico [16], [24]- [27]. Estas enfermedades y otras dolencias producidas por el estrés afectan de manera indirecta la eficiencia de las personas y por ende la productividad en las organizaciones, representando un aumento del ausentismo laboral por reposos médicos debido a accidentes de trabajo y/o enfermedades ocupacionales [28], [17], [29].

Muchas de las experiencias de estrés laboral surgen cuando las demandas del trabajo son altas, y a su vez los recursos en ese entorno son bajos [30]. Así mismo otros autores [30]- [31] señalan que existen principalmente cuatro posibles situaciones laborales en relación a las exigencias y a los recursos; aquellos donde los recursos son escasos y las exigencias son elevadas, aquellos con bajos niveles de exigencias y bajo niveles de recursos, y los menos estresantes serían aquellos donde hay altos niveles de demanda pero con elevados niveles de recursos, y donde hay altos niveles de recursos y pocas exigencias.

Las nuevas tecnologías, las nuevas exigencias laborales y las necesidades de globalización, en consonancia con los cambios socioeconómicos y sociopolíticos causan cada vez más impacto en los mercados laborales [32]. El trabajo mental y emocional así como la poca flexibilidad laboral, las necesidades sociales cada vez más exigentes, y las nuevas formas de actividad laboral, son algunos de los desencadenantes de situaciones emocionales intensas y frecuentes que conducen al estrés en el trabajo [31]- [33].



Un algoritmo para evaluación de estrés podría contribuir a la asistencia médica, así como a los servicios médicos laborales, para evitar situaciones de salud que afecten al desempeño de los trabajadores o que afecten la vida social y familiar de los mismos. A su vez podría contribuir a las mejoras en los objetivos de las organizaciones, desde la perspectiva de la salud de sus empleados.

Los estresores laborales pueden ser muy diversos, varían desde su frecuencia, intensidad, duración y predictibilidad [34]. En función del análisis de contenido es posible establecer ocho categorías de estresores laborales [32], [34]: los relacionados al entorno físico y el ambiente, los a la organización laboral, con el puesto y las funciones del cargo, los causados por el rol que se desempeña, aquellos que se refieren a las relaciones personales, los méritos laborales, la estructura y el clima laboral de la organización, las relaciones trabajo-familia.

En el último año han surgido importantes estudios en torno al estrés [35], y diversos organismos han evaluado la presencia o no de las emociones estresantes, en entornos laborales, familiares y sociales. Por lo antes planteado, algunos autores han desarrollado cuestionarios para evaluar el estrés [36], [37], en su mayoría con resultados confiables, que ofrecen una visión ampliada de los altas cifras de personas estresadas que han desmejorado su salud.

En base a las preocupaciones que el estrés puede causar en la población, también se han desarrollado sistemas electrónicos para su detección [38]. Estas investigaciones señalan que algunas de las variables que pueden incidir en las manifestaciones de estrés son la frecuencia respiratoria, la actividad muscular del ojo, la conductividad de la piel y la presión arterial. En [38] han logrado medir todas estas variables y realizar una implementación con dispositivos arduinos que permiten evaluar dichos parámetros y arrojar resultados al respecto del estrés, adicional se incluyeron máquinas de soporte vectorial que ofrecen resultados confiables para discriminar la presencia o no de estrés, específicamente estrés matemático que los autores catalogaron como uno de los factores de mayor estrés en la población de estudio.

Otro método utilizado en la detección de estrés ha sido la visualización hiper-dimensional [4], que combina el escalamiento multidimensional y la realidad virtual. Este método utilizado arrojó resultados empíricos que evidencian la relevancia de las variables de pulso en la detección de estrés, así como la importancia en la reducción del error al emplear métodos poco tradicionales para dicha detección. El sistema desarrollo además permite distinguir los modos de funcionamiento psicológico en las personas, pudiendo detectar incluso la mentira como una emoción conducente al estrés.



La recepción de señales eléctricas del organismo para la detección de estrés es otro procedimiento utilizado por los investigadores [39]. Este método capta las señales eléctricas cerebrales a través de electrodos en los lóbulos frontal y temporal, las cuales son procesadas por amplificadores operacionales que en conjunto con un microprocesador realizan el análisis de las variables, que puede más adelante ser validado por el especialista médico. En la referencia [39] se consideró el estudio de las señales alfa, beta, theta y delta que se producen en el lóbulo frontal, que es el encargado de las expresiones de conducta emocional. Los resultados mostraron que el procesamiento de estas señales requiere un mayor rigor, ya que el método utilizado agrega ruidos e interferencias producto de las señales adicionales que podrían encontrarse en el ambiente, como señal wifi, luz eléctrica, entre otras. Sin embargo el uso de las señales alfa y beta para la clasificación de estrés, fue bastante significativo y permitió la detección de los estados de relajación y excitación propios del estrés.

El uso de la ingeniería en los procesos biológicos humanos se presenta como una amplia tecnología médica que da solución a problemas de salud. Una de las aplicaciones computacionales más comunes para este tipo de análisis, es el uso de la computación inteligente, más específicamente de las redes neuronales artificiales, que ofrecen una amplia capacidad de respuestas para el estudio de los sistemas biológicos [40]. Otros métodos, como la regresión logística y la realidad virtual, han brindado resultados satisfactorios en los escenarios de salud ocupacional y medicina en general [41]- [42]. En este trabajo se muestra que la computación inteligente aporta a los procedimientos no invasivos para lograr categorizar los estados de estrés, clasificarlos y evaluar situaciones médicas a partir de los resultados obtenidos.

La computación inteligente, también llamada computación evolutiva y su aplicación en los procesos de ingeniería y medicina han permitido la vinculación de los procesos naturales a los sistemas electrónicos [42], [43]. La evolución ha inspirado modelos de cómputo asociados al comportamiento de los organismos, provocando el desarrollo de nuevas formas de procesamiento de la información y manejo de datos [42], [44]. Este tipo de sistemas se propone como soluciones a problemas médicos que garanticen procedimientos no invasivos con un mínimo margen de error.

Las aplicaciones inteligentes, constituidas dentro de la computación evolutiva, permiten dar respuestas a una amplia variedad de situaciones médicas de difícil categorización por métodos tradicionales. La inteligencia artificial encierra todos los procesos de computación evolutiva, que a su vez involucra las aplicaciones inteligentes e incluye técnicas como las redes neuronales artificiales, la optimización con enjambres de partículas o la misma compu-



tación evolutiva, es decir, estrategias basadas en la inteligencia que emerge de las interacciones en un colectivo, sean neuronas, partículas u organismos [44], [45].

El campo de la medicina es ejemplo paradigmático de dominio incierto, donde la información implícita, la multiplicidad, la ambigüedad y la imprecisión hacen imprescindible el tratamiento de la incertidumbre [46], [47]. La incertidumbre es, junto con el conocimiento y el aprendizaje, uno de los problemas fundamentales de la inteligencia artificial y de sus complementos en la computación inteligente [47].

El estudio de algoritmos inteligentes en áreas de la medicina y la salud, es importante para el bien clínico porque tanto la medicina como la salud representan procesos de la vida diaria complejos, para los cuales resulta difícil interpretarlas y reconocerlas a simple vista o mediante algún procedimiento matemático clásico. Diversas investigaciones, donde se muestra el análisis, identificación y clasificación de este tipo de variables médicas y procesos clínicos han sido publicadas y entre las cuales se puede mencionar los siguientes: Análisis espectral de parámetros fisiológicos para la detección de emociones, modelo híbrido para el diagnóstico de enfermedades cardiovasculares basado en inteligencia artificial, árboles de decisión como herramienta en el diagnóstico médico, inteligencia artificial para asistir el diagnóstico clínico en medicina, etc. Todos estos procesos demandan algoritmos inteligentes para resolver su naturaleza no lineal [22], [27], [29].

Caracterizar variables no lineales requiere de un procesamiento computacional específico, reconocer las variables físicas y fisiológicas, y en la mayoría de los casos se requiere la validación de expertos. Esto ha conducido que en tiempos recientes se incorporen técnicas de inteligencia artificial para desarrollar sistemas, donde con el apoyo de expertos se describan características intrínsecas de las personas, en combinación con el análisis psicológico, lográndose así un sistema adecuado para diferentes aplicaciones [18]- [29]. La asociación de análisis psicológico en combinación con métodos inteligentes aplicados a la medicina, ha hecho posible el desempeño eficiente de sistemas que antes no eran posible poner en práctica [18]- [31].

Estos sistemas cuyas variables son no lineales y de gran complejidad numérica, demandan algoritmos inteligentes de control difuso y redes neuronales artificiales, que permitan que el sistema tenga un resultado satisfactorio, accediendo así al alcance de los objetivos que hasta el momento no se han logrado con técnicas estadísticas clásicas. Esta eficiencia se refiere, además de otros, al seguimiento de recorridos y caracterización de variables de forma eficiente, a la optimización de valores y procesos inciertos, a la depuración de perturbaciones, y a la asociación de parámetros fisiológicos



con parámetros actitudinales [24]- [36].

Una amplia gama de técnicas inteligentes son utilizadas para la caracterización de variables fisiológicas humanas, la clasificación de expresiones emocionales y el reconocimiento de estrés [28], [39], [38], [41]. Entre otras pueden mencionarse las siguientes: Robotic gesture recognition [37], Robust Real-Time Face Detection [38], Burnout research in the social services: a critique [4], Maslach Burnout Inventory [21]. Career Burnout. Causes and cures [40], etc. Estas técnicas de manera individual no producen precisión en los resultados por la complejidad de las variables asociadas, como las indicadas en [9], [22], [26], [34]. Combinando apropiadamente metodologías médicas convencionales con técnicas inteligentes, específicamente lógica difusa del tipo Mandani y redes neuronales del tipo adaptativa, es factible lograr los resultados esperados cuando se trata de manejar este tipo de variables complejas [1]- [11], [26].

Partiendo de estas premisas, en este libro se describe un algoritmo de estimación de emociones utilizando Redes Neuronales Artificiales con un algoritmo de Gradiente Orientado, como una herramienta para el tratamiento médico preventivo, que permite la evaluación permanente de las personas para atender su estado de salud. Este tipo de sistema pretende aportar considerablemente al tratamiento de enfermedades crónicas, enfermedades del sistema cardíaco, del sistema dérmico y gastrointestinal, donde se ha visto mayor prevalencia en presencia del estrés.

Se valoran las ecuaciones dinámicas asociadas al estrés, que permiten reconocer la importancia de los estresores en el organismo y su impacto en las enfermedades. Los autovalores del sistema linealizado ofrecen un equilibrio inestable tipo silla, y confirman que la prevalencia de estresores es un factor fundamental para la presencia de enfermedades. Este trabajo se centra en el análisis de emociones que desencadenan el estrés, y se evalúa la presencia o no de estrés en las personas, sin considerar las especificidades de las enfermedades causadas ni otros aspectos del estrés.

Debido al gran crecimiento de las máquinas y el avance de la tecnología en la sociedad actual, se hace necesaria la integración de las mismas para dar solución a situaciones humanas que requieren especial atención, como es el caso de las emociones y las situaciones estresantes, que pueden ocasionar importantes consecuencias negativas en el desempeño de las personas y en la producción de las empresas e instituciones. Actualmente es posible utilizar la inteligencia artificial para simular sistemas reales de la conducta humana y obtener resultados óptimos con el mínimo error, entre las más comunes se encuentran las redes neuronales artificiales y los sistemas de optimización por enjambre de partículas [48]. De esta forma se pueden generar herra-



mientas que faciliten la evaluación médica y permitan la caracterización de los parámetros actitudinales en las personas para asegurar que el desempeño y las funciones que deba cumplir en determinado cargo se ejecutan de manera ideal. Fadin y Álvarez [38] plantean la caracterización de actitudes humanas a partir de la implementación de la transformada de coseno discreta y el perceptrón multicapa, como herramientas híbridas para la determinación de factores de difícil percepción, como la segmentación de imágenes.

En este libro se definen y aplican algoritmos inteligentes de reconocimiento de estrés en las personas como una herramienta para la evaluación médica preventiva. Se documentan los criterios que permiten proyectar las variables que afectan la conducta humana y se definen los sistemas en el marco de la inteligencia artificial aplicada en la ingeniería y la medicina. Además, se eligen los sistemas inteligentes integrados a los requerimientos médicos y se establecen los algoritmos inteligentes de redes neuronales artificiales para la caracterización de las emociones humanas. Finalmente, se generan y verifican algoritmos, mediante la técnica de autoconsistencia, para el estado emocional de las personas a través de un sistema inteligente accesible a los usuarios.

La investigación es aplicada tecnológica, ya que se generan conocimientos que se pudiesen poner en práctica en el sector salud. Es cuantitativa, ahonda en los fenómenos a través de la recopilación de datos y fundamentos matemáticos. Es cuasi experimental, controla solo algunas variables del fenómeno a estudiar.

Se utilizan modelos de inteligencia artificial basados en reconocimiento de patrones y morfología facial, que caracterice las expresiones humanas. Mehrabian [16] sostuvo que en la transmisión de mensajes con carga emocional la parte verbal (palabras) contribuye sólo un 7% en la arquitectura del texto o mensaje, mientras que la fonética (rasgos fónicos que afectan a la métrica de la voz) contribuye un 38% y las expresiones faciales del individuo lo hacen en un 55%. Adicional a ello, Ekman [4] señaló que las siete emociones principales se pueden reconocer en las expresiones faciales de manera universal, y que éstas pueden depender de la cultura, pero fundamentalmente tienen origen biológico, son: ira, asco, miedo, alegría, tristeza, sorpresa y la emoción nula. Así mismo, para desarrollar sistemas inteligentes que interactúen naturalmente con los seres humanos, se debe implementar el reconocimiento de emociones humanas a partir de las expresiones faciales [14].

El trabajo está centrado en diseñar un algoritmo general aplicable en sistemas inteligentes utilizando una herramienta computacional matemática para que sea capaz de reconocer las emociones humanas a partir de las variables fisiológicas y de las expresiones faciales y la adquisición de datos reales



en las experiencias de campo de las áreas de aplicación.

Desde la invención del polígrafo, el detectar las emociones del ser humano así como interpretarlas y reconocerlas a simple vista o mediante algún procedimiento ha sido una necesidad social. Esta actividad se ha desarrollado utilizando encuestas metodológicas, arboles de decisión y tratamientos matemáticos clásicos. Tal como se refieren en trabajos previos que muestran el análisis, identificación y clasificación de este tipo de variables como son análisis espectral de parámetros fisiológicos para la detección de emociones, modelo híbrido para el diagnóstico de enfermedades y basado en inteligencia artificial, entre otros métodos. Todos estos procesos demandan algoritmos inteligentes para resolver su naturaleza no lineal con el enfoque metodológico descrito, en tal sentido, se propone un método digital y computacional para estas detecciones

Este trabajo está centrado en diseñar un algoritmo computacional selectivo aplicado en un sistema inteligente que sea capaz de reconocer las emociones humanas a partir de las variables fisiológicas y de las expresiones faciales. Se utilizarán los datos y resultados reportados por diversos autores y obtenidos por ellos mediante encuestas reconocidas para demostrar que el modelo desarrollado refleja la realidad observada.

Para la aplicación del algoritmo se seleccionó el campo de la medicina, sin ser limitativo ni exclusivo para este campo de la actividad humana, y para su validación se aplicó la técnica de auto consistencia para demostrar que el modelo desarrollado es consistente con las mediciones de campo y los resultados de los experimentos analizados de la literatura estudiada.

A tales efectos se diseñó el siguiente esquema de trabajo.

1. Analisis del comportamiento humano

- a.Comportamiento
- b.Espectro de voz

2.Revisión de emociones

- a.Voz y estados emocionales (Variable independiente)
- b.Relación expresión facial y estrés (Variable independiente)
- c.Relación Voz y Estrés (Variable dependiente)

3.Analisis Revision de los fundamentos matematicos

- a.Modelado matemático
- b.Algoritmos conocidos



- 4.Revisión del concepto de redes neuronales
- 5.Diseño del algoritmo
- 6.Pruebas computacionales
 - a.Estimación de emociones
 - b.Comparación con bases de datos de imágenes faciales
- 7.Pruebas de campo

8.Apoyo de expertos en el campo de aplicación seleccionado, para así conjuntamente con la adquisición y análisis comparativo de datos de personas reales para la validación de los datos y resultados obtenidos de la simulación al aplicar el algoritmo obtenido,

9.El trabajo está centrado en diseñar un sistema inteligente en un software de procesamiento matemático que sea capaz de reconocer las emociones humanas a partir de las variables fisiológicas y de las expresiones faciales y se tuvo el apoyo de los expertos para la validación de los datos y resultados, así como la adquisición de datos reales en pacientes.

La aplicación de este programa de trabajo califica la tesis como una investigación cuasi experimental, ya que combina datos y resultados de otros investigadores y se combinan con datos empíricos obtenidos de encuestas, observación y aplicación directa del algoritmo.

En lo que resta, el trabajo está distribuido de la siguiente manera: en el capítulo 2 se muestra la fundamentación matemática de las redes neuronales empleadas, además del tratamiento matemático de Fourier para el proceso de valoración de la frecuencia de voz, así como las ecuaciones que relacionan el estrés con la enfermedad. En el capítulo 3 se muestran los aspectos experimentales, los códigos que sustentan el desarrollo realizado. Los resultados se observan en el capítulo 4 y finalmente se exponen las conclusiones y recomendaciones.



Capítulo 2: La ingeniería en el campo de las emociones.

En este capítulo se exponen los fundamentos de ingeniería, metodológicos y teóricos para el desarrollo de un algoritmo de estimación de emociones, utilizando una Máquina de Soporte Vectorial con un algoritmo de gradiente orientado y una red Adams para optimización. Se reconoce el aporte significativo de la ingeniería en las áreas médicas.

Fundamentación Metodológica.

Al hacer la fundamentación epistemológica y metodológica, se evidencia que el libro se ha inclinado más por los postulados de la racionalidad técnica más que por los postulados de la racionalidad interpretativa y los de la racionalidad crítica en el campo epistemológico, así como que se ha identificado más con la tradición cuantitativa en el aspecto metodológico, ya que se trata de conocer el valor y no las intenciones y significados de las emociones de los participantes en la investigación, elementos que son cuantificables con el uso de software matemáticos.

Enfoque Epistemológico.

Resulta necesaria una fundamentación epistemológica del presente trabajo ya que la Epistemología como ciencia estudia cuál es la entidad del conocimiento científico que se va a tratar [52], y da cuenta del cómo, cuál ha sido el proceso de constitución y desarrollo de los conocimientos científicos. Además en la Epistemología no sólo deben establecer los criterios formales sancionadores de la validez del conocimiento científico sino que ha de buscar los criterios de demarcación y los procesos a seguir para el mismo.

Algunos autores [53] puntualizan, que el carácter de una investigación resulta acorde, no sólo con los métodos que utiliza, sino también con los planteamientos epistemológicos y la definición del objeto de la que parte.



En este libro se estudiaron los parámetros que influyen en el estrés para establecer algoritmos computacionales inteligentes. Esto requirió de bases matemáticas, técnicas de computación inteligente y lenguaje de programación, además fue necesario un tratamiento estadístico de datos en las fases del proceso de medición de las variables. De tal forma, que en la investigación desarrollada, se utilizó un enfoque predominantemente cuantitativo, para estimar los parámetros que influyen en el estrés. De acuerdo con [54], el enfoque cuantitativo utiliza la recolección de datos y el análisis de los mismos para desarrollar los procesos investigativos, confiando además en la medición de variables.

La investigación de este libro se realizó bajo el paradigma de la modernidad, enfocado en el positivismo según [55]. Este se establece como una relación precisa entre observación y teoría. La investigación desarrollada permite definir patrones cuya abstracción de voz, expresión facial, lleva a establecer conclusiones por comparación y semejanza, por lo que se identifican categorías que se interconectan. La postura epistémica adoptada en la presente Tesis Doctoral es el positivismo.

En la investigación del material que compone este libro, se utilizó un razonamiento deductivo, que incluyó la aplicación de técnicas de tipo matemático. Es importante mencionar que el método deductivo está asociado frecuentemente con la investigación cuantitativa [56]. También para [57], en términos del pragmatismo, la ciencia se construye por razonamientos deductivos.

La postura epistémica adoptada en la investigación y en el desarrollo experimental es el pensamiento metateórico. Al tratar de ir más allá del conocimiento del estrés, no bastan las explicaciones del pasado, es necesario predecir el futuro incluyendo otras variables al estudio que no habían sido consideradas. Se asume que los supuestos filosóficos son un medio para ordenar los paradigmas destacando sus cualidades de manera más clara y precisa.

Supuesto ontológico: El estudio del estrés debe ser abordado en una nueva manera fundamentada en el desarrollo de nuevos algoritmos, basados en técnicas de computación inteligente. Existen estudios previos, pero parten del supuesto de cómo afectan la voz, la sudoración, la fisionomía facial, entre otras, desde la perspectiva que no todos los individuos reaccionan o se afectan de una misma manera, ante estímulos generadores de estrés.

En resumen, el enfoque epistemológico aplicado en la investigación, consistió en un método cuantitativo, bajo una lógica deductiva y asumiendo el positivismo como paradigma, para dar una solución práctica a la problemática planteada por el vacío en el conocimiento, representado por la no



existencia de un método para medir y explicar los parámetros que determinan el estrés, para ello se utilizaron de manera conveniente diversas técnicas de computación inteligente para establecer los algoritmos y métodos de análisis de datos, obteniéndose los algoritmos probados y verificados en el área de la medicina.

Diseño de Experimento.

La investigación es de tipo aplicada tecnológica ya que sirve para generar conocimientos que se puedan poner en práctica en el sector productivo, con el fin de impulsar un impacto positivo en la vida cotidiana, además es de tipo cuantitativa, porque ahonda en los fenómenos a través de la recopilación de datos y se vale del uso de herramientas matemáticas, estadísticas e informáticas para medirlos. Esto permite hacer conclusiones generalizadas que pueden ser proyectadas en el tiempo. De acuerdo al tipo de variable manipulada en la investigación es cuasi experimental, porque controla solo algunas variables del fenómeno a estudiar, por ello no llega a ser totalmente experimental. Finalmente en su carácter temporal la investigación es de tipo transversal ya que se aplica para observar los cambios ocurridos en los fenómenos, individuos o grupos durante un momento concreto.

Diseño de la Investigación.

Esta investigación se concentra en el diseño cuasi experimental, que corresponde a aquella en la que existe una exposición, una respuesta y una hipótesis para contrastar, pero no hay aleatorización de los sujetos a los grupos de tratamiento y control, o bien no existe grupo control propiamente dicho [58].

¿Qué aporta este trabajo al conocimiento y a la investigación sobre el tema?

Con esta tesis se pretende contribuir al conocimiento científico en Ciencias e ingeniería basado principalmente en los siguientes aspectos:

1. Aplicar técnicas de computación inteligente en ingeniería, estableciendo algoritmos implementados en el campo de aplicaciones médicas.
2. Establecer las bases matemáticas necesarias para llevar a cabo procedimientos para expresar las variables de voz y rostro algorítmicamente, de tal forma que permitan su simulación.
3. Mostrar cómo la inteligencia artificial contribuye al tratamiento de se-



ñales médicas no lineales que caracterizan las emociones.

4. Vincular las variables de rostro y voz con el fin de establecer nuevos algoritmos para la estimación de emociones.

5. Formular el algoritmo estable y convergente que caracteriza las emociones y su efecto en el estrés de las personas.

6. Generar una línea de investigación en ciencias e ingeniería para profundizar en el estudio de distintas variables, además de voz y rostro, en distintos contextos: Industrias básicas, aeronáutica, y en diferentes campos laborales y profesionales.

¿Qué tipo de conocimiento y de qué naturaleza se ha generado con esta investigación?

Mediante este estudio queremos generar un conocimiento de tipo idiográfico sobre lo que acontece en las nuevas tendencias de la ingeniería y su aplicación en las distintas áreas profesionales. Se pretende generar un conocimiento de tipo cuantitativo principalmente, a partir de la aplicación de técnicas de inteligencia artificial, reflexionando sobre la aplicación de estas en las distintas ramas profesionales, más específicamente para este trabajo, en las ramas de la medicina y la salud. Se trata de describir el uso de la ingeniería en la medicina, resaltando las nuevas técnicas de la computación inteligente y sus aplicaciones no invasivas para el tratamiento de datos no lineales, teniendo en cuenta las perspectivas técnicas y desarrollos de otras investigaciones relacionadas. Es un conocimiento sobre lo general, que puede ser particularizado en otras áreas profesionales, pudiendo ser aplicable de distintas formas en una amplia diversidad de soluciones técnicas, por tanto está contemplado un algoritmo que se orienta a la implementación de una línea de investigación interpretativa y crítica, para el desarrollo de nuevas propuestas en ingeniería.

Fundamentación Matemática

Las ecuaciones matemáticas del estrés describen los efectos que producen los estímulos estresores en el organismo, cuando se presentan de forma consecutiva y no son atendidos oportunamente. Esta fundamentación matemática resulta indispensable para comprender la naturaleza del estrés en el organismo.

En medicina, el estrés o síndrome general de adaptación, es visto como una preocupación por parte de la Organización Mundial de la Salud (OMS) [59], quienes definen el estrés asociado a las exigencias laborales, sociales o



familiares, lo que hace que este pueda variar de un individuo a otro, ya que se ve afectado por la cultura, el nivel académico, las habilidades individuales, el entorno y las concepciones primarias de la persona desde su formación. La detección del estrés ocupó a gran parte de la comunidad científica, de tal forma que se lograron avances significativos para su inclusión y reconocimiento como enfermedad ocupacional en 1950 [60].

En la última década la ingeniería ha tenido una importante participación en el aporte de soluciones médicas, más específicamente en aquellas relacionadas a los problemas de estrés, que se han venido desencadenando de forma progresiva en todos los sectores sociales [28], [38], [61]. Diversas investigaciones de ingeniería han contribuido con desarrollos para mitigar o identificar el estrés en las personas, pero no resulta de fácil reconocimiento por la diversidad de variables asociadas a los estímulos emocionales.

Según [59], [60], [62] el estrés laboral es el único riesgo ocupacional capaz de afectar al 100% de la población. Por lo tanto, es catalogado como una epidemia. El estrés es una enfermedad que Hans Selye predijo en 1930 en el contexto médico [63]. Las personas experimentan situaciones de ansiedad, astenia y desánimo general, llamado en un principio el síndrome de estar enfermo. Estas observaciones estuvieron presentes en pacientes de distintas enfermedades y de distintas características físicas y psicológicas.

Si bien el estrés se escapa de las manos de las ciencias médicas, ya que depende de muchos factores particulares de las personas, la ingeniería asume el reto de ofrecer cada vez soluciones más eficientes. El uso de la computación inteligente en los procesos médicos es un camino incipiente para el largo trayecto que le espera a la fusión de ingeniería y medicina para los nuevos tiempos. Equipos electromecánicos y una amplia variedad de algoritmos se utilizan actualmente para dar soporte a las necesidades de la medicina.

Si empleamos la ingeniería para considerar al estrés como una variable que afecta a la salud, entonces es posible que esta relación se materialice en ecuaciones que definan su pertinencia o no, incluyendo así los estímulos estresores y las características de salud que podrían verse afectadas ante una permanencia del estrés en el organismo. Por lo tanto, se plantea desarrollar un algoritmo inteligente para captar emociones de carácter electrosensorial.

La solución del modelo dinámico del estrés permite, por una parte, obtener un conjunto de variables físicas: estresor, mecanismos psicológicos, etc. Y por otra, información sobre el impacto del estrés en las enfermedades. Rojas [66] sugiere un modelo causal para el análisis dinámico del estrés, evaluando a través del mismo la retroalimentación del estrés y las enfermedades. Rahe [67] afirma que es importante el conocimiento de la historia natural de los estresores psicosociales y su relación con los estados de salud y enferme-



dad de los individuos, considerando el impacto psicosocial de los estresores en el tratamiento del síndrome general de adaptación. En [67] se desarrolla una tabla de vida, que documenta cronológicamente los principales acontecimientos de la historia de una persona y el estado de salud concomitante a lo largo de su existencia, pudiendo comprender cómo se desarrollaron las enfermedades en consonancia con las situaciones estresantes suscitadas.

En la referencia [64] se analiza principalmente la respuesta activa a la percepción inicial de amenaza, describiendo un modelo del proceso de afrontamiento, que incluye la toma de decisiones y la cantidad de estrés que se espera en diferentes situaciones, además se consideran los efectos de la propia respuesta al estrés sobre las facultades cognitivas.

Algunos autores [68] han evaluado el concepto de síntomas psicósomáticos, como un efecto que relaciona de forma muy estrecha el estrés y un conjunto de factores tanto cognitivos como emocionales y sociales [69], [70]. Esta condición de salud se ha hecho frecuente en las últimas décadas, sobre todo en los ambientes laborales, trayendo importantes consecuencias para las personas. Los efectos que puede desencadenar el estrés incluyen cambios de ánimos fluctuantes, irritación, depresión, problemas para socializar, desánimo, bajo rendimiento en el trabajo y problemas de salud como presión alta, erupciones en la piel, problemas cutáneos, entre otros [71].

El estrés y las emociones son variables asociadas inevitablemente [72], ya que el primero se produce dependiendo de las reacciones que las emociones desencadenan en el organismo. Así, es posible que un mismo estímulo no produzca la misma emoción en un conjunto de personas, pudiendo producir reacciones diferentes en otras personas, las cuales a su vez están condicionadas a las características propias de cada una, desde la formación familiar hasta la formación académica y las relaciones sociales, son factores que influyen para que una emoción tenga ciertas reacciones fisiológicas en determinados individuos y no las tenga en otros. De esta manera, no es posible categorizar los estímulos, aunque si es posible clasificarlos probabilísticamente.

De esta manera, es pertinente decir que un alto porcentaje de personas siente estrés cuando pierde el empleo, pero esto deberá estar sujeto a las características individuales de cada uno, ya que aquella persona que tiene diversas fuentes de ingreso no sentirá el mismo estrés que aquellas personas que no tienen ninguna otra fuente de ingresos. Lo mismo ocurrirá con todos los estímulos estresores, estarán condicionados a las características propias de cada individuo.

Es posible entonces asegurar un comportamiento matemático del estrés, como variable electrosensorial tiene un proceso de avance desde su origen hasta el colapso del organismo. Partiendo desde la aparición de los estímu-



los estresores, su presencia y constancia desencadenan reacciones internas en el organismo, estas a su vez suelen tratarse con otro tipo de reacciones de auto defensa, como la ingesta de alcohol y de drogas, que aunque no resuelven el problema, lo disfrazan. Es así como el organismo se auto engaña, mientras el estrés se hace cada vez mayor, logrando afectar algunos órganos y funciones naturales, principalmente las relacionadas con el sistema circulatorio, problemas cardíacos, problemas dérmicos, y situaciones emocionales en general.

Los principales aportes de la ingeniería en la medicina, están asociados a procesos no invasivos, que son aquellos que no involucran instrumentos para romper la piel o que penetran físicamente en el cuerpo. Los ejemplos abarcan: las radiografías, un examen oftalmológico estándar, una tomografía computarizada, una resonancia magnética, un ECG y un monitor Holter. Los dispositivos no invasivos abarcan: audífonos, tablillas externas y férulas [77]. Todos ellos producto de la ingeniería.

Se han considerado en este trabajo algunos parámetros que permiten medir el estrés [78], que son variables no invasivas. Sin embargo otros estudios [79] han demostrado que son muchos los parámetros que pudieran caracterizar el estrés, pero que uno de los más relevantes es el espectro de voz [79], [80]. Masip, Garrido y Herrero [81] han valorado la importancia de las vibraciones de la voz ante las reacciones estresantes del organismo como la mentira, asegurando que todo proceso de estrés conduce a pequeñas vibraciones en el timbre de voz que permiten diferenciar a una persona en esta situación. Otros autores [82] señalan la importancia de evaluar el habla para caracterizar el estado emocional de las personas, pudiendo ser posible con métodos computacionales como front-end y back-end, tomando en cuenta la voz como variable de entrada. El sistema reconoce las emociones y sus efectos en el habla natural de las personas.

En este libro se ha considerado la evaluación del espectro de voz y de la fisonomía facial para el reconocimiento de estrés en las personas, a partir de una aplicación médica que permite la evaluación y el seguimiento de los pacientes. Además se muestran las ecuaciones dinámicas del estrés y su relación con las enfermedades.

Análisis de comportamiento

Cuando se presentan estímulos estresores (electro sensación medible), el organismo percibe estos como situaciones que afectan el estado de ánimo [83]. Cuando se presentan estos estímulos de forma permanente se hace más complicado para el sujeto evadir los efectos de los mismos, por esto



busca refugio en otras soluciones que pueden conducir al consumo de alcohol, entre otras variantes que podrían empeorar su situación de salud y su estado emocional. Cuando la influencia de los estímulos estresores es muy fuerte, se empiezan a manifestar otras reacciones en el organismo, ocasionando problemas más serios de salud, problemas de conducta y problemas emocionales que podrían llegar a ser irreparables si no se atienden de forma apropiada.

Modelado matemático

Según [66]- [3] las ecuaciones dinámicas para establecer el algoritmo son productos de una relación causa-efecto, basado en los flujos de entrada y salida. En este libro se ha considerado un diagrama de estados, como se observa en la figura 1. Donde los estados varían desde el estrés hasta la eliminación del estresor, que produce el afrontamiento a la enfermedad, como consecuencia de una solución médica profesional. La dinámica propuesta por Forrester [83]- [3] supone la vinculación de variables de un determinado sistema con el fin de alcanzar las dependencias entre ellas y sus efectos.

Para formular el algoritmo se ha considerado un diagrama de estados que consiste en una representación del estrés, E, estresor, Es, y también aquellas variables asociadas a los estados de salud de las personas. Se ha tomado en cuenta que cuando el estresor aumenta, el nivel de estrés también aumenta, y en consecuencia la persona afectada tiende a buscar soluciones para intentar equilibrar esta situación, estas soluciones no son tales, pues en su mayoría suelen ser el alcohol, la ingesta excesiva de alimentos, las salidas a fiestas y demás actividades que suelen utilizarse para disfrazar la situación de estrés, que además se vinculan con un mecanismo de autodefensa del organismo y resistencia para abordar la situación de estrés. Cuando estas posibles soluciones actúan es posible observar una posible mejora del estado de salud, que a su vez, irá en decremento a medida que el estresor siga presente y que las soluciones aparentes no puedan producir el efecto real que se espera, para finalmente ser indispensable la ayuda de un profesional médico. En la figura 1 se muestran las relaciones entre el estresor y el desencadenamiento de un mal estado de salud en las personas, siendo necesaria la reducción del estresor para mejorar las condiciones del estrés y en consecuencia, el mejoramiento de la salud del paciente.



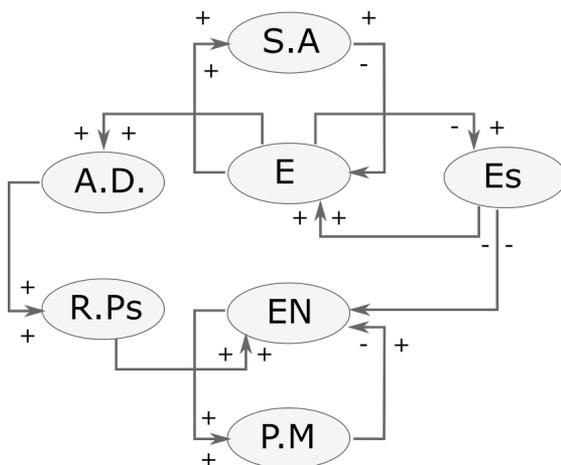


Figura 1. Diagrama de estados; SA: solución aparente, E: estrés, Es: Es-tresor; AD: AutoDefensa; RPs: Respuesta Psicosomática; EN: Enfermedad; PM: Profesional Médico

Analizando los estados del diagrama de la figura 1, es posible coincidir con [66], [3] y constatar las ecuaciones de la (1) a la (6):

$$\frac{d(En)}{dt} = R.Ps - P.M \quad (1)$$

$$E = \alpha(E * EN) \quad (2)$$

$$S.A = \beta(E) \quad (3)$$

$$RPs = \delta(f(E)) \quad (4)$$

$$P.M = \gamma(EN) \quad (5)$$



$$(6)$$

Las constantes α , β , δ y γ , refieren a parámetros propios de (1) y (2). Cuando analizamos [3] y [84], es posible cuantificar el estrés y el estado de salud, a partir de la tasa de incidencia del mal estado de salud (α), la frecuencia de manifestación del estrés (β), índice del mal estado de salud en contraste con la escala de estrés [84], y la frecuencia de la enfermedad (γ).

De las referencias [85], [86] es posible afirmar que a medida que el nivel de estrés se hace mayor, el organismo intenta solucionar la situación, con mecanismos de autodefensa. Así pues, es posible generalizar las ecuaciones en función de x y y que representan el estrés (x) y la enfermedad o mal estado de salud (y).

$$\frac{dx}{dt} = \alpha xy - \beta x \tag{7}$$

$$(8)$$

$$\alpha xy - \beta x = 0$$

Para atender estas ecuaciones es necesario hallar el jacobiano:

$$(9)$$

$$\begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} = \begin{pmatrix} \alpha y - \beta & \alpha x \\ -\alpha x & -d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

representación linealizada de (7) y (8) es

$$(10)$$

Las ecuaciones (7) y (8) representan un sistema no lineal autónomo, en el que es posible evaluar los puntos estacionarios o críticos, y se evidencia que estos estarían dados por (11)

$$\text{Pto2: } \left(\frac{bd}{ac}; \frac{b}{a} \right) \tag{11}$$



La intersección de las nuclinas dadas en (7) y (8) definen estos puntos críticos dados por (13). Es posible entonces que este sistema no lineal pueda ser linealizado con un desarrollo de Taylor en el punto crítico, en este punto las variables que definen el comportamiento del sistema no manifiestan cambios o variaciones y por tanto, las derivadas asociadas a ellas se anulan. Por lo que las ecuaciones en el punto de funcionamiento se expresan como:

$$ax_0 = b \quad ; \quad x_0 = \frac{b}{a} \quad \text{con } x' = 0 \tag{12}$$

$$y' + dy - cx = 0 \tag{13}$$

$$y_0 = \frac{cx}{d} \quad ; \quad y_0 = \frac{cb}{ad} \quad \text{con } y' = 0 \tag{14}$$

$$\tag{15}$$

Ya definidas las ecuaciones en el punto de operación, es posible desarrollar Taylor, y linealizar con el jacobiano (16) en los puntos críticos.

En (18) se observa la matriz jacobiana con los autovalores en el origen dada por:

$$J(\frac{b}{a}, \frac{cb}{ad}) = \begin{pmatrix} -\frac{b}{a} & 0 \\ c & -d \end{pmatrix} \quad \lambda = \begin{pmatrix} \frac{bd}{ac} & b \\ a & -d \end{pmatrix} \tag{16}$$

$$= -\frac{d}{2} \pm \frac{1}{2} \sqrt{d^2 + \frac{b^2 d}{a}} \quad \lambda_1 = -\frac{d}{2} - \frac{1}{2} \sqrt{d^2 + \frac{b^2 d}{a}}, \lambda_2 = -\frac{d}{2} + \frac{1}{2} \sqrt{d^2 + \frac{b^2 d}{a}}$$

Mientras que los autovalores para el punto crítico vienen dados por obteniendo así un

Puede observarse que el discriminante es siempre positivo, como también aseguran otros autores [3], y también es evidente que λ_{-1} es negativo y λ_{+2} es positivo. Lo que ocasiona un equilibrio circunstancial, esto implica que si el estresor aumenta, el nivel de estrés aumenta produciendo estados



de ánimos crecientes en la persona afectada, mientras que si el estresor se reduce, el nivel de estrés se hace menor y el estado de salud del paciente mejora.

El sistema linealizado en L'place se muestra en (17) y (18); por convencionalismo se han cambiado las letras griegas por letras latinas ($a \rightarrow \alpha$; $b \rightarrow \beta$; $c \rightarrow \delta$; $d \rightarrow \gamma$)

$$X(S) = \frac{bd}{acs} \tag{17}$$

$$\tag{18}$$

Diseño del algoritmo

El algoritmo propuesto consta de cinco partes principales descritas en la figura 2, en esta es posible apreciar que el sistema completo consta de una etapa de detección de rostro, otra etapa de detección de voz y finalmente una etapa de interfaz de usuario.

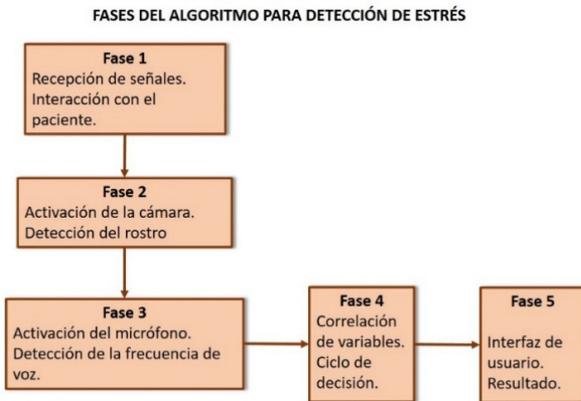


Figura 2. Fases del algoritmo para detección de estrés

El espectro de voz

En el proceso de proyección de la voz se involucran órganos del sistema respiratorio y digestivo, que son controlados por el sistema nervioso central [87]. La excitación generada en las cuerdas vocales es propagada a través de la faringe, la cavidad bucal y la cavidad nasal [87], [88]. Estas cavidades determinan las características acústicas de la voz [89].

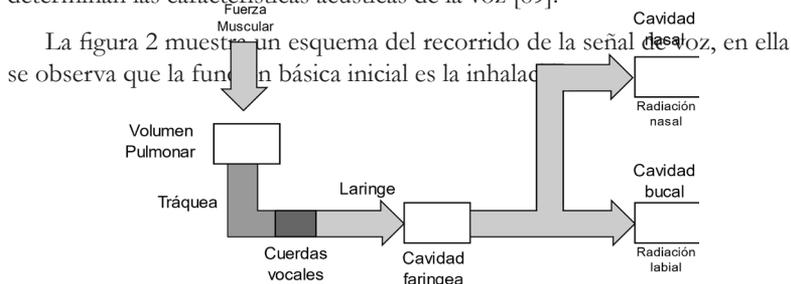


Figura 3. Diagrama de producción de voz.

La expulsión del aire depende de la energía de los músculos del tórax, cuando el tórax se contrae se produce un aumento de la presión pulmonar, lo que expulsa el aire y lo hace atravesar los bronquios y la tráquea, actuando como excitatriz del conducto vocal. Una vez realizado este proceso es posible producir una voz sonora, como consecuencia de la tensión en las cuerdas vocales, que vibran producto del flujo de aire. También es posible generar una voz sorda, producto de una obstrucción en el flujo de aire que atraviesa la cavidad vocal.

El modelo matemático de generación de voz se basa en el análisis del tracto vocal, como una concatenación de tuberías de sección variable, produciendo una función de transferencia [87].

La transformada de Fourier es una de las herramientas más utilizadas en el procesamiento de señales, la cual consiste en pasar una señal del dominio del tiempo al dominio de la frecuencia, lo cual facilita el análisis de las dife-



rentes frecuencias presentes en la señal.

$$\hat{f}(w)$$

La transformada de Fourier está dada en (20) y en ella se observa $f(t)$ en función del tiempo y $e^{-j2\pi wt}$ en el dominio de la frecuencia.

$$(19)$$

Una de las debilidades de la Transformada de Fourier es que no permite el análisis frecuencial localizado [89]. El método de la función ventana permite analizar el tiempo y frecuencia de la señal. Para este proceso se multiplica la señal original por la función ventana, que es la encargada de acotar la señal en un intervalo de tiempo. Esta transformación es llamada Transformada de Gabor y se define en (21).

$$(20)$$

$g(t)$ es la función gaussiana que define entonces la transformada Gabor.

Además, se aplica una descomposición de modo empírico, para completar el proceso con un FILTRADO de la señal. Los componentes obtenidos de dicha filtración se llaman función de modo intrínseco [90].

La voz y los estados emocionales

Duque y Morales [15] afirman que existe una relación estrecha entre las características de la voz y los estados emocionales de las personas (figura 3). Algunas investigaciones [91], [92], han demostrado que varios aspectos del estado físico y emocional, incluyendo edad, sexo, nivel de inteligencia, aspecto físico y personalidad, pueden ser identificados solo con la voz.





Figura 4. Relación voz-estados emocionales

Gracias a estudios recientes [93], ha sido posible asegurar que algunos de los componentes de la voz son característicos para expresar emociones, entre las que se puede mencionar:

La frecuencia fundamental.

EL tiempo de duración.

La calidad de la voz.

De estos aspectos, la frecuencia fundamental es la más resaltante para determinar las emociones. La curva del tono de voz podría suponerse discontinua para las emociones consideradas como negativas (miedo, enfado) y es suave para las emociones positivas (por ejemplo, la alegría).



El estrés y la voz

La Organización Mundial de la Salud (OMS) define el estrés como el conjunto de reacciones fisiológicas que prepara al organismo para la acción [94].

La clasificación del estrés tomando en cuenta el promedio de frecuencia de la voz, es considerada por [95], [96], [15] en la forma siguiente:

Estado neutral: Es el estado en el que la persona no tiene un estrés significativo, está en estado de relajación, la frecuencia de voz está en el rango $12\text{Hz} > f_v \geq 8\text{Hz}$

Estado medio: Es el estado en el que la persona siente cierto nivel de estrés significativo, angustia, inquietud moderada, la frecuencia de su voz se encuentra en el rango de $15\text{Hz} > f_v \geq 12\text{Hz}$

Estado alto: Es el estado en el que la persona presenta un alto nivel de estrés, inquietud elevada, ansiedad, factores físicos como sudoración, la frecuencia de su voz está en el rango de $17\text{Hz} \geq f_v \geq 15\text{Hz}$.

El estrés y la expresión facial

Las emociones humanas han estado en estudios desde hace muchos años [7]. Lazarus [69] propuso una lista con requisitos que debían cumplirse para que la emoción se considere válida:

Tener un sustrato neural específico y distintivo

Tener una expresión o configuración facial específica y distintiva.

Poseer sentimientos específicos distintivos.

Derivar de procesos biológicos evolutivos.

Manifiestar propiedades motivacionales y organizativas de funciones adaptativas.

El autor señala que las emociones más destacadas son la ira, la tristeza, la alegría, el asco, la sorpresa, el miedo y la emoción nula.



Capítulo 3: Algoritmos de evaluación y clasificación.

Las redes neuronales artificiales

En los procesos con redes neuronales se tienen cuatro elementos básicos:

1.Las conexiones, pesos o sinapsis que definen el comportamiento de la neurona. Dichas conexiones pueden estar representadas por un signo positivo cuando se consideren excitadoras, o presentar un signo negativo cuando se consideren inhibitoras.

2.Un elemento sumador que suma las entradas multiplicadas por las sinapsis correspondientes.

3.La función de activación no lineal para condicionar la amplitud de la señal de salida.

4.Un nivel de umbral que determina la activación de la neurona.

La figura 4 describe de forma gráfica el comportamiento y los elementos de la neurona.

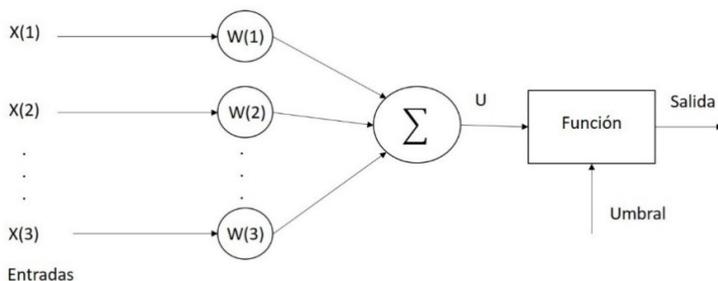


Figura 5. Modelo de la red neuronal



Matemáticamente se pueden reconocer las expresiones dadas en (21) y (22)

$$U = \sum_{j=1}^k w(j).x(j) \tag{21}$$

Y

$$\text{Salida} = \rho(U - \text{umbral}) \tag{22}$$

Donde ρ es una función no lineal conocida como función de activación. Por lo general esta se asocia al umbral de la salida U, mediante una entrada y un peso adicional, como se observa en (23).

$$\text{umbral} = \sum_{j=0} w(j).x(j), x(0) = 1 \tag{23}$$

El modelo neuronal descrito es considerado el modelo general, sin embargo también son posibles otros modelos que no llevan a cabo un promedio de las entradas directamente, sino que antes de multiplicar por los pesos realizan una transformación de las entradas, que puede ser cuadrática, polinómica o esférica.

El modelo planteado en (23) es un modelo estático, por lo que un modelo más completo debería considerar salidas anteriores, dando origen a un modelo dinámico, originando una neurona con memoria, como describe (24):

$$\text{Salida} = F(\text{salidas}_{n-k}, \text{entradas}), k=1, \dots, n-1 \tag{24}$$

Lo que conduce a afirmar que la salida no solo depende de las entradas como en (23) sino que además depende de las salidas anteriores.



Redes neuronales convolucionales

Este tipo de red está asociado al principio del perceptrón multicapa, y se enfoca en el funcionamiento de la corteza visual, específicamente de las células responsables de la selectividad de orientación y detección de bordes en los estímulos visuales dentro de la corteza primaria [83]. Estas redes funcionan con matrices bidimensionales, que se enfocan en la visión artificial y permiten una amplia aplicabilidad para sistemas de clasificación y segmentación de imágenes.

En el proceso de extracción de características las neuronas simples de un perceptrón son sustituidas por matrices que actúan sobre los datos bidimensionales que pasan por ellas, produciendo una salida en cada neurona como la ecuación (25)

$$Y_j = g(b_j + \sum_i K_{ij} \otimes Y_i) \tag{25}$$

donde la salida Y_j que corresponde a una neurona j está dada por una matriz que se estima por una combinación lineal de las salidas dadas por Y_i de las neuronas de la capa anterior, que han sido calculadas con el núcleo de convolucional K_{ij} que hace posible dicha conexión. Este resultado es sumado a un dominio b_j que son a su vez activadas por una función g no lineal.

Este operador de convolución se encarga de filtrar la imagen entrante con un núcleo que ha sido entrenado previamente. Este procedimiento permite abstraer características abstractas de las imágenes, a partir de la asignación de valores a los píxeles y creando núcleos que resaltan la gradiente en una única dirección definida.

Redes de memoria a corto plazo

Las redes de memoria a corto plazo aprenden a largo plazo a partir de datos de secuencia. Para este trabajo se utiliza la capa bidireccional, que permite evaluar la secuencia en direcciones de las señales EKG, consideradas para la medición de error del sistema. Las redes a corto plazo son elementos esenciales para las redes recurrentes, la figura 5 muestra la arquitectura básica de este tipo de red.



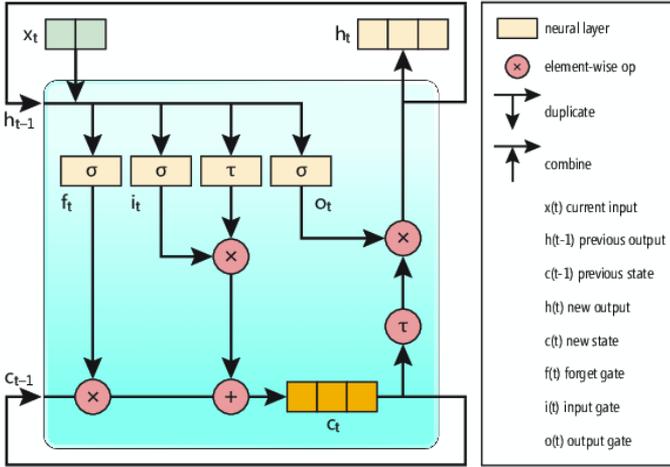


Figura 6. Arquitectura de la red de corto plazo.

Matemáticamente es posible destacar las ecuaciones (26) a la (30) que describen los elementos de entrada, elementos de salida, elementos de pesos, y funciones de activación.

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{26}$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{27}$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{28}$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tau(W_c x_t + U_c h_{t-1} + b_c) \tag{29}$$

$$h_t = o_t \circ \tau(C_t) \tag{30}$$



donde:

n : tamaño de la entrada

m : tamaño del estado de la celda y salida.

x_t : vector de entrada, tiempo t , tamaño $n \times 1$

f_t : vector de la puerta de olvido, tamaño $m \times 1$

i_t : vector puerta de entrada, tamaño $m \times 1$

o_t : vector puerta de salida, tamaño $m \times 1$

h_t : vector de salida, tamaño $m \times 1$

C_t : vector de estado de la celda, tamaño $m \times 1$

$W_{(f)}$, $W_{(i)}$, W_o, W_c : matrices de pesos de la puerta de entrada, tamaño $m \times n$

$U_{(f)}$, $U_{(i)}$, U_o, U_c : matrices de peso de la puerta de salida, tamaño $m \times m$

$b_{(f)}$, $b_{(i)}$, b_o, b_c : vectores bias, tamaño $m \times 1$

σ : función de activación sigmoide logistica

τ : función de activación tangente hiperbólica

Algoritmo K-Nearest Neighbors

El algoritmo K-Nearest-Neighbor es del tipo supervisado, y corresponde a la categoría de Machine Learning. Es utilizado para la clasificación de muestras bien sean del tipo discretas o del tipo continuo en el caso de las predicciones. Se trata de un algoritmo sencillo basado en la búsqueda por el vecino más cercano, tratando de ubicar la similitud entre los puntos adyacentes, aprendidos en el proceso de entrenamiento.

A pesar de presentar ciertas similitudes con el algoritmo K-means, que se trata de un proceso no supervisado, donde la k se refiere al número de clústers que se desean clasificar, en el caso del K-Nearest Neighbors (KNN) la k se trata del número de puntos vecinos que se deben considerar para lograr la clasificación los grupos que ya se conocen, por tratarse de un algoritmo supervisado.

Las características esenciales del algoritmo KNN son:



Supervisado: significa que ya los clúster han sido etiquetados, y solo están a la espera de datos.

Basado en instancias: el algoritmo memoriza las instancias de entrenamiento que son usadas como base de reconocimiento para el proceso de predicción.

Una de las ventajas que tiene este algoritmo es que requiere mucho consumo de memoria para el procesamiento, por lo que resulta mejor el tratamiento con bases de datos pequeñas.

El proceso de medición de distancias para la clusterización es realizada a través de la distancia Euclidiana (31) o la Cosine Similarity, que mide el ángulo vectorial, considerando aquellos que sean menores.

$$d(x_i, y_j) = \sqrt{\sum_{r=1}^p (x_{ri} - x_{rj})^2} \quad (31)$$

Cada vector está descrito en términos de p atributos considerando q clases para la clasificación. Los valores de los atributos del i -ésimo vector (donde $1 \leq i \leq n$) se representan por el vector p -dimensional.

La fase de entrenamiento del algoritmo consiste en almacenar los vectores característicos y las etiquetas de las clases de los ejemplos de entrenamiento. En la fase de clasificación, la evaluación del vector (del que no se conoce su clase) es representada por un vector en el espacio característico. Se calcula la distancia entre los vectores almacenados y el nuevo vector, y se seleccionan los k vectores más cercanos. El nuevo ejemplo es clasificado con la clase que más se repite en los vectores seleccionados.

Algoritmo SVM

Una máquina de soporte vectorial (SVM) es un algoritmo de aprendizaje supervisado que se puede emplear para clasificación binaria o regresión. Las máquinas de soporte vectorial son muy populares en aplicaciones como el procesamiento del lenguaje natural, el habla, el reconocimiento de imágenes y la visión artificial.

Una máquina de soporte vectorial construye un hiperplano óptimo en



forma de superficie de decisión, de modo que el margen de separación entre las dos clases en los datos se amplía al máximo. Los vectores de soporte hacen referencia a un pequeño subconjunto de las observaciones de entrenamiento que se utilizan como soporte para la ubicación óptima de la superficie de decisión.

Las máquinas de soporte vectorial pertenecen a una clase de algoritmos de Machine Learning denominados métodos kernel y también se conocen como máquinas kernel.

El entrenamiento de una máquina de soporte vectorial consta de dos fases:

Transformar los predictores (datos de entrada) en un espacio de características altamente dimensional. En esta fase es suficiente con especificar el kernel; los datos nunca se transforman explícitamente al espacio de características. Este proceso se conoce comúnmente como el truco kernel.

Resolver un problema de optimización cuadrática que se ajuste a un hiperplano óptimo para clasificar las características transformadas en dos clases. El número de características transformadas está determinado por el número de vectores de soporte.

Para construir la superficie de decisión solo se requieren los vectores de soporte seleccionados de los datos de entrenamiento. Una vez entrenados, el resto de los datos de entrenamiento son irrelevantes.

El entrenamiento de una máquina de vectores de soporte consta de dos fases:

Transformar los predictores (datos de entrada) en un espacio de características altamente dimensional. En esta fase es suficiente con especificar el kernel; los datos nunca se transforman explícitamente al espacio de características. Este proceso se conoce comúnmente como el truco kernel.

Resolver un problema de optimización cuadrática que se ajuste a un hiperplano óptimo para clasificar las características transformadas en dos clases. El número de características transformadas está determinado por el número de vectores de soporte.

Para construir la superficie de decisión solo se requieren los vectores de soporte seleccionados de los datos de entrenamiento. Una vez entrenados, el resto de los datos de entrenamiento son irrelevantes.



Algoritmo HOG

HOG significa Histogramas de Gradientes Orientados. HOG es un tipo de descriptor de características. El objetivo de un descriptor de características es generalizar el objeto de tal forma que el mismo objeto (en este caso un rostro) produzca lo más cerca posible del mismo descriptor de características cuando se vea bajo diferentes condiciones. Esto hace que la tarea de clasificación sea más fácil.

El algoritmo HOG utiliza una función global para describir un rostro en lugar de una colección de características locales. En pocas palabras, esto significa que todo rostro está representado por un único vector de características, a diferencia de muchos vectores de características que representan partes más pequeñas de ese rostro.

El detector de rostros HOG usa una ventana de detección deslizante que se mueve alrededor de la imagen. En cada posición de la ventana del detector, se calcula un descriptor HOG para la ventana de detección. Este descriptor se muestra luego al SVM entrenado, que lo clasifica como “rostro” o “no rostro”.

El algoritmo HOG Es un descriptor de características, usado para detectar objetos en visión por computadora y procesamiento de imágenes. Cuenta las ocurrencias del gradiente de orientación en porciones localizadas de una imagen - ventana de detección, o región de interés (ROI).

Implementación del algoritmo del descriptor HOG:

1. Dividir la imagen en pequeñas regiones conectadas llamadas celdas, y por cada celda calcular un histograma de direcciones de gradiente u orientaciones de aristas para los píxeles dentro de la celda.

2. Discretizar cada celda en los contenedores angulares de acuerdo con la orientación del gradiente.

3. Cada píxel de la celda contribuye al gradiente ponderado de su correspondiente contenedor angular.

4. Los grupos de celdas adyacentes son considerados como regiones espaciales llamadas bloques. Los grupos de celdas en un bloque es la base para la agrupación y normalización de histogramas.

5. Los grupos normalizados de histogramas representan el histograma



del bloque. El conjunto de estos histogramas de bloques representan el descriptor.

Algoritmo LBP

Los patrones binarios locales (LBP) son un tipo de descriptor visual utilizado para la clasificación en visión artificial. Además puede decirse que es una herramienta poderosa para la clasificación de textura, se ha determinado que cuando LBP se combina con el descriptor Histograma de gradientes orientados (HOG), mejora considerablemente el rendimiento de detección en algunos conjuntos de datos.

El vector de características puede procesarse utilizando la máquina de soporte vectorial, máquinas de aprendizaje extremo o algún otro algoritmo de aprendizaje automático para clasificar imágenes. Dichos clasificadores se pueden usar para el reconocimiento facial o el análisis de textura.

Las expresiones matemáticas, conceptos y descripciones planteadas en este capítulo son la base fundamental para desarrollar el algoritmo mostrado en el capítulo siguiente.

LBP proporciona un operador de análisis de textura que se define como una medida de la textura en una escala de grises invariante, derivado de una definición general de textura mediante vecinos locales.

La forma actual del operador LBP es muy diferente de su versión básica: la definición original se extiende a un conjunto de vecinos arbitrarios circulares, y se han desarrollado nuevas versiones del mismo, sin embargo, la idea principal es la misma: un código binario que describe el patrón de la textura local que es construido por el umbral de un conjunto de vecinos por el valor de gris de su centro. El operador tiene que ver con muchos otros métodos conocidos de análisis de texturas, en la figura 7 se observan las relaciones del operador LBP con los otros métodos.



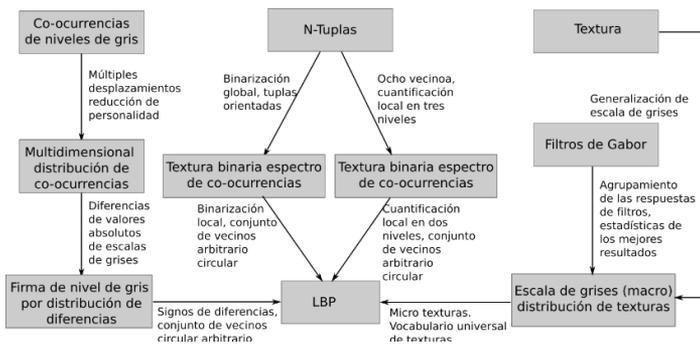


Figura 7. LBP en el campo de análisis de imágenes

Otra versión para el operador original, es la llamada patrones uniformes, que puede ser usado para reducir la longitud del vector de características y para implementar un descriptor sencillo e invariante frente a rotaciones. Esta versión fue inspirada por que algunos patrones binarios son más frecuentes en las imágenes de textura que otros. Un código LBP se llama homogéneo si el patrón binario contiene un máximo de dos transiciones a nivel de bits, de 0 a 1 o viceversa, cuando el patrón de bits es atravesado de manera circular. Por ejemplo, los patrones de 00000000 (0 transiciones), 01110000 (2 transiciones) y 11001111 (2 transiciones) son uniformes, mientras que los patrones de 11001001 (4 transiciones) y 01010010 (6 transiciones) no lo son. En el cálculo de las etiquetas LBP, cuando los patrones uniformes son utilizados, se utiliza una etiqueta para cada uno de los patrones uniforme y todos los patrones no uniformes están etiquetados con una sola etiqueta. Por ejemplo, cuando se utiliza (8, R) de vecindad, hay un total de 256 patrones, 58 de los cuales son uniformes y el resto son no uniformes, por lo que resulta un total de 59 etiquetas diferentes [97].

Fases para el desarrollo del algoritmo

FASE I: Caracterización de las variables asociadas a las emociones humanas, rostro y voz.

FASE II: Caracterización de las variables matemáticas asociadas a la valoración del estrés y las emociones y su impacto en la generación de



enfermedades.

FASE III: Diseño del algoritmo de estimación de emociones con la fisionomía facial.

FASE IV: Diseño del algoritmo de estimación de emoción con el espectro de voz.

FASE V: Integración de algoritmos de detección de emociones.

FASE VI: Optimización del algoritmo a través de simulaciones y ajustes de variables.

FASE VII: Validación del algoritmo.

Fase VIII: Documentación y presentación de resultados. En esta fase se procedió al análisis de los resultados a fin de establecer las conclusiones del tema.

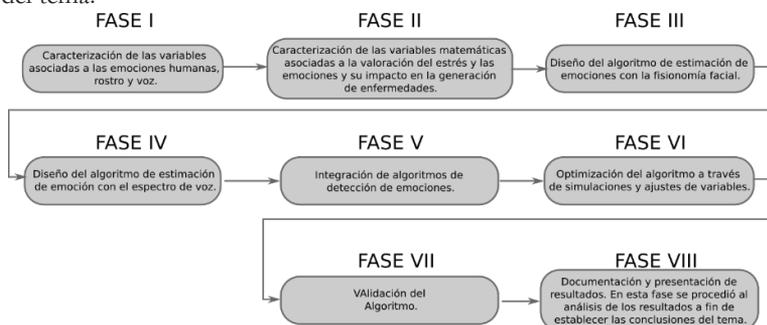


Figura 8. Secuencia de las fases del trabajo



Capítulo 4: Pruebas computacionales.

El estrés se puede detectar a partir de numerosas variables fisiológicas como secreción ácida estomacal, salivación, actividad motora intestinal, irregularidades menstruales, humedad de la piel, alteración del trabajo cardíaco, alteración de las expresiones faciales, alteraciones en la frecuencia de voz, alteraciones en la respuesta galvánica de la piel, entre otras. Sin embargo los estudios previos han demostrado que la fisonomía facial [5], las alteraciones en la voz [15] , [81] son determinantes y exclusivas para la valoración de estrés en las personas.

En este trabajo se ha utilizado un software de reconocimiento de expresiones faciales y se ha modificado para realizar una captura de expresiones en movimiento, con la finalidad de optimizar el proceso y lograr una captura real de las expresiones. Además se realizó una adaptación de un software para detección y clasificación de las emociones a partir de la frecuencia de voz, este software fue desarrollado tomando en cuenta la fundamentación bibliográfica asociada a las frecuencias de voz [95] y las emociones.

Tomando en cuenta la fundamentación bibliográfica se realizó el procesamiento de voz y rostro para la detección de emociones [71] [69]. La figura 7 muestra la implementación realizada en Matlab® para el algoritmo.

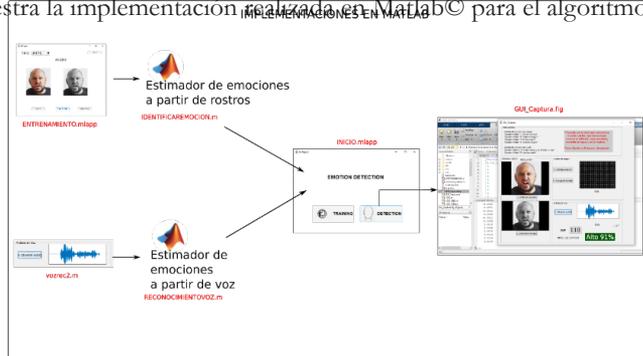


Figura 9. Implementación de variables en Matlab



Estos dos procesos se unen en el algoritmo planteado en esta tesis, que estaría entonces compuesto de dos elementos básicos para la detección de emociones. Este algoritmo reúne dos variables no invasivas que en conjunto componen un sistema sencillo para la detección, además se ha incorporado un código que combina los dos resultados para ofrecer una respuesta más acertiva (figura 8).

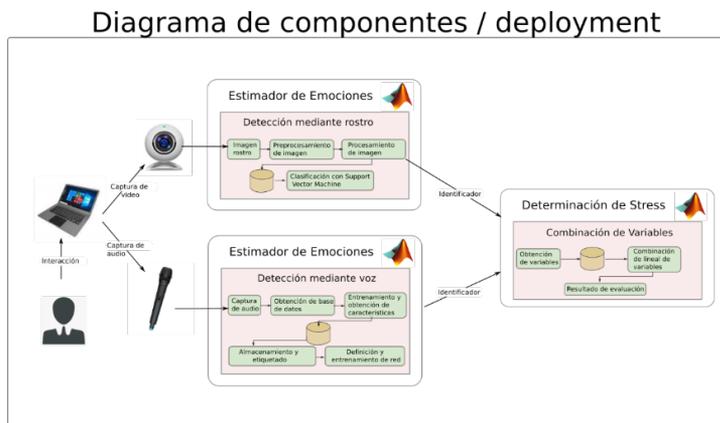


Figura 10. Diagrama de componentes

A continuación, se presentan los algoritmos utilizados para el sistema de detección de emociones. El sistema consta de un proceso de dos fases, que determinan las variables más relevantes para la detección emocional [28], [13], [15], [16], [43]. Estas son compiladas en un sistema alterno que caracteriza la combinación de dichas variables para definir el porcentaje de estrés en las personas (figura 6).

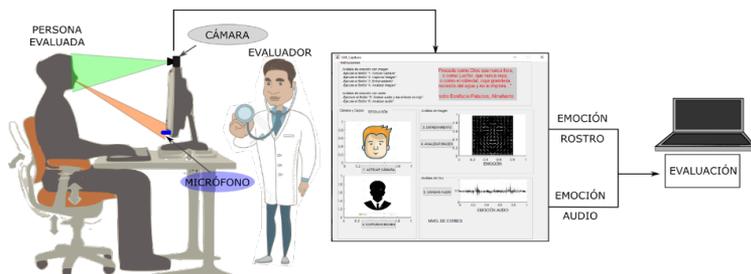


Figura 11. Diagrama general del sistema de detección de emociones y clasificación



ficación del estrés.

En la figura 7 se observa la descripción gráfica del sistema desarrollado, en ella se puede observar la forma en que se lleva a cabo la toma de datos y posterior análisis de la información. El paciente se sienta frente a la computadora que contiene la interfaz de usuario que le permite leer en voz alta el texto que hará posible la detección de la frecuencia de voz, mientras que la cámara estará dispuesta para la captura de la imagen del rostro. El médico encargado podrá hacer los ajustes necesarios de la cámara y activar el sistema de detección de estrés cuando lo considere apropiado. Él podrá observar los resultados y en base a esto tomar las consideraciones que crea conveniente para el tratamiento del paciente.

La figura 9 muestra el entorno de ejecución que hace posible la detección de emociones utilizando el algoritmo planteado.



Figura 12. Arquitectura de la interfaz de determinación de estrés

Estimación de emociones mediante análisis de imagen facial.

De los múltiples métodos que existen para el reconocimiento de emociones [1]-[4], el propuesto en [1], presenta más que un buen desempeño, exactitud, considerando las múltiples etapas de entrenamiento y evaluación para siete emociones básicas como son ira, desprecio, asco, vergüenza ,



miedo, felicidad, neutral, orgullo, tristeza y sorpresa.

En la figura 7 se observa el proceso de aprendizaje y clasificación de la fisonomía del rostro, considerando la extracción de características.

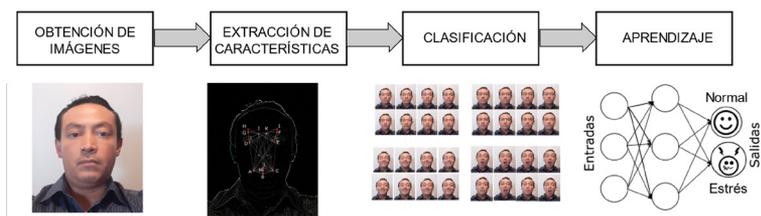


Figura 13. Esquema del proceso de extracción de características, clasificación y aprendizaje de expresiones faciales para detección de estados emocionales normales y de estrés.

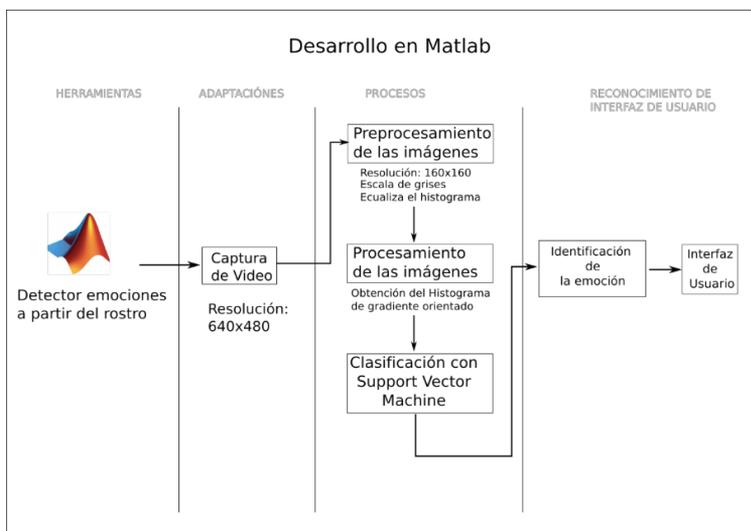


Figura 14. Desarrollo en Matlab para la detección de rostros

Selección de la base de datos de imágenes faciales

Se revisaron las bases de datos existentes y se analizaron en función de su accesibilidad y el tipo de datos, la tabla 1 muestra algunas bases de datos



analizadas.

Tabla 1. Selección de la base de datos

Nombre	Tamaño	N° emociones	Acceso	Encontrada
Humaine Database	40 imágenes	6	Disponible	No hay registros
Belfast Database	125 imágenes	6	Disponible	No hay registros
Sal Database	80 imágenes	1	Bajo pedido	Si
Indian Institute of tecnology	240 imágenes	6	Bajo pedido	Si
The Yale Face database	165 imágenes	3	Bajo pedido	No hay registros
Caltech Frontal Face	450 imágenes	No definido	Disponible	Si
Human Scan	400 imágenes	No definido	Disponible	Si
JAFEE	27 Por cada emoción	7	Disponible	Si
KDEF	27 por cada emoción	7	Disponible	Si
MUG	28 por cada emoción	7	Disponible	Si
WSEFEP	28 por cada emoción	7	Disponible	Si
TFEID	138 imágenes	7	Disponible	Si
ADFES	138 imágenes	7	Disponible	Si

Para el estudio con la imagen facial se describe a continuación la fase inicial del sistema que incluye el tratamiento de las imágenes y la clusterización de las mismas según el estado emocional. Para ello se han considerado 4 bases de datos; JAFEE, KDEF, MUG, WSEFEP.

Es importante resaltar que se utilizan los algoritmos SVM y KNN porque arrojan un valor de precisión de 71,42% con la base de datos JAFEE, sin embargo, con el HOG y SVM arroja un valor de precisión de 75% con la base KDEF. El algoritmo HOG con el KNN y con la base MUG arroja 65,71%.



Para la parte de reconocimiento facial se emplean varias etapas, la primera corresponde al preprocesamiento de las imágenes, para luego normalizarlas y eliminar los efectos de la luz (Figura 7) Luego, en una siguiente fase, se utiliza el procesamiento de extracción de características, que se emplea a través de un Histograma de Gradiente Orientado, que es el encargado de propiciar los datos de entrada para el entrenamiento y a su vez los datos de prueba para cada una de las emociones a clasificar. Más adelante en una tercera fase se configura un Support Vector Machine (SVM) como clasificador, junto con el algoritmo del vecino más cercano, k-Nearest Neighbours (KNN) que procesa cada dato nuevo evaluando los k vecinos cercanos, con el fin de determinar las distancias del nuevo elemento a cada uno de los ya existentes, de tal forma que hace posible la organización de las distancias en orden descendente y ubicar los elementos según el grupo al que correspondiera. Este algoritmo no supervisado, se basa en varios atributos descriptivos y un solo atributo objetivo conocido como clase. Se evalúa además el desempeño del sistema considerando la técnica de la matriz confusión (CM). Esta matriz considera el número de predicciones de cada clase en cada una de sus columnas, y las instancias en la clase real están dadas por las filas de la matriz, así uno de los beneficios de las matrices de confusión es que facilitan la revisión del sistema y permiten determinar si este está confundiendo dos clases. El clasificador se evalúa tomando en cuenta las muestras de clases de la entrada, si estas producen muchos cambios en la tasa de error, entonces no serán representativas en cuanto la eficiencia del clasificador. El sistema es evaluado considerando las bases de datos JAFFE, KDE, WSEFEP y MUG, arrojando una tasa de predicción aproximada del 96,13%.

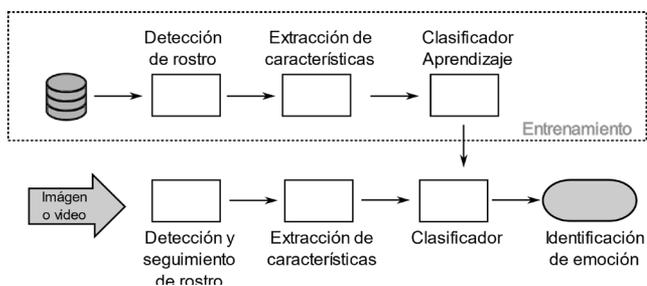


Figura 15. Esquema de funcionamiento de reconocimiento de rostro para la detección de emociones

El mejor resultado fue posible empleando el algoritmo HOG, tomando



en cuenta el histograma de gradiente orientado para la obtención de características (Figura 8). Los métodos incluyen los extractores de características HOG y LBP, y los clasificadores KNN y SVM. Pudiendo combinar los métodos y las bases de datos.

Para este tipo de procesamiento es necesario crear una base de datos con más de 20 imágenes, que puedan ser almacenadas de forma clasificada, una vez que estas ya han sido tratadas y normalizadas. Las imágenes pueden ser de mediana y baja resolución, como se observan en la figura 9, con la finalidad de no invertir mucho coste de procesamiento computacional. Para cada estado emocional se han considerado un conjunto de opciones diferentes, que da lugar a 25 imágenes aproximadas por cada expresión.

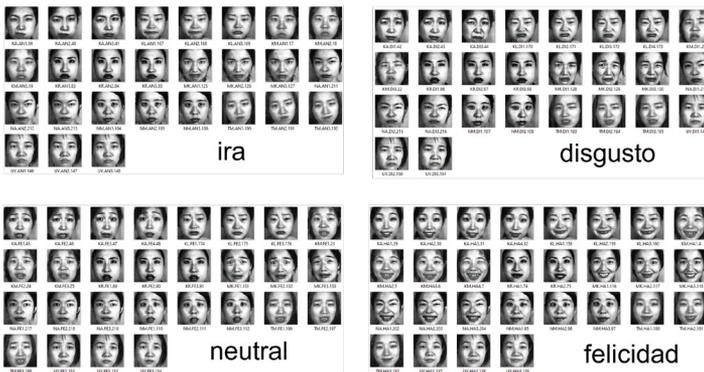


Figura 16. Ejemplos de imágenes para la base de datos

El sistema es alimentado con los datos de entrenamiento, pero además es necesario la generación de datos para las verificaciones, que no deben corresponder a la base de datos pero que si corresponden a la misma persona que es analizada. Las imágenes son procesadas transformándolas a escala de grises (figura 11^a), para luego normalizarlas eliminando las variaciones de intensidad de luz. Luego se aplica el algoritmo HOG para determinar las características a partir del histograma de gradiente orientado (figura 11b).



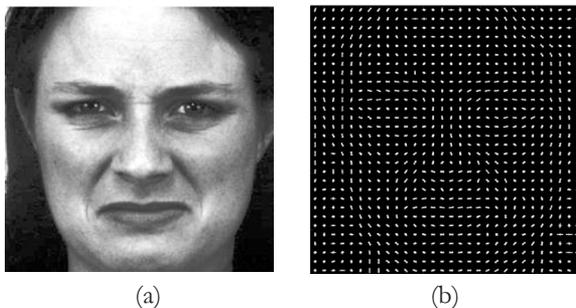


Figura 17. Aplicación de los algoritmos de normalización de la imagen.

Se destaca que el clasificador kNN (k-nearest neighbors algorithm) permite obtener la distancia a los puntos más cercanos permitiendo comparar por semejanza dos imágenes. Esta táctica se emplea para realizar una comparación sencilla entre resultados de otras técnicas de búsqueda, esta técnica se ha empleado de forma eficiente en investigaciones sobre: Bioinformática, procesamiento de imágenes y compresión de datos, recuperación de documentos, visión por computadora, base de datos multimedia, Análisis de datos de marketing.

La figura 12 representa la funcionalidad del algoritmo de HOG para determinar la distancia al punto más próximo de un conjunto de datos. En el caso de estas imágenes, la evaluación se realiza comparando los puntos obtenidos por el algoritmo de HOG de la imagen comparándola con los obtenidos mediante el entrenamiento con el uso de la base de datos.

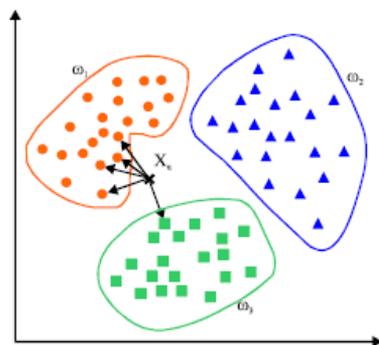


Figura 18. Funcionalidad del algoritmo de HOG para la evaluación de la



distancia más cercana.

Es posible evaluar una persona específica [1] para la determinación de su estado emocional para lo que se desarrolla un proceso de normalización de la imagen (figura 12), las imágenes utilizadas deben tener un tamaño de 160x160 pixeles (función `imresize` en Matlab); se lee la imagen con la función `imread` y se recorta el tamaño deseado de la cara con la función `imcrop`. Se convierte en imagen en escala de grises con la función `rgb2gray` para que pueda ser reconocido por el HOG.

La imagen capturada es comparada con todas las emociones almacenadas. El vector de salida `multisvm` arroja la ubicación de la imagen que es utilizada por la GUI para mostrar el estado emocional.

Elementos:

Home.m: ejecuta la interfaz para validar los métodos de clasificación y de extracción de características. Generando el `home.fig`

Export.m: crea una tabla de Excel en la que se van a escribir en la ventana de comandos todo lo que se ve en `home.fig`. Explica en la ventana de comandos la enumeración de las emociones.

HOG.m: extrae las características a partir de una imagen en escala de grises.

HOG_KNN: ejecuta el HOG y clasifica con KNN.

HOG_SVM: ejecuta el HOG y clasifica con SVM

Interfaz.m: permite la obtención de la imagen con la cámara, captura, entrenamiento de la red, detección de la emoción y visualización del tratamiento a través del audio y resultado final del paciente. Ejecuta la gráfica `interfaz.fig`

LBP.m: ejecuta el extractor de característica por el método LBP

Normalizar: permite el ajuste y normalización de las imágenes.

Paths.m: enruta a la lectura de las bases de datos para el proceso de entrenamiento. Carga las imágenes de las bases de datos.

Select_DB.m: selecciona el archivo que se va a extraer de la base de datos desde el Excel.



Xlswrite1.m: general el Excel para escribir las emociones con las características.

WSEFEP.xls: genera el formato excel

La figura 13 describe el proceso de extracción de características de la fisonomía facial, desde el proceso de entrenamiento hasta la detección final. En ella se observa el filtrado de la señal con el histograma HOG, así como la clasificación con la máquina de vector de estados.

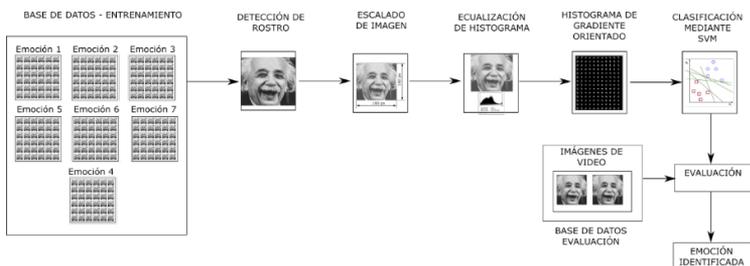


Figura 19. Esquema de funcionamiento de la caracterización de rostros.

Para lograr una efectiva caracterización facial, se ha utilizado una base de datos de entrenamiento con más de 300 imágenes que describen distintas emociones (ira, desprecio, asco, vergüenza, miedo, felicidad y neutral). Para la evaluación de la efectividad del entrenamiento se utilizaron datos recolectados de 80 estudiantes de la carrera de mecatrónica, que hicieron un total de 560 imágenes para pruebas.

Estimación de emociones mediante análisis de voz

La figura 18 muestra el desarrollo en Matlab para análisis de voz. En ella se perciben los procesos realizados para la estimación de emociones.



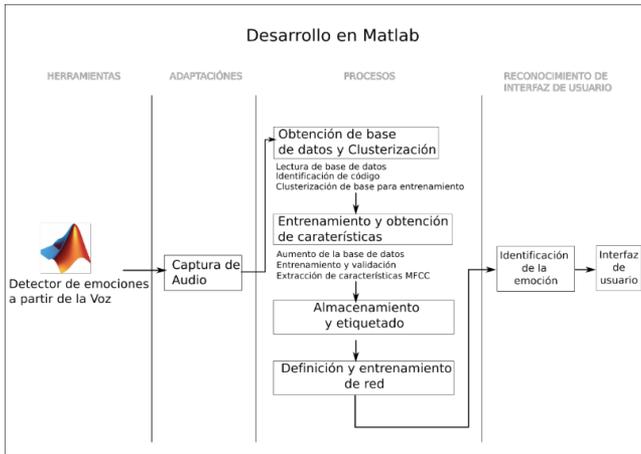


Figura 20. Estimador de voz

El análisis de la frecuencia de voz es uno de los factores más fundamentales en la determinación de emociones humanas [28], [89], [80]. Para el análisis del espectro de voz se consideró un proceso sencillo de determinación del fonema de voz, para el cual se realiza un proceso de lectura por parte de los pacientes, para extraer de forma clara y precisa el espectro de voz. La figura 19 muestra un esquema del proceso de extracción característica de la VOZ.

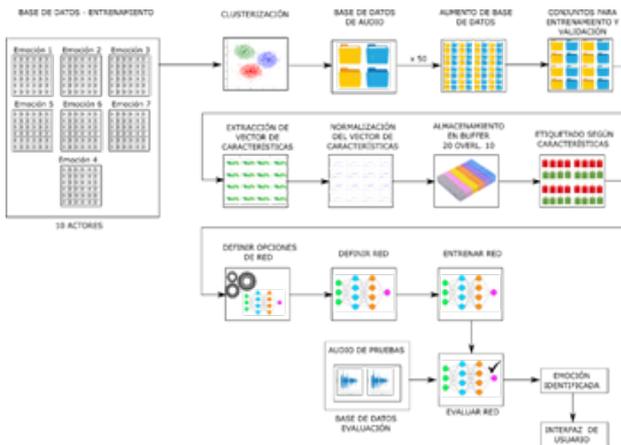


Figura 21. Diagrama general del proceso de extracción de características de la voz



Para el algoritmo de entrenamiento de voz se utilizó una base de datos de Berlin con 565 datos, que corresponden a las siete emociones básicas: ira, aburrimiento, disgusto, miedo, felicidad, tristeza y neutral. Los archivos se guardan en la variable datafolder y a continuación se guarda en la variable ads, los archivos de extensión .WAV.

```

Down Load Data Set

datafolder = PathToDatabase;
ads = audioDatastore(fullfile(datafolder,"wav")) ;
    
```

Figura 22. Código de recepción de datos de voz.

Los archivos de audio se identifican empleando letras y números para distinguir al hablante, el texto hablado, la emoción y la edad (tabla 2).

Tabla 2. Modelo de Identificación de hablantes

IDENTIFI CADOR	GÉNER O	EDAD (AÑOS)
03	Masculin o	31
08	Femenin o	34
09	Femenin o	21
10	Masculin o	32
11	Masculin o	26
12	Masculin o	30
13	Femenin o	32
14	Femenin o	35
15	Masculin o	25
16	Femenin o	31



En la tabla 2 se observan los identificadores de emociones para el sistema, tanto en idioma inglés como en germánico. El idioma español tiene los mismos identificadores que el inglés.

Tabla 3. Modelo de identificadores de emociones

LETRA	EMOCIÓN (SPANISH-ENGLISH)
A	Ira/anger
B	Aburrimiento/boredom
D	disgusto/disgust
F	Ansiedad/ anxiety/fear
H	felicidad/happiness
S	tristeza/sadness
N	Neutral

Para el análisis de espectro de voz se consideró la lectura de un texto fijo, que permita a la persona pronunciarlo para que su voz sea retenida en la base de datos (Tabla 3).

Tabla 4. Codificación de las frases de prueba.

CÓDIGO	FRASES PARA LA PRUEBA
A01	El mantel está sobre la mesa.
A02	Ella lo entregará el miércoles.
A04	Esta noche podría decirle a él.
A05	LA hoja de papel negra se encuentra allí, además del trozo de madera.
A07	En siete horas podría ser.
B01	Qué pasa con las bolsas debajo de la mesa?
B02	Simplemente lo llevaron arriba y ahora están bajando de nuevo.
B03	Actualmente siempre iba a casa y la veía.
B09	Descartaré esto y luego iré a tomar una copa con Omar.
B10	Podría estar en el lugar donde siempre lo guardamos.



Cada archivo de voz estará identificado por todas estos identificadores, así pues, la primera posición corresponderá al número que representa al hablante, las posiciones 3 al 5 corresponden al identificador del texto hablado, la posición seis a la emoción y la posición siete, a las versiones que representan, y se caracterizará con letras del abecedario.

Por ejemplo, el archivo de nombre 03A01Ha.wav corresponde a un audio de un hablante de 31 años de edad (03 en la tabla 1), que ha pronunciado el texto A01 (Tabla 3), con una emoción de felicidad (H) (Tabla 2). La versión del audio será a, que corresponde a la primera grabación de voz de ese individuo.

Se reemplazan los códigos de las emociones por sus nombres y se tabulan a fin de visualizar cuantos hablantes generan cuantas emociones.

```
Files ID
filepaths = ads.Files;
emotionCodes = cellfun(@(x)x(end-5),filepaths,'UniformOutput',false);
emotions = replace(emotionCodes,{'W','L','E','A','F','T','N'},...
{'Anger','Boredom','Disgust','Anxiety/Fear','Happiness','Sadness','Neutral'});

speakerCodes = cellfun(@(x)x(end-10:end-9),filepaths,'UniformOutput',false);
labelTable =
cell2table([speakerCodes,emotions],'VariableNames',{'Speaker','Emotion'});
labelTable.Emotion = categorical(labelTable.Emotion);
labelTable.Speaker = categorical(labelTable.Speaker);
head(labelTable)
```

Figura 23. Código de tabulación de hablantes y emociones.

Pruebas y funcionamiento

Dado que en los ensayos los distintos audios no concidirán exactamente con los pliegos de las ondas, es necesario incrementar en 50 veces, la base de datos, mediante la función `audioDataAugmenter`. Para crear un objeto `audioDataAugmenter` se puede considerar el cambio de tono a 0.5 o un rango determinado, el tiempo se cambia a 1seg y se usa un rango de -0.3 y 0.3 segundos. Se establece la posibilidad de agregar ruido a 1 y especificar su rango entre [-20, 40] dB.

Para cada archivo de audio:

1. Se crean 50 aumentos



2. Normaliza el audio a un valor máximo absoluto de 1
3. Se escribe el audio aumentado como archivo .WAV, se etiqueta luego de la versión con _augK, en donde k es el número de repetición como en la Tabla 4.

Tabla 5. Nombres de archivos de audio

Nombre
03*01Ha_aug01.wav
03*01Ha_aug02.wav
03*01Ha_aug03.wav
03*01Ha_aug04.wav

Los coeficientes cepstrales en las frecuencias de Mel (Mel Frequency Cepstral Coefficients) permiten representar el habla basado en la percepción auditiva humana. Para ello se realiza lo siguiente:

1. Separa la señal en pequeños tramos
2. Se le aplica la transformada de Fourier a cada tramo y obtiene la potencia espectral de la señal.
3. Aplicar filtros de la escala Mel y sumar las energías
4. Se toman el algoritmo de las energías de cada frecuencia mel
5. Se aplica la transformada de coseno discreta a esos logaritmos.
6. Evaluando 5 se obtienen los coeficientes.

Para realizar lo descrito anteriormente se emplea un objeto `audiFeatureExtractor`. Con una ventana periódica de 30ms, sin traslape. Se define una red `BiLSTM` empleando la función `bilstmLayer`, se coloca una capa `DropoutLayer`, antes y después del `bilstmLayer` para prevenir un sobre ajuste. La función `BiLSTM` permite que la red aprenda de la serie de tiempo completa en cada paso de tiempo. La `DropoutLayer` es una capa que permite establecer aleatoriamente los elementos de entrada en cero con una probabilidad dada.

El proceso para la estimación de emoción a través de la voz, se puede resumir en los siguientes pasos:



1. Se divide la base de datos de audio en conjuntos de entrenamiento y validación.
2. Se extraen vectores de características de los conjuntos de entrenamiento y validación.
3. Se normalizan los vectores de características.
4. Se protegen los vectores de características en secuencias de 20 con superposiciones de 10.
5. Se replican las etiquetas para que estén en correspondencia uno a uno con los vectores de características. Se crea una matriz vacía para contener las categorías emocionales y se agrega a las etiquetas de validación a fin de que la matriz categórica contenga las emociones.
6. Se definen opciones de entrenamiento.
7. Se define la red neuronal, que en este caso es una red de tipo gradiente descendiente.
8. Se entrena a la red.
9. Se evalúa la red.

En este capítulo se ha descrito el algoritmo que propuesto en esta tesis, se han considerado las descripciones individuales de las fases del trabajo, con los aspectos fundamentales que componen cada fase del algoritmo de detección de estrés. De esta manera se han resaltado los aportes realizados en el desarrollo del software, tomando en cuenta la detección de rostros a partir de imágenes de video y la vinculación de dos procesos; la detección de rostros y la detección de voz.

En el capítulo siguiente se destacan los resultados del trabajo aquí planteado, se describen los aspectos del sistema de reconocimiento de estrés diferenciando las dos fases que lo componen, voz y rostro.



Capítulo 5: Hallazgos de la ingeniería en la detección de emociones.

En este capítulo se presentan las simulaciones realizadas, mediante el planteamiento de casos de estudio y pruebas experimentales.

Computación inteligente en el análisis de emociones humanas

El sistema de clasificación de las emociones se logra a partir de dos variables: rostro y espectro de voz, esta clasificación es condicionante para la determinación de estrés en las personas. La figura 9 (Capítulo III) muestra la interfaz de usuario con los elementos propios del sistema, esta permite al especialista médico obtener los datos del paciente y estimar su estado de salud emocional. La interfaz cuenta con una opción para activar la cámara y capturar la foto del paciente, además muestra la normalización de la imagen, que es la que se va a procesar para ser tratada por la red neuronal de tipo gradiente descendiente y entregar ese resultado a la fase de concatenación de variables y clasificación. También cuenta la interfaz con el botón de grabación de voz, para lo cual se ha incluido un texto que pueda ser leído por la persona evaluada, de tal manera que el sistema pueda captar el espectro de voz y procesarlo.

Los métodos de reconocimiento de emociones se evaluaron mediante estimadores con múltiples métodos, tanto para análisis facial como reconocimiento de voz.

Existe una amplia variedad de estimadores entre los cuales los más eficientes trabajan mediante redes neuronales, lo que le permite refinar la identificación de la emoción y por tanto su desempeño en el tiempo. Se realizaron pruebas con diferentes bases de datos y diferentes estimadores. Como se describió en el capítulo anterior. Por otra parte se consideraron cuatro bases



de datos, llamadas JAFFE, KDEF, MUG, WSEFEP, y se tomaron en cuenta dos algoritmos para extracción de características, HOG y LBP, para dos clasificadores, KNN y SVM. Para poder determinar la mejor combinación de estos elementos, se llevaron a cabo pruebas que permitían evaluar el error porcentual de los mismos (figura 9, capítulo III). El desempeño del sistema con el algoritmo SVM mostró una efectividad del 71,42% mientras que su mejor combinación HOG+SVM, ofrece un 96,13% de precisión, presenta incertidumbres no verificadas dado que las emociones no son constantes en el tiempo, en un período corto de 25 segundos se ha observado cambios de hasta 3 emociones generadas posiblemente por recuerdos o asociaciones mentales del actor y considerando que las emociones no son constantes en todos los individuos.

Se determinó que los estimadores son susceptibles de rasgos faciales referentes a la zona geográfica, etnia y características de la población. Además, un factor determinante en análisis de imagen es el entorno y condiciones de iluminación, deben ser estandarizados a fin de mejorar las estimaciones.

En múltiples ensayos realizados en la identificación de emoción, en base a la voz, se identifican como aspectos que afectan de forma relevante para una buena estimación a los parámetros como sexo, edad y condición física y emocional.

Las pruebas realizadas proporcionan un resultado aceptable para la identificación del estrés, sin embargo, este análisis debe refinarse con un sistema que mantenga la emoción por un tiempo prolongado a fin de mejorar esa predicción e incertidumbre. Una alternativa considerada fue el análisis en video [23], logrando capturar de manera adecuada mucha más información pero debe tomarse en cuenta que según pruebas, la precisión se reduce en el mejor de los casos a un 87%, sin embargo al combinar las variable de voz este valor mejora.

El clasificador empleado en este trabajo, es el tipo Adam, que es una combinación del Stochastic Gradient Descent with momentum y RMSprop [98]. Este algoritmo de optimización promete una mayor eficiencia que otros algoritmos, haciendo que el error de la red llegue al mínimo en el menor tiempo posible, y posee similar precisión que métodos como:

-Convolutional neural network (CNN) con un 97%



-Patched based 3D Gabor con 94.48% (Base C-K)

-SVM with clusterin algoritms, 94.34%

-PHONG (Pyramid Histogram of Oriented Gradient), con 96.33%

-Sparse Representiom con 94.70% (JAFFE)

Como se aprecia y según se pudo corroborar, la calidad de la base de datos empleada afecta directamente la precisión, por tanto, este método empleado debe considerarse necesariamente para entornos estructurados, iluminación constante e invariante.

Análisis de emociones a través de la voz

Con respecto al reconocimiento por voz, el estimador contempla para los ensayos, personas mayores a 20 años, también se basa en el género, sin embargo, la incertidumbres en este ámbito pueden reducirse asociando las emociones similares como por ejemplo la sorpresa y disgusto, pueden generar patrones de sonido similares, por tanto unirlos podrían disminuir el error considerablemente.

Con las pruebas realizadas se detrmirió que los actores aportan archivos de audio de referencia, más no son archivos que representen una emoción real, lo cual aporta en la incertidumbre, por su calidad de ser forzados a mostrar una emoción para la creación y entrenamiento de las redes neuronales empleadas. En la figura 23 se puede apreciar el pseudocódigo del análisis de voz.



```
ESTIMADOR EMOCIÓN MEDIANTE VOZ, speech emotion recognition (SER)

Entrenamiento de estimador
| leer audio de base de datos clasificada
| clusterización
| asignar etiquetas
| aumentar base de datos
| crear base de entrenamiento y testeo
| extracción del vector de características
| clasificación mediante SVM
| estandarización del vector de características
| guardar en buffer
| etiquetado según características
| definir red de entrenamiento
fin
evaluación de la red
| leer audio de base de datos para test
| Estimación de emoción
| Si emoción=1
|   Clase=='Enojado'
| Si emoción=2
|   Clase=='Disgustado'
| Si emoción=3
|   Clase=='Asustado'
| Si emoción=4
|   Clase=='Feliz'
| Si emoción=5
|   Clase=='Neutral'
| Si emoción=6
|   Clase=='Triste'
| Si emoción=7
|   Clase=='Sorprendido'
| fin
| escribir en tabla de excel
| matriz de confusión
fin
```

Análisis de fisionomía del rostro.

En lo que respecta a la emoción mediante análisis del rostro, el gran número de músculos existentes en la cara se prestan para múltiples variantes de expresiones en una misma emoción, y además estas dependen del entor-



no circundante que puede afectar notoriamente en las múltiples pruebas y creación de bases de datos.

Se evidenció que si bien, el manejo de imágenes de mediana resolución, aporta mejor información, se debe balancear el coste computacional versus la calidad de la información a fin de optimizar los resultados para una aplicación específica.

Se deduce que la identificación de la emoción mediante análisis de imagen puede mejorarse y constituye un área de continuo desarrollo y mejoramiento mediante aplicación de técnicas de inteligencia artificial.

Se constató que una persona evaluada bajo la presencia de otra persona, en calidad de supervisor, afecta la emoción dado que puede desarrollar conductas espontáneas, lo que añade errores en el sistema, dificultando el estimador.

La unión de tres variables para la identificación de las emociones resulta un método más certero que el uso de un estimador único, dado que corrobora y aporta con información, de forma homóloga a la participación de varias opiniones en observadores. Por tanto la estimación, evidencia mayor certeza. La figura 24 muestra el pseudocódigo de procesamiento de imagen del rostro.



```
ESTIMADOR EMOCIÓN MEDIANTE IMAGEN FACIAL (HOG+SVM)

Entrenamiento de estimador
  leer imagen de base de datos clasificada
  detección de rostro
  escalado de imagen
  equalización de histograma
  histograma de gradiente orientado HOG
  clasificación mediante SVM
  Clusterización
    Si emoción==1
      Clase=='Enojado'
    Si emoción==2
      Clase=='Disgustado'
    Si emoción==3
      Clase=='Asustado'
    Si emoción==4
      Clase=='Feliz'
    Si emoción==5
      Clase=='Neutral'
    Si emoción==6
      Clase=='Triste'
    Si emoción==7
      Clase=='Sorprendido'
  fin
fin

evaluación del clasificador
  leer imagen de base de datos clasificada
  detección de rostro
  escalado de imagen
  Estimación de emoción
    Si emoción==1
      Clase=='Enojado'
    Si emoción==2
      Clase=='Disgustado'
    Si emoción==3
      Clase=='Asustado'
    Si emoción==4
      Clase=='Feliz'
    Si emoción==5
      Clase=='Neutral'
    Si emoción==6
      Clase=='Triste'
    Si emoción==7
      Clase=='Sorprendido'
  fin
  escribir en tabla de excel
  matriz de confusión
fin
```

Figura 26. Pseudocódigo de procesamiento de imagen de rostro.



Identificación del rostro dependiendo de la distancia

Se consideró la evaluación del rostro desde diferentes distancias entre el usuario y el computador donde está la cámara web, en principio se evaluó con el usuario en una postura normal típica de una sala de atención médica, pero la detección del rostro presentó fallas para el reconocimiento de las imágenes. Arrojando en estos casos datos incongruentes sin información precisa. De tal manera que se procedió a evaluar una distancia más corta y con una postura más firme, con la columna más recta y con el rostro frontal a la cámara. De esta manera se obtuvieron resultados más precisos, pero sin embargo esta necesidad de ajuste de la postura afecta las condiciones emocionales naturales de la persona. Se logró establecer una distancia óptima de 45cm entre la cámara y el paciente.

Iluminación del ambiente en la captura de la imagen del rostro.

Se realizaron pruebas con entornos de trabajo normales, sin ningún tipo de modificaciones y se observó que las imágenes capturadas no tenían las características apropiadas para poder procesarse efectivamente, la luz incidía de forma diferente según el tipo de piel y no era posible una uniformidad en las imágenes que permitiera incluirlas en la base de datos. Fue necesario la adecuación de un espacio exclusivo para la captura de imágenes, por lo que al entrevistar al especialista médico se acordó que dicho análisis debe realizarse en un espacio único del consultorio donde la iluminación sea la apropiada y el fondo de la imagen sea de color blanco, para lograr una mejor captura del rostro con todas las características naturales de la persona a estudiar.

Análisis del espectro de voz

La figura 25 muestra el pseudocódigo para análisis de voz.



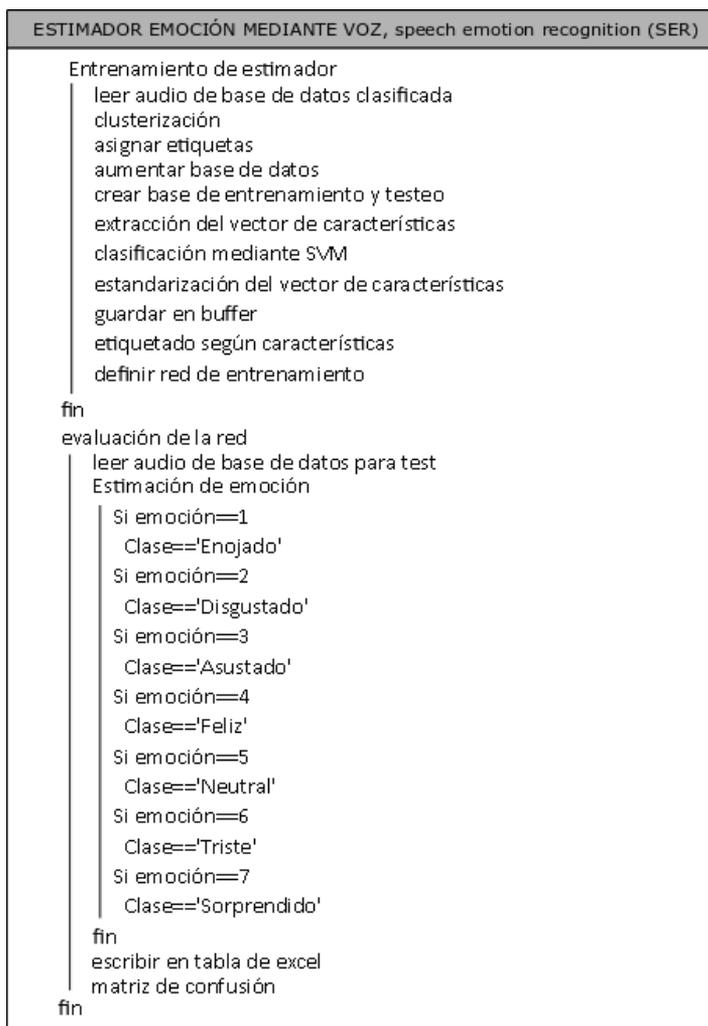


Figura 27. Pseudocódigo de espectro de voz.

Optimización de los algoritmos

Se realizó un proceso de optimización con el código de métricas de Mat-



lab, logrando evaluar la velocidad de procesamiento con la función complexity, además esta evaluación de métricas avisa cuáles líneas de código deben ser mejoradas a través de la función mlint. Este código de evaluación permite reconocer los conflictos que pudieran presentarse, usando la función clash. Por otro lado con la función clone, se reconocen cuáles sentencias están repetidas, y además existe una función comments expresa las recomendaciones del sistema, finalmente la función details reporta el código con la métrica. En la figura 26 se observa la evaluación de métricas del código.

ANÁLISIS CON CODEMETRIC
Analizar mediante codemetric
leer dirección de archivo
crear dirección de destino para resultados
busqueda de archivos
Generación de estados de archivos
evaluar líneas
evaluar número de comentarios
evaluar numero de ayudas
evaluar mlint
evaluar clash
evaluar clone
escribir tabla con resultados evaluados
creación de reporte

Figura 28. Análisis de métricas del código empleado

Esta evaluación de métricas fue complementada con en el cálculo de la precisión, para lo cual se realizó un análisis entre total de predicciones y el número de aciertos de estas. Así la precisión fue analizada siguiendo el proceso de predicciones del algoritmo (Figura 27). Luego se verificó el cálculo de forma manual con la ecuación (32) y con la matriz de confusión (Figura 28).

Figura 29. Estimación de precisión con las predicciones arrojadas por el algoritmo.

Para el cálculo manual de precisión se ha considerado la ecuación (32), tomando en cuenta la matriz de confusión.



	Train images	Test images	True Expressions	Accuracy %
Angry	42	1	1	100
Disgusted	42	1	0	0
Fear	36	1	0	0
Happy	44	1	1	100
Neutral	42	1	1	100
Sad	37	1	1	100
Surprised	32	1	1	100
Total	275	7	5	71.4286

Figura 29. Estimación de precisión con las predicciones arrojadas por el algoritmo.

Para el cálculo manual de precisión se ha considerado la ecuación (32), tomando en cuenta la matriz de confusión.

$$p_{\text{precisión}}^{\zeta} = \frac{\text{predicciones acertadas}}{\text{predicciones totales}} \times 100\% \quad (32)$$

$$\text{precisión} = \frac{5}{7} \times 100\% = 71,42\%$$

La matriz de confusión de la figura 28 muestra cinco emociones acertadas del total de siete emociones evaluadas.

Angry	1	0	0	0	0	0	0
Disgusted	0	0	0	0	0	0	1
Fear	0	1	0	0	0	0	0
Happy	0	0	0	1	0	0	0
Neutral	0	0	0	0	1	0	0
Sad	0	0	0	0	0	1	0
Surprised	0	0	0	0	0	0	1
	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised

Figura 30. Matriz de confusión con 5 predicciones acertadas de las siete predicciones totales.



Demostrando que la precisión arrojada por el algoritmo es exactamente la misma que la encontrada de forma manual, a través de la matriz de confusión.

Para el proceso de optimización se realiza la correlación entre imágenes y la correlación entre espectros de voz, así como entre señales de pulso, esta correlación mide la relación de una variable respecto a la otra para evaluar la confiabilidad de la misma, obteniéndose que el porcentaje de correlación estará entre cero y uno; para valores cercanos a uno se logra una correlación bastante similar entre ambas variables, para valores lejanos de uno, la similitud entre ambos datos no existe o es distante. Una muestra de esto puede observarse en la figura 29, de 500x600 píxeles, la correlación entre el estado emocional molesto (Angry) y el estado feliz (happy), en esta evaluación se obtuvieron coeficientes de correlación de 0,863. Mientras que en la figura 230 se observa la correlación entre dos imágenes de la misma emoción (angry) que arrojaron una correlación de 1.

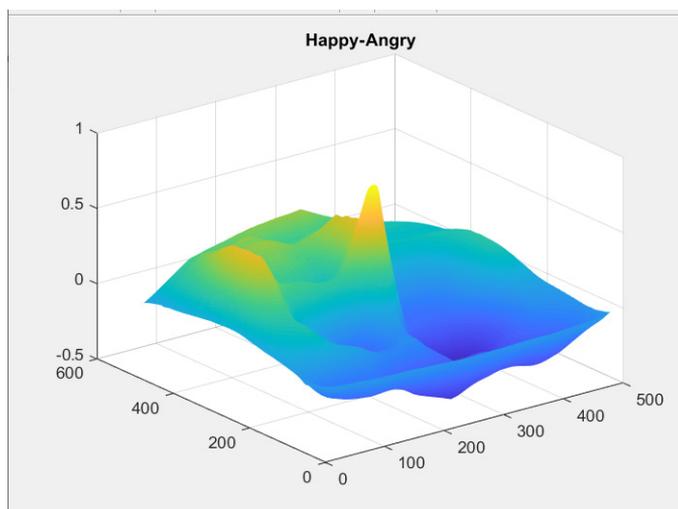


Figura 31. Análisis de correlación happy-angry con coeficiente de correlación 0,863.



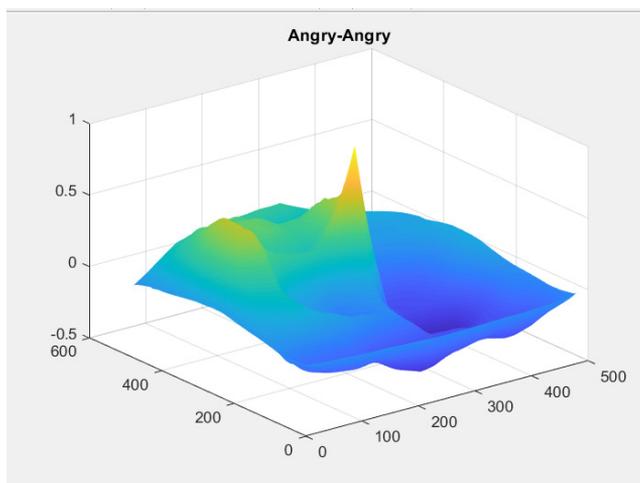


Figura 32. Análisis de correlación angry-angry con coeficiente de correlación 1.

Los códigos utilizados para este proceso de optimización, se observan en el anexo 3.

Este proceso permite evaluar la similitud entre dos datos de una misma variable, y permite a su vez realizar los ajustes necesarios cuando se considere que dicha similitud no es la apropiada.

Validación

Pruebas con especialista médico

Se hicieron pruebas con especialista médico, el Dr. Simón Bolívar Grueso, médico general de la ciudad de Esmeraldas, realizó pruebas de funcionamiento y además aportó en el diseño de la interfaz de usuario. Considerando que esta debía ser más amigable al usuario, con señalizaciones más sencillas y con más indicaciones que le permitan al especialista de salud entender el proceso y hacer el seguimiento de la toma de datos. Además, el Dr. Simón Grueso sugirió que se disponga un espacio específico en el consultorio para el procesamiento de estrés en la persona, porque la misma sensación de estar



siendo evaluado puede causar estrés. El especialista también agregó que ha conocido varios casos de muertes por estrés, que esta herramienta le parece oportuna para contribuir al diagnóstico de enfermedades y que podría ser de gran ayuda para los pacientes y sus médicos tratantes.

Entorno de trabajo

El entorno de trabajo consta de una interfaz de usuario donde se pueden apreciar los elementos de la captura de variables para la estimación de la emoción que puede definir el estrés en las personas. El sistema consta de un párrafo para la lectura y captura de la frecuencia de voz. Se hicieron pruebas para definir el texto de este párrafo y se observó que algunos textos no eran lo más apropiado porque interferían en la receptividad de las personas para determinados temas, por ejemplo, un texto infantil puede causar incomodidad a algunas personas, así como textos de carácter político, o con alguna tendencia específica. Se optó por un texto plano sin tendencias definidas.

Se incluyó en la interfaz un proceso por pasos para la ejecución del sistema, que permita a un especialista médico realizar el análisis sin mayores complicaciones, de forma autónoma y sin que esto le permita alguna pérdida de tiempo en el proceso.

Por la entrevista sostenida con el Dr. Simón Bolívar Grueso, especialista de medicina general, fue posible descubrir que la interfaz debía tener cada proceso por separado que le permita al médico hacer la gestión de forma fácil y con botones explícitos del proceso. Por lo que se incluyeron botones que describen el proceso que realiza cada acción en vez de botones de acción generalizada. La figura 31 muestra los elementos de la interfaz de usuario.



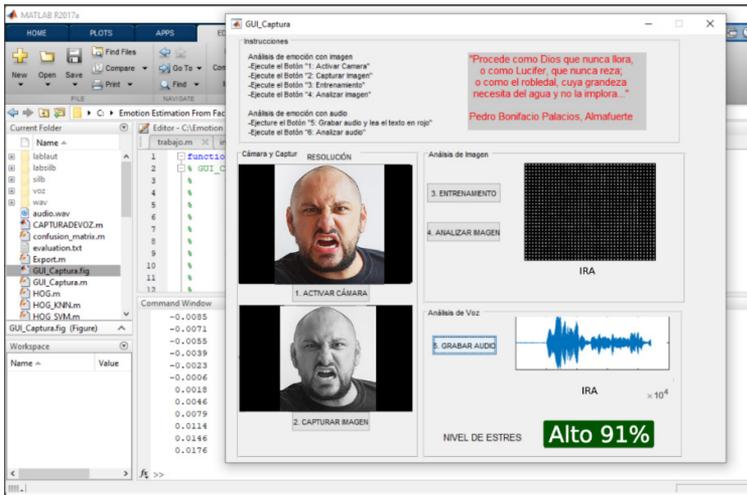


Figura 33. Interfaz de usuario para la estimación de emociones que expresan estrés.

Estos datos se ajustan en la fase de concatenación de variables y clasificación (figura 32), donde a través de un código sencillo se combinan para definir el estado de estrés de la persona.

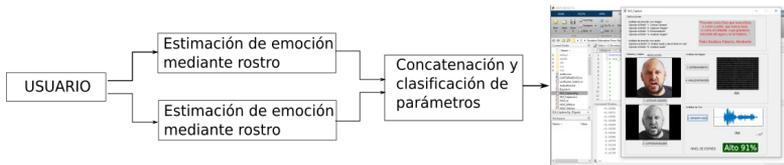


Figura 34. Diagrama de funcionamiento con la fase de concatenación y clasificación.

El sistema cuenta con una integración de tecnologías y variables biomédicas que hacen posible un resultado más óptimo. Algunos autores [39], [15], [41], [82], [76], [60] sugieren que no es posible una detección eficaz del estrés con un único parámetro, y que para lograr una mejor respuesta era necesaria la integración de otros parámetros fisiológicos que intervengan que el proceso de manifestación de emociones.



Análisis Estadístico

Se realizó una prueba tipo test a 80 estudiantes universitarios, la cual consistió en un conjunto de preguntas desarrolladas por el instituto de medicina de salud social de México. El test cuenta con 12 preguntas relacionadas con las principales manifestaciones de estrés, y según estos resultados el cuestionario permite conocer sobre el estado de salud emocional de las personas, pudiendo clasificar en seis tipos de estrés, que van desde el más leve hasta el más elevado. La prueba realizada a los estudiantes fue contrastada con los resultados obtenidos del sistema informático con este mismo grupo de personas. Partiendo de esta premisa la muestra fue de 80 personas voluntarias.

El test fue desarrollado por los investigadores del Ministerio de Salud de México [99]. Para la validación del test se utilizó el criterio de expertos. Se empleó el Método Delphi de validación de expertos para el tratamiento del cuestionario. Para el efecto se crearon dos grupos de trabajo cooperativo y se consideró inicialmente un cuestionario estructurado, que sometido a la valoración de los expertos se le fueron agregando o modificando los criterios siguiendo las opiniones de los expertos durante tres rondas sucesivas. Todas las opiniones se sometieron a un análisis estadístico. Se obtuvo un coeficiente de correlación máxima de 0.94, lo que evidencia un alto grado de validez. Para la fiabilidad se utilizó el método de consistencia interna alfa de Cronbach con un coeficiente de 0.811, y un nivel de confianza del 95% ($p \leq 0.05$), mientras que para el método de dos mitades, el coeficiente de consistencia interna en el alfa de Cronbach de la primera mitad es de 0.766 y el de la segunda mitad de 0.738, los coeficientes resultantes de la prueba de dos mitades indican un índice de fiabilidad muy alta. Se realizó una rúbrica para el tratamiento estadístico, compuesto por 10 ítems.

El test aplicado (Anexo 2) estuvo compuesto por un cuestionario de 12 preguntas básicas sobre las condiciones de la persona para asumir un día común en los últimos tres meses. Este test fue adaptado del cuestionario de problemas psicosomáticos propuesto por el Ministerio de Salud de México [100]. Se consideró este cuestionario porque consta de preguntas sencillas sin mayores complicaciones y expone respuestas claras que son entendidas con facilidad por cualquier persona. El test se empleó en 80 personas mayores de 20 años y menores de 30 años que fueron voluntarias. La tabla 6 muestra los resultados del test clásico que se realizó a las 80 personas.



Tabla 6. Resultados del test.

PREGUNTAS	RESULTADOS DEL TEST					
	1 Nunca	2 Casi nunca	3 Pocas veces	4 Algunas veces	5 Relativamente frecuente	6 Muy Frecuente
Imposibilidad de conciliar el sueño.	15	22	18	16		9
Jaquecas y dolores de cabeza	12	26	33			9
Indigestiones o molestias gastrointestinales.	34	25	12		9	
Sensación de cansancio extremo o agotamiento.		41		16	14	9
Tendencia de comer, beber o fumar más de lo habitual	2	28	43			7
Disminución del interés sexual.	46	20	5	9		
Respiración entre cortada o sensación de ahogo.	71	9				
Disminución del apetito.	64	9	7			
Temblores musculares (por ejemplo tics nerviosos o parpadeos).	67	9	2	2		
Pinchazos o sensaciones dolorosas en distintas partes del cuerpo.	69	9	2			
Tentaciones fuertes de no levantarse por la mañana.					71	9
Tendencias a sudar o palpitaciones.	17	37	17		9	
Coefficiente de variación	0,7	0,5	0,9	0,6	1,2	0,1

Se observa en la tabla 6 que de 80 personas encuestadas el 11,25% presentan un elevado nivel de estrés. Mientras el 88,75% restante presenta estrés moderado, con tendencia a nulo.

El coeficiente de correlación (Tabla 5) indica que hay correlación negativa entre las variables, observándose que hay una correlación inversa entre algunas opciones seleccionadas por los entrevistados. Sin embargo, se observa en la tabla 4 el coeficiente de variación con 0,2 en la opción muy frecuente, lo cual indica que una pequeña parte de la muestra presenta estrés (9 personas).

Por otra parte el alfa de cronbach determinó un 0,81% de confiabilidad, lo cual permitió considerar que el cuestionario realizado era confiable para el estudio.

Tabla 7. Coeficiente de correlación para datos recolectados en el test

	Nunca	Casi nunca	Pocas veces	Algunas veces	Relativamente frecuente	Muy Frecuente
Nunca	1					
Casi nunca	-0,90448	1				
Pocas veces	-0,8894	0,67025498	1			
Algunas veces	-0,99932	0,83297294	0,965351284	1		
Relativamente frecuente	N/A	0,68204829	N/A	N/A	1	
Muy Frecuente	0,979076	0,10136755	-0,80295507	N/A	N/A	1



Se compararon los datos observados en el test con los datos observados en el sistema, obteniendo un error de 3,75% para un conjunto de 80 personas evaluadas.

El sistema informático logró detectar 12 personas con nivel de estrés elevado (96,13% de confiabilidad), mientras el test escrito detectó solo 9 personas. Esto se debe posiblemente a que las interrogantes del test pueden estar sujetas a condiciones culturales, sociales, formativas que no son generales para todos los evaluados. Además el sistema informático puede captar elementos fisiológicos de las personas sin necesidad de considerar las características socio-culturales de las mismas, pudiendo aislar el análisis de estrés presente sin evaluar las causas que lo producen.

La herramienta no ofrece un diagnóstico médico, es un desarrollo de ingeniería que sirve de complemento para la evaluación médica profesional.

Los resultados obtenidos en el sistema informático presentan un porcentaje de acierto del 96,13% en comparación con los resultados del test escrito. Logrando una efectiva respuesta en la detección del problema de salud. De forma aislada los dos parámetros reflejan resultados parecidos, con un porcentaje de acierto de 95, 56% en comparación con el test escrito.

Otros resultados de medición de estrés

Otros estudios de medición de estrés han podido demostrar que el algoritmo de SVM es el más eficiente y con mínimo error para el tratamiento de imágenes, así como para la clasificación de frecuencias de audio (voz). La tabla 8 muestra algunos resultados obtenidos en cuanto a algoritmos de SVM [101] y en evaluaciones de estrés.



Tabla 8. Comparación con otros estudios.

Nº	Nombre del investigador	Método	Base de datos	Confiabilidad
1	Philipp & Rana [102]	SVM	(CK&CK+) and live video	87.90%
2	Tommaso, Caifeng, Vincent & Ralph [103]	HOG+SVM; LBP+SVM; LTD+SVM	CK	92.70%
3	KHARAT & DUDUL [104]	DCT+SVM; FFT +SVM; SVD +SVM	JAFFE	94.29 %
4	Zhiguo & Xuehong [105]	WPCA+SVM;PPCA+SVM	CKACFEID	84.75%
5	Abhinav, Akshay, Yogesh & Tom [106]	PHOG+SVM	GEMEPFERA	67%
6	Junkai, Zenghai, Zheru & Hong [107]	HOG+SVM	JAFFE	94.30%
7	Manar, Aliaa & Atallah [108]	HOG+SVM	CK2 & random DB	95%
8	Carcagni, Coco, Leo & Distanto [109]	HOG+SVM	CK2	95.9%
9	Sandeep, Shubh, Yogesh & Neeta [110]	DCT+WT+GF-GT+ADB	JAFFEE	93,40%
10	M. Trianes, M. Blanca, F. Fernández, M. Escobar, E. Maldonado y Á. Muñoz. [32]	Encuesta	60 Personas	96%
11	C. Betancourt y I. Yépez [38]	SVM considerando la edad y el género	40 Personas	87%
12	J. Aróztegui Vélez [4]	Realidad virtual	30 Personas	70%
13	C. Valderrama, C. Ulloa [43]	Transformada discreta wavelet	20 personas	72,34%
14	F. Suárez y L. Rosales	SVM+HOG+ADAM	Base de datos	96,13%

La aplicación del algoritmo y método matemático definido con datos válidos de otros estudios permite concluir en base al nivel de confiabilidad que tiene una validez sólida.

Los mejores resultados observados en la tabla 8 son aquellos donde el SVM es combinado con el algoritmo HOG, ofreciendo una confiabilidad superior al 90% en casi todos los casos [111], [112], [113], [114], [115], [6],



[116], [117].

Es relevante destacar que se constató que la toma de datos de las imágenes de rostro no puede realizarse en entornos oscuros o con poca luz, ya que la definición de los gestos no es perceptible por la cámara, por tanto el reconocimiento facial posee un mejor desempeño para entornos que posean una luz uniforme y que posea un fondo constante para la evaluación. Se propone en trabajos posteriores la implementación de un detector de rostros para lo cual también debe considerarse una mayor resolución superior a los 150x150 píxeles, lo cual obliga a emplear procesadores de más capacidad y superiores a los 36 GB de memoria RAM a fin de no demorar mucho en el análisis.

El análisis de las siete emociones realizado, puede reducirse a las emociones más representativas para cuestiones de identificación del estrés, así, el miedo y la sorpresa poseen ciertas similitudes, de igual manera con el enojo y el asco. Si se agrupan estas emociones, se pueden reducir hasta en un 30% los errores según Bartual [82].

El reconocimiento facial es de gran utilidad para la identificación de personas, ya que es un método no invasivo y que no requiere de mucha colaboración por parte del usuario al momento de la detección.

Contar con una cámara de una gran resolución colocada en una posición favorable y una buena iluminación para la identificación de los usuarios, permite que el reconocimiento del usuario sea más eficaz, con mayor precisión y que no se vea afectado por elementos externos.

Por otra parte, las redes neuronales artificiales ofrecen una opción para los problemas de clasificación, su complejidad dependerá del número de capas que ellas tengan. Entre más capas tenga la red mayor será su capacidad de acierto, pero también requerirá un mayor consumo de cómputo. Cuando la red neuronal tiene un gran número de capas ocultas se les llama Deep Learning (aprendizaje profundo). Este tipo de entrenamiento requiere técnicas específicas y costosas tanto en tiempo como recursos computacionales. Su nivel de complejidad radica en que no permite la actualización de parámetros de forma simultánea, sino que los pesos son actualizados de capa en capa, ocasionando lentitud en el tiempo de procesamiento. El gran problema de este método es que conforme queremos visualizar neuronas



de capas más profundas, los patrones visuales generados tienden a volverse irrealistas y no interpretables.

Para problemas de clasificación se conocen múltiples algoritmos que ofrecen soluciones aceptables para lograr los objetivos. Uno de los algoritmos más sencillos es el de regresión logística, que permite estimar un modelo en el que la variable de salida se expresa linealmente en términos de la variable de entrada, asociado a un conjunto de parámetros para el entrenamiento [95], pero a medida que aumenta el número de datos se hace imposible el procesamiento para lograr el error mínimo aceptable. De ahí que se emplee el método de descenso de gradiente, este método consiste en partir de unos valores aleatorios de los parámetros del modelo e ir actualizando dichos valores iterativamente de modo que converjan hacia el valor en el que el gradiente del error con respecto al parámetro sea cero.

La evaluación de modelos es fundamental en Machine Learning, sin embargo no se menciona cómo evaluarlos. Existen varias métricas para analizar el error de los modelos. La que se ha utilizado en este trabajo es el código de evaluación de métricas de Matlab, que permite conocer los aspectos más relevantes del código desarrollado como velocidad, posibles conflictos, repetibilidad de líneas, sentencias que deben mejorarse, entre otros. Esta métrica a pesar de ser útil, no tiene en cuenta el número de datos de cada variable y esto impide reflejar fielmente la calidad del clasificador. Una métrica sin ese problema y que determina la precisión, es la matriz de confusión, que consiste en calcular el porcentaje de aciertos en los ejemplos predichos como una determinada. La matriz de confusión de un problema de clase n es una matriz $n \times n$ en la que las filas se nombran según las clases reales y las columnas, según las clases previstas por el modelo. De esta manera nos ha permitido mostrar de forma explícita cuándo una clase es confundida con otra.

Delimitación

El trabajo se limitó al uso de SVM con algoritmo de HOG. La investigación tampoco trata la comparación o escogencia de otros softwares ni otros lenguajes de programación. La investigación solo fue probada en aplicaciones médicas, únicamente para adultos mayores de 19 años. Solo se valoraron



las emociones conducentes al estrés, mas no fue interés en categorizarlo. Se limitó además, el número de personas para la muestra por las restricciones de permiso, consideraciones éticas y consentimiento informado. No fue objetivo de esta investigación diseñar un prototipo.



Bibliografía

- [1]C. Ibañez, «Charles Darwin,» Sociedad, pp. 38-42, 2009.
- [2] D. Palacios, Contribución al estudio de selección de parámetros para identificación de estrés en la voz, España: Universidad Politécnica de Madrid, 2017.
- [3]J. A. Amador Moncada, H. A. Granada Díaz, J. M. Redondo Ostegón y G. O. Tost, «Dinámicas no lineales y no suaves en procesos estrés-enfermedad,» Rev. Ciencia y Desarrollo., vol. 8, n° 1, pp. 9-19, 2017.
- [4]J. Aróztegui, J. Fernández y A. Manzanero, «Visualización hiper-dimensional en la investigación sobre la detección,» Repositorio Universidad de Madrid, Madrid, 2018.
- [5]R. Bartual y J. Herranz, «Reconocimiento facial y reconocimiento anímico mediante análisis facial,» Universidad Politécnica de Valencia, Valencia, 2017.
- [6]J. Chen y e. al., «Facial expression recognition based on facial components detection and hog features,» de International Workshops on Electrical and Computer Engineering Subfields. , 2014.
- [7]P. Fernández-Berrocal, Darwin y el misterio de las emociones, Málaga: UMA. SEDOC, 2009.
- [8]E. Ramírez, «Antropología “compleja” de las emociones humanas,» Isegoría, vol. 25, pp. 177-200, 2001.
- [9]H. Harlow y M. Harlow, «Social Deprivation in Monkey,» Scientific American, n° 207, pp. 136-146, 1962.
- [10]M. Reite y R. Short, «Attachment, loss and depression,» Journal of Child Psychology and Psychiatr, n° 2, pp. 141-170, 1981.
- [11]A. Montagu, «Natural Selection and the Origin and Evolution of Weeping in Man,» Science, n° 130, pp. 1572-1573, 1959.
- [12]A. Surrallés, «Afectividad y epistemología de las ciencias huamanas,» Revista de Antropología Iberoamericana, pp. 1-15, 2005.
- [13]G. Bourdin, «antropology of the emotions: cocepts and trends,» Revista de Ciencias Antropológicas, n° 67, pp. 55-74, 2016.
- [14]A. Wierzbicka, «Human Emotions: Universal or Culture-Specific?,» American Anthropologist, vol. 88, n° 3, pp. 584-594, 1986.
- [15]C. Duque y M. Morales, «Caracterización de voz empleando análisis tiempo-frecuencia aplicacda al reconocimiento de emociones,» Universidad Tecnológica de Pereira, Pereira, 2007.
- [16]P. Moure, «De lo psicológico a lo fisiológico en la relación entre emociones y salud,» Revista Psicología Científica, vol. 13, n° 19, pp. 1-8, 2011.
- [17]M. Naranjo, «Una revisión teórica sobre el estrpes y algunos aspectos



- relevantes de este en el ámbito educativo.» *Revista Educación*, vol. 33, n° 2, pp. 171-190, 2009.
- [18]E. Martínez y D. Díaz, «Una aproximación psicosocial al estrés escolar.» *Educación y Educadores*, vol. 10, n° 2, 2007.
- [19]A. Sánchez y M. Vaquero, «Burnout, variables fisiológicas y antropométricas: un estudio en el profesorado.» *Revista medicina y seguridad en el trabajo*, vol. 54, n° 210, 2008.
- [20]M. González y R. Landero, «Síntomas psicósomáticos y teoría transaccional del estrés.» *Revista Ansiedad y estrés*, vol. 12, n° 1, pp. 45-61, 2006.
- [21]N. Berrío y R. Mazo, «Estrés Académico.» *Revista Dialnet*, vol. 3, n° 2, 2011.
- [22]C. Maslach y A. Pines, «The burnout syndrome in day care setting.» *Rev. Child care quarterly*, vol. 62, pp. 100-113, 1977.
- [23]J. Kulandai y T. Sree, «Reconocimiento de emociones faciales en videos usando HOG y LBP.» Bangalore, India, 2019.
- [24]F. Lavallo-González, J. Villarreal-Pérez, G. González-González, J. Montes-Villarreal, L. Mancillas-Adame, H. Tamez-Pérez, M. Bautista-Medina y J. Valencia-García, «Validación de la medición de cortisol en saliva de una población de adultos jóvenes.» *Revista de endocrinología y nutrición.*, vol. 19, n° 4, pp. 146-148, 2011.
- [25]M. Naranjo, «Unaq revisión teórica sobre el estrés y algunos aspectos relevantes de este en el ámbito educativo.» *Rev. Educación*, vol. 33, n° 2, pp. 171-190, 2009.
- [26]C. Flórez, «Estrés laboral en empresas de producción.» Universidad de Manizales, Colombia, 2014.
- [27]O. Slipak, «Estrés laboral.» *Revista Psicología y Psiquiatría*, 1991.
- [28]E. Antón, «Estrés laboral y variables biomédicas.» *Revista My Science Work*, 2013.
- [29]M. Sánchez y L. Maldonado, «Estrés en docentes universitarios.» *Revista de Ciencias Sociales*, vol. IX, n° 2, pp. 323-335, 2003.
- [30]R. Karasek, «Job demands, job decision latitude, and mental strain. Implications for job redesign.» *Rev. Administrative Science*, vol. Quaterly, n° 24, pp. 285-308, 1979.
- [31]J. Peiró, «Estrés Laboral, Liderazgo y Salud Organizacional.» *Rev. Papeles del psicólogo*, vol. 29, n° 1, pp. 68-82, 2008.
- [32]J. Peiró y I. Rodríguez, «Estrés laboral, liderazgo y salud organizacional.» *Rev. Papeles del psicólogo*, vol. 29, n° 1, pp. 68-82, 2008.
- [33]J. Peiró, «El estrés laboral: una perspectiva individual y colectiva.» *Rev. Investigación administrativa*, vol. 30, n° 88, pp. 30-40, 2001.
- [34]L. Pratt y J. Barling, «Differentiating between dailyevents, acute and



- chronic stressors. A framework and its implications.» Issues and Development in research, pp. 41-53, 1988.
- [35] CDC, «Centro para el control y la prevención de enfermedades.» 2019. [En línea]. Available: <https://espanol.cdc.gov/coronavirus/2019-ncov/daily-life-coping/managing-stress-anxiety.html>.
- [36] M. Trianes, M. Blanca, F. Fernández, M. Escobar, E. Maldonado y Á. Muñoz, «Evaluación del estrés infantil: Inventario Infantil de Estresores,» *Rev. Psicothema*, vol. 21, n° 4, pp. 598-603, 2009.
- [37] J. Mcade, M. Lumley y R. Casey, «Stress, emotional skill, and Illness in children: The importance of distinguishing between Children's and Parents' Reports of Illness,» *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 42, pp. 405-412, 2001.
- [38] C. Betancourt y I. Yépez, «Diseño e implementación de una plataforma multisensado para el monitoreo de sujetos sanos sometidos a un test de inducción de estrés,» *Repositorio ESPE, Quito-Ecuador*, 2018.
- [39] S. Avalos, M. Muñoz y K. García, «Desarrollo de dispositivo portátil para la detección de señales eléctricas cerebrales y su registro,» *Repositorio Universidad de Puebla, Mexico, Mexico*, 2018.
- [40] J. Trujillano, J. March y A. Sorribas, «aproximación metodológica al uso de redes neuronales artificiales para la predicción de resultados en la medicina.» *Rev. Medicina Clínica*, vol. 122, pp. 59-67, 2004.
- [41] M. Gutiérrez, «Aplicaciones de realidad virtual en psicología clínica,» *Aula médica psiquiatría*, vol. 4, n° 2, pp. 92-126, 2002.
- [42] A. Martínez y J. Goddard, «Definición de una red neuronal para la clasificación por medio de un programa evolutivo,» *Rev Mexicana de ingeniería biomédica.*, vol. XXII, n° 1, pp. 4-11, 2001.
- [43] E. Valderrama y G. Ulloa, «Análisis espectral de parámetros fisiológicos para la detección de emociones,» *Rev. Sistemas y Telemática*, vol. 10, n° 20, pp. 27-49, 2011.
- [44] W. Alfonso, M. Muñoz, E. López y E. Caicedo, «Optimización de funciones inspiradas en el comportamiento de búsqueda de néctar en abejas,» *de Congreso Internacional de Inteligencia Computacional.*, Colombia, 2007.
- [45] E. Aguirre-Slancas, «La memoria y el aprendizaje y su relación con la masticación,» *Rev. Mex Neuroci*, vol. XXII, n° 42, pp. 131-138, 2014.
- [46] A. Sandoya, «La incertidumbre en la medicina,» *Rev. Tendencias en Medicina*, vol. XXII, n° 42, pp. 131-138, 2014.
- [47] F. Diez, J. Gámez y M. Puerta, «Aplicaciones de los modelos gráficos probabilísticos en medicina,» *Sistemas expertos probabilísticos*, pp. 239-263, 1998.
- [48] G. Giraldo, «Método del Enjambre de Partículas y Evolución Diferencial



Para el Diseño de Peso Mínimo de un Panel Naval Reforzado.» Universidad Tecnológica de Bolívar, Carategna de Indias, 2011.

[49]J. Rubio, E. Guerrero y F. Castro, «Fuentes de estrés, síndrome de Burnout y actitudes disfuncionales de orientadores de institutos de enseñanza secundaria.» Universidad de Extremadura, España, 2003.

[50]B. Domínguez y Y. Olvera, «Estados emocionales negativos; dolor crónico y estrés.» Rev. Ciencias, pp. 67-75, 2006.

[51]F. Lavalle-González, J. Villarreal-Pérez, G. González, M. Montes-Villarreal, L. Mancillas-Adame, H. Tamez-Pérez y e. al., «Validación de la medición de cortisol en saliva de una población de adultos jóvenes.» Rev. Endocrinología y Nutrición, vol. 19, n° 4, pp. 146-158, 2011.

[52]E. Martínez, «Estudio de la integración de los medios informáticos en los currículos de educación infantil y primaria: sus implicaciones en la práctica educativa.» Universidad Complutense, Madrid, 1998.

[53]M. Sola, «La formación del profesorado en el contexto del espacio Europeo de educación superior. avances.» Rev. Interuniversitaria de formación de profesorado, vol. 18, n° 3, pp. 91-105, 2004.

[54]M. E. N. E. y V. A. Ñaupas H., Metodología de la investigación cuantitativa-cualitativa y redacción de la tesis.; Bogotá: Kindle Edition, 2014.

[55]Delgado, Métodos y técnicas cualitativas y cuantitativas de las ciencias sociales, 1994.

[56]S. Pita y S. Pértegas, Investigación cuantitativa y cualitativa. Unidad de epistemología, España: Hospital Universitario Juan Canalejo, 2002.

[57]V. Orozco, «El marketing y la ciencia.» Rev. Pensamiento crítico, n° 9, pp. 107-113, 2002.

[58][En línea]. Available: http://www.ujaen.es/investiga/tics_tfg/estu_cua-si.html.

[59]G. Alcántara Moreno, «La definición de salud de la Organización Mundial de la salud y la interdisciplinariedad Sapiens.» Rev. Universitaria de Investigación, vol. 9, n° 1, pp. 93-107, 2008.

[60]O. E. Slipak, «Historia y concepto del estrés.» Rev. Argentina de clínica neuropsiquiatría., vol. 03, pp. 355-360, 1991.

[61]B. De Camargo, «Estrés, Síndrome General de Adaptación o Reacción General de Alarma.» Rev. Médico Científica, vol. 17, n° 2, pp. 123-135, 2004.

[62]E. Comercio, «Estrés laboral en el Ecuador.» EL Comercio, 26 enero 2019.

[63]F. Suárez y L. Rosales, «Evaluación de estrés en la generación de enfermedades laborales.» Revista Espirales, vol. 3, n° 30, pp. 100-111, 2019.

[64]R. Neufeld y R. Paterson, «Advances in the investigation of psychological stress.» Wiley series on health psychology/behavioral medicine., pp.



43-67, 1989.

[65]M. Moscoso, «Estrés, salud y emociones: estudio de la ansiedad, cólera y hostilidad.» Rev. De la facultad de psicología de la Universidad Nacional Mayor de San Marcos., vol. II, n° 2, pp. 47-68, 1998.

[66]C. A. Rojas Sierra, «Procesos Complejos del Estrés: Dinámica no lineal.» Tesis de MAestría. Universidad Nacional de Colombia, Colombia, 2012.

[67]R. Rahe, «Estrésores psicosociales y trastorno de adaptación: la tabla de vida de Van Gogh ilustra el estrés y la enfermedad.» The Journal of Clinical Psychiatry, vol. 51, n° 11, pp. 13-19, 1990.

[68]K. Kroenke, R. Spitzer y J. Williams., «The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms.» Psychosomatic Medicine, vol. vol. 64, pp. 258-266., 2002.

[69]R. Lazarus. y S. Folkman., «Stress, coping and adaptation.» New York, Springer., 1984.

[70]B. Sandín., «El estrés psicosocial.» Madrid: Klinik., 1999.

[71]M. González y R. Landero., «Síntomas psicósomáticos y teoría transaccional del estrés.» vol. 12, n° 1, pp. 45-61, 2006.

[72]P. Costa y R. McCrae., «Neuroticism, somatic complaints, and disease: When are somatic complaints unfounded?» American Psychologist, vol. 40, pp. 19-28, 1987.

[73]C. Rojas., «Procesos Complejos del Estrés: Dinámica no lineal.» Tesis de Maestría. Universidad Nacional de Colombia, Colombia, 2012.

[74]E. Wheeler., «El estrés en estudiantes de EGB.» Trabajo de Tesina. Universidad abierta interamericana., Buenos Aires, Argentina, 2012.

[75]F. Suárez y L. Rosales., «Simulación de estrés en la generación de enfermedades laborales.» Espirales, vol. 2, n° 9, 2019.

[76]R. Rahe, «Life change measurement clarification.» Psychosomatic Medicine, vol. 40, pp. 95-98, 1978.

[77]M. Plus, «Biblioteca nacional de medicina de los EEUU.» Medline Plus, 2020. [En línea]. Available: <https://medlineplus.gov/spanish/ency/article/002269.htm>.

[78]F. Suárez, L. Rosales y J. Sayago., «Artificial neural network for the evaluation of vital signs.» Universidad, Ciencia y Tecnología., vol. 22, n° 89, pp. 103-107, 2018.

[79]G. Velásquez., «Sistema de reconocimiento de voz en Matlab.» Tesis de grado. Universidad de San Carlos de Guatemala., Guatemala, 2008.

[80]D. Giraldo. y O. Quintero., «Análisis de señales de audio utilizando la transformada de Gabor.» 2010. [En línea]. Available: https://repository.ea-fit.edu.co/.../29%20Análisis_de_senales_audio_utilizando_transformada.

[Último acceso: junio 2019].



- [81]J. Masip, E. Garrido y C. Herrero, «La detección de la mentira mediante la medida de la tensión en la voz: una revisión crítica.» *Estudios de psicología*, vol. 25, n° 1, 2004.
- [82]C. Ortego, «Detección de emociones en voz espontánea.» Trabajo de fin de carrera. Universidad Autónoma de Madrid., Madrid., 2009.
- [83]j. Forrester, «La dinámica de sistemas y el aprendizaje del alumno en la educación escolar.» *Rev. Academia*. .
- [84]J. Forrester, «Industrial Dynamics, Cambridge.» Productivity Press., p. 464, 1961.
- [85]J. Forrester, «Counterintuitive Behavior of Social Systems,» *Technology Review*, vol. 73, n° 3, pp. 53-68, 1971.
- [86]C. Merino.Soto y C. G. Ruiz-Del Castillo, «Explorando el vínculo entre la inteligencia emocional y la satisfacción con la vida en adultos peruanos.» *Rev. Ansiedad y Estrés.*, vol. 24, n° 2, pp. 140-143, 2018.
- [87]C. D. Sánchez y M. M. Pérez, «Caracterización de voz empleando análisis tiempo-frecuencia aplicada al reconocimiento de emociones.» Trabajo de grado. Universidad tecnológica de Pereira., Pereira., 2007.
- [88]F. Vargas, «Selección de características en el análisis acústico de voces.» Masters Thesis, Universidad Nacional de Manizales., Colombia, 2003.
- [89]I. Cárdenas., H. Ceballos., L. Shang-Hsueh., W. Pavez. y C. Terrisse., «Estudio acústico de la variación interlocutor en sujetos hablantes nativos del español de Santiago de Chile.» Tesis de Grado. Universidad de Chile., Chile, 2010.
- [90]L. Estrada., A. Torres. y J. Raimon, «Evaluación de la asincronía bilateral y toracoabdominal mediante señales mecanomiográficas.» de A: Congreso Anual de la Sociedad Española de Bioingeniería. “Libro de Actas del CA-SEIB 2014 XXXII Congreso Anual de la , España, 2014.
- [91]E. Roldan., «Calidad y dinámica de la voz en grupos sociales en la ciudad de Valdivia (Chile).» *Estudios Fisiológicos.*, vol. 33, pp. 111-118., 1998.
- [92]F. Núñez, P. Cortéz y C. Suárez., «Índice de incapacidad vocal: predictivos.» *Acta Otorrinolaringol.*, vol. 57, pp. 101-108, 2006.
- [93]J.González., T.Cervera. y J.Miralles., «Análisis acústico de la voz: Fiabilidad de un conjunto de parámetros multidimensionales.» vol. 53, n° 4, pp. 256-268, 2002.
- [94]S. Mental, «Unidad de salud mental,» 2019. [En línea]. Available: https://www.salud.gob.sv/archivos/pdf/seguridad_ocupacional_2016_presentaciones/presentacion09082016/MANEJO-DEL-ESTRES-TALENTO-HUMANO-EN-SALUD.pdf.
- [95]N. Elisei., «Análisis acústico de la voz normal y patológica utilizando dos sistemas diferentes: ANAGRAF Y PRAAT.» *Interdisciplinaria*, vol. 29, n°



2, pp. 339-357, 2012.

[96]C. Ortego, «Detección de emociones en voz espontánea,» Universidad Autónoma de Madrid, Madrid, 2009.

[97]E. Ramón, «Transformaciones basadas en el algoritmo Local Binary Pattern de imagenes capturadas con la Kinect para clasificacion facial,» 2011.

[98]D. Kingma y J. Ba, «Adam: un método para la optimización estocástica,» Cornell University, vol. 9, pp. 1-15, 2017.

[99]IMSS, «Test de Estrés,» [En línea]. Available: <http://www.imss.gob.mx/sites/all/statics/salud/estreslaboral/Test-Estres-Laboral.pdf>.

[100]IMSS, «IMSS,» 2020. [En línea]. Available: <http://www.imss.gob.mx/sites/all/statics/salud/estreslaboral/Test-Estres-Laboral.pdf>.

[101]G. Najah, «Emotion Estimation from Facila Images,» Atilin University, 2017.

[102]P. Michel y R. El Kaliouby, « Facial expression recognition using support vector machines.,» de The 10th International Conference on Human-Computer Interaction, , Crete, Greece, 2005.

[103]T. Gritti y e. al., «Gritti, T., et al. Local features based facial expression recognition with face registration errors. in Automatic Face & Gesture Recognition,» de 8th IEEE International Conference on. 2008. IEEE, 2008.

[104]G. Kharat y S. Dudul, «Human emotion recognition system using optimally designed SVM with different facial feature extraction techniques,» WSEAS Trans. Comput, vol. 7, n° 6, pp. 650-659. , 2008.

[105]Z. Niu y X. Qiu, «Facial expression recognition based on weighted principal component analysis and support vector machines.,» de Conference on Advanced Computer Theory and Engineering , 2010.

[106]A. Dhall y e. al., «Emotion recognition using PHOG and LPQ features. in Automatic Face & Gesture Recognition and Workshops (FG 2011),» de IEEE International Conference on. 2011. IEEE, 2011.

[107]S. Gupta y e. al., «A hybrid method of feature extraction for facial expression recognition. in Signal-Image Technology and Internet-Based Systems (SITIS),» de 2011 Seventh International Conference on. 2011. IEEE, 2011.

[108]J.Chen y e. al., «Facial expression recognition based on facial components detection and hog features,» de International Workshops on Electrical and Computer Engineering Subfields., 2014. .

[109]M. Donia, A. Youssif y A. Hashad, «Spontaneous facial expression recognition based on histogram of oriented gradients descriptor.,» Computer and Information Science, vol. 7, n° 3, p. 31, 2014.

[110]P. Carcagnì y e. al., «Facial expression recognition and histograms of oriented gradients: a comprehensive study.,» SpringerPlus, vol. 4, n° 1, p. 15,



2015.

[111]P. Michel y R. El Kaliuby, «Facial expression recognition using model-based feature extraction and action parameters classification.» Journal of visual communication and image representation., vol. 8, n° 3, pp. 278-290, 1997.

[112]T. Gritti y e. al, «Local features based facial expression recognition with face registration errors in automatic face & gesture recognition.» de IEEE International Conference , 2008.

[113]G. Kharat y S. Dudul, «Human emotion recognition system using optimally designed SVM with different facial feature extraction techniques.» WSEAS Trans. Comput, vol. 7, n° 6, pp. 650-659, 2008.

[114] Z. Niu y X. Qiu, «Facial expression recognition based on weighted principal component analysis and support vector machines.» de 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE). IEEE., 2010.

[115]A. Dhall y e. al., «Emotion recognition using PHOG and LPQ features. in Automatic Face & Gesture Recognition and Workshops,» de 2011 IEEE International Conference on. 2011. IEEE. , 2011.

[116]M. Donia, A. Youssif y A. Hashad, «Spontaneous facial expression recognition based on histogram of oriented gradients descriptor.» Computer and Information Science, vol. 7, n° 3, p. 31, 2014.

[117]P. Carcagni y e. al., «Facial expression recognition and histograms of oriented gradients: a comprehensive study,» SpringerPlus, vol. 4, n° 1, p. 1, 2015.

[118]D. E. Arroyo y M. G. Idoate, «Visualizando neuronas en redes neuronales convolucionales,» ETS de Ingeniería Industrial, Informática y de Telecomunicación, Pamplona, 2019.

[119][wikipedia.org/wiki/Coeficiente_de_correlaci%C3%B3n_de_Pearson#:~:text=De%20manera%20menos%20formal%2C%20podemos,ambas%20sean%20cuantitativas%20y%20continuas.](https://es.wikipedia.org/wiki/Coeficiente_de_correlaci%C3%B3n_de_Pearson#:~:text=De%20manera%20menos%20formal%2C%20podemos,ambas%20sean%20cuantitativas%20y%20continuas.) 2020. [En línea]. Available: https://es.wikipedia.org/wiki/Coeficiente_de_correlaci%C3%B3n_de_Pearson#:~:text=De%20manera%20menos%20formal%2C%20podemos,ambas%20sean%20cuantitativas%20y%20continuas





AutanaBooks
Engineering & Services



ISBN: 978-9942-38-662-5

