



Universidad Autónoma de Querétaro
Facultad de Ingeniería
Maestría en ciencias línea terminal mecatrónica

“Detección de estrés en conductores automovilísticos a través de medidas estadísticas y señales electromiográficas”

Opción de titulación
Tesis

Que como parte de los requisitos para obtener el Grado de
Maestría en mecatrónica

Presenta:
Ing. Olivia Vargas López

Dirigido por:
Dr. Juan Pablo Amézquita Sánchez

Centro Universitario, Querétaro, Qro.
Febrero, 2022
México



Universidad Autónoma de Querétaro
Facultad de Ingeniería
Maestría en ciencias línea terminal mecatrónica

“Detección de estrés en conductores automovilísticos a través de medidas estadísticas y señales electromiográficas”

Opción de titulación
Tesis

Que como parte de los requisitos para obtener el Grado de
Maestría en mecatrónica

Presenta:
Ing. Olivia Vargas López

Dirigido por:
Dr. Juan Pablo Amézquita Sánchez

Dr. Juan Pablo Amézquita Sánchez
Presidente

Dr. Martín Valtierra Rodríguez
Secretario

Dr. Luis Alberto Morales Hernández
Vocal

Dr. Roque Alfredo Osornio Ríos
Suplente

Dr. Jesús Rooney Rivera Guillen
Suplente

Centro Universitario, Querétaro, Qro.
Febrero, 2022
México

RESUMEN

Las consecuencias económicas y personales derivados de los accidentes automovilísticos generan efectos negativos en la sociedad, los cuales han ido en aumento recientemente. Una de las causas que pueden generar estos accidentes de tráfico es el estrés presente en los conductores mientras van manejando sus automóviles, por ello, es fundamental desarrollar técnicas para la detección de niveles estrés en conductores automovilísticos. Las metodologías desarrolladas anteriormente son capaces de distinguir si el conductor presenta estrés o no con una mala relación efectividad y carga computacional, lo cual no permite que pueda ser implementada en tiempo real. En este trabajo, se investiga la eficiencia de indicadores estadísticos tales como moda, mediana, desviación estándar, varianza, entre otros, debido a que estos son capaces de medir cambios sutiles que se dan en una señal fisiológica como las señales electromiográficas. En este sentido, los resultados que se han obtenido con la metodología desarrollada en esta investigación mostraron que la varianza y la desviación estándar en conjunto con un clasificador de tipo máquina de soporte vectorial con un kernel cúbico son los más efectivos para la detección de estrés en conductores con una efectividad del 96%, lo cual muestra una mejora en la eficiencia, además de ser una metodología de baja carga computacional debido a que no tiene una etapa de preprocesamiento de las señal. Es importante mencionar, que las máquinas de soporte vectorial pueden ser entrenados con diferentes kernels, por lo tanto, se entrenó el clasificador con cada uno de ellos para conocer cual daba el mejor resultado utilizando como entrada los indicadores estadísticos.

(Palabras clave: detección de estrés, señales EMG, indicadores estadísticos; máquina de soporte vectorial)

SUMMARY

In recent years, the economic and personal consequences caused by traffic accidents generate negative effects in society have been increased. Level of stress is one of the causes that can generate these traffic accidents drivers, therefore, it is very important the development of techniques to detect stress levels in drivers. Previous methodologies are able to distinguish whether the driver presents stress or not with a poor effectiveness and computational load ratio, which does not allow it to be implemented in real time. In this work, the efficiency of statistical indicators such as mode, median, standard deviation, variance, among others, is investigated because they are able to measure subtle changes that occur in a physiological signal such as electromyographic signals which are used in this work. In this sense, the results obtained with the methodology developed in this thesis work showed that the variance and standard deviation in conjunction with a support vector machine type classifier with a cubic kernel are the most effective for the detection of stress in drivers with an effectiveness of 96%, which shows an improvement in efficiency, besides being a methodology of low computational burden because it does not have a preprocessing stage of the signal. It is important to mention that support vector machines can be trained with different kernels, therefore, the classifier was trained with each one of them to know which one gave the best result using the statistical indicators as input.

(Key words: stress detection; EMG signals; statistical time features; support vector machine)

AGRADECIMIENTOS

A mi mamá, Olivia, porque siempre me ha apoyado en todas las decisiones que he tomado y por darme el apoyo necesario para terminar la maestría.

A mi papá, que ha creído en mi en todo momento y por apoyarme para estudiar el posgrado.

A mis tías, María y Ana, por creer en mí y animarme a dar lo mejor a pesar de todo.

A Carlos, por dame el apoyo y valor para estudiar un posgrado y superarme en cada momento.

A mis asesores, Dr. Amézquita, por creer en mi capacidad y por la valiosa orientación a través de la investigación para culminar este trabajo, al Dr. Valtierra, Dr. Morales, Dr. Osornio y al Dr. Rivera, por apoyarme en el proceso y por sus comentarios que hacen que el trabajo sea mejor y culminar mi posgrado con una mejor visión de la investigación en México.

A la Universidad Autónoma de Querétaro y al Consejo Nacional de Ciencia y Tecnología (CONACyT) por bríndame la oportunidad de estudiar el posgrado mediante las becas otorgadas tanto de la institución como la del consejo.

TABLA DE CONTENIDOS

1.	INTRODUCCIÓN	12
1.1	ANTECEDENTES	13
1.2	DESCRIPCIÓN DEL PROBLEMA	19
1.3	JUSTIFICACIÓN	21
1.4	HIPÓTESIS Y OBJETIVOS	22
1.4.1	<i>Hipótesis</i>	22
1.4.2	<i>Objetivo general</i>	22
1.4.3	<i>Objetivos particulares</i>	22
2.	REVISIÓN DE LITERATURA	24
2.1	SISTEMA MUSCULOESQUELÉTICO	25
2.2	SEÑALES ELECTROMIOGRÁFICAS	27
2.3	SISTEMA DE ADQUISICIÓN DE DATOS	29
2.4	AMPLIFICADORES OPERACIONALES	30
2.5	FILTROS ANALÓGICOS	31
2.5.1	<i>Filtros activos</i>	32
2.5.2	<i>Filtros Notch doble T con amplificadores</i>	33
2.6	PROCESAMIENTO DIGITAL DE SEÑALES	34
2.6.1	<i>MATLAB</i> ®	34
2.6.2	<i>Indicadores estadísticos</i>	35
2.6.3	<i>Normalización</i>	37
2.6.4	<i>Método de Kruskal-Wallis</i>	37
2.6.5	<i>Máquinas de soporte vectorial</i>	38
3.	METODOLOGÍA	40
3.1	METODOLOGÍA PROPUESTA	41
3.2	SISTEMA DE ADQUISICIÓN DE DATOS	42
3.2.1	<i>Sistema de acondicionamiento de la señal</i>	42
3.2.2	<i>Tarjeta de adquisición de datos</i>	45
3.2.3	<i>Software de control para adquisición de datos</i>	46
3.3	MATERIALES.....	46
3.3.1	<i>Base de datos empleada</i>	46
3.3.2	<i>Preparación de las señales EMG de la base de datos</i>	47

4.	RESULTADOS Y DISCUSIÓN.....	49
4.1	PRESENTACIÓN DE RESULTADOS METODOLOGÍA	50
4.2	PRESENTACIÓN DE RESULTADOS SISTEMA DE ADQUISICIÓN DE DATOS	57
4.3	DISCUSIÓN DE RESULTADOS	60
5.	CONCLUSIONES Y PROSPECTIVAS.....	63
5.1	CONCLUSIONES	64
5.2	PROSPECTIVAS	66
6.	REFERENCIAS.....	67
A.1	ARTÍCULO PUBLICADO	80

g

ÍNDICE DE TABLAS

Tabla 1. Amplificadores operacionales (Boylestad, 2009).....	31
Tabla 2. Indicadores estadísticos (Abdullah, 2005; Phinyomark et al., 2012; Choi et al., 2020; Yanez-Borjas et al., 2020)	36
Tabla 3. Tipos de SVM y sus kernels (Noble, 2006).	39
Tabla 4. Valores p obtenidos mediante el método de Kruskal-Wallis.....	52
Tabla 5. Efectividad obtenida para cada kernel con las diferentes ventanas.	56
Tabla 6. Comparación entre la metodología propuesta con trabajos similares.	61

ÍNDICE DE FIGURAS

Figura 2.1 Músculos posteriores del cuerpo humano (Marieb, 2008).....	25
Figura 2.2 Electrodoos superficiales (Piña,2015).....	26
Figura 2.3 Electrodoos invasivos (a) de fibra simple, (b) de aguja concéntrica, (c) electrodo mono polar y (d) macro electrodo (Merletti, 2004).....	26
Figura 2.4 Posicionamiento de los electrodoos para medir el músculo trapecio (Autoría propia).....	27
Figura 2.1 Amplificador operacional básico (Boylestad y Nashelsky, 2009).....	28
Figura 2.2 Configuraciones del filtro (a) paso bajas, (b) paso altas y (c) paso banda, donde se observan las frecuencias de corte (Autoría propia).....	30
Figura 2.3 Configuración del filtro Sallen-Key (a) paso bajas y (b) paso altas ambos de segundo orden (Autoría propia).....	31
Figura 2.4 Configuración simple del filtro notch doble T (lu et al., 2011).....	31
Figura 2.5 Ventana principal de MATLAB® (Autoría propia).....	33
Figura 3.6 Metodología propuesta.....	39
Figura 3.2 Vista superior de los 8 pines del amplificador LT1920.....	41
Figura 3.3 Esquemático filtro notch de 60 Hz.....	41
Figura 3.4 Esquemático filtro sumador no inversor.....	42
Figura 3.5 Esquemático filtro Antialiasing.....	42
Figura 3.6 Tarjeta PCB del electromiograma.....	43
Figura 3.7 Tarjeta ARDUINO MEGA (ARDUINO, 2021).....	44
Figura 3.8 Descripción de la ruta de manejo.....	45

Figura 3.9 Localización de los sensores EMG.....	46
Figura 3.10 Segmentos extraídos de cada periodo de la señal.....	46
Figura 4.1 Diagramas de caja de los 17 indicadores estadísticos de los tres periodos de los segmentos sin normalizar.....	51
Figura 4.2 Diagramas de caja de los 17 indicadores estadísticos de los tres periodos de los segmentos normalizados.....	52
Figura 4.3 Diagramas de caja de los 17 indicadores estadísticos de ambas condiciones (con y sin estrés).....	53
Figura 4.4 Electrodo utilizado para las pruebas del EMG.....	56
Figura 7.5 Ubicación de electrodo en el paciente para las pruebas del EMG.....	56
Figura 8 Electromiógrafo conectado para la realización de las pruebas.....	57
Figura 9 Señales obtenidas mediante el EMG al realizar contracciones del músculo del paciente.....	58

1. INTRODUCCIÓN

Este capítulo se enfoca en la revisión de otras metodologías propuestas en la literatura para la detección de estrés en automovilistas, así como la justificación y descripción del problema que se abarcara en este trabajo de tesis. Además, se describe el problema que se pretende atacar y la justificación para desarrollar el proyecto, así como la hipótesis que se plantea y los objetivos a cumplir durante el desarrollo de este trabajo de tesis.

1.1 Antecedentes

El estrés es la reacción del cuerpo humano ante cualquier tipo de amenaza, liberando hormonas como la adrenalina y el cortisol. Esto induce al cuerpo a entrar en un estado de alerta y emergencia, lo que provoca un aumento en el ritmo cardíaco, tensión muscular, incremento en la presión sanguínea, aceleración en la respiración e incremento en la agudez de los sentidos (Alić, 2016). Debido a estos cambios es posible utilizar parámetros fisiológicos para detectar el estrés mediante la respuesta galvánica de la piel, frecuencia cardíaca o respiratoria, actividad muscular, entre otros (Rigas et al., 2012). El estrés es un problema de suma importancia debido a que afecta a la población en general en su vida diaria; además, juega un papel crítico en la adaptación de las personas a conducir vehículos ya la hora de realizar decisiones de maniobras propias de la conducción (Useche et al., 2017). Los conductores tienen consecuencias drásticas que afectan a su comportamiento, siendo el principal causante de accidentes de tráfico (Ghandour et al., 2020), dando como resultado muertes o lesiones graves y produciendo efectos negativos como la disminución de la capacidad de tomar decisiones ante cualquier evento (Healey y Picard, 2005).

La Organización Mundial de la Salud en 2018 alrededor de 1.3 millones de personas a nivel mundial fallecieron debido a eventos relacionados con accidentes automovilísticos; también, se estima que 50 millones de personas han sufrido lesiones derivadas de estos accidentes, siendo la principal causa de decesos en personas de entre 5 y 29 años, además es la octava causa de muerte en todas las edades, lo cual causa estragos en el ámbito social y económico (Useche et al., 2017). Por otro lado, la exposición continua al estrés en los conductores puede

incrementar problemas de salud como envejecimiento acelerado (Maurya et al., 2016), depresión y/o ansiedad (Smoller, 2016), diabetes tipo II (Hackett, 2017), asma (Marshall, 2019) y problemas cardiovasculares (Esler, 2017). Hinojosa et al. (2012) menciona que México tiene la segunda tasa de mortalidad por accidentes de tráfico a nivel Latinoamérica, siendo la principal causante de muerte en jóvenes de entre 19 y 29 años. Debido a estas razones, es de suma importancia el generar métodos capaces de identificar características relevantes en señales fisiológicas provenientes del ser humano para detectar estrés de manera certera con el fin de disminuir accidentes de tráfico y, a su vez, las muertes que éstos ocasionan. Este estudio se enfocará en el análisis de señales fisiológicas provenientes de conductores.

Teniendo en cuenta la importancia descrita anteriormente sobre la necesidad de generar metodologías de detección y cuantificación del estrés en conductores, Healey y Picard (2005) propusieron el desarrollo de un experimento que permitiera adquirir las señales fisiológicas de diversos conductores en situaciones reales de manejo. Con esta finalidad, una ruta fue diseñada, la cual incluyó situaciones que implicaron 3 niveles de estrés: bajo, medio y alto. Cada prueba contiene un marcador que permite identificar el nivel de estrés que tiene el conductor, mismo que fue determinado por un experto que acompañaba a cada conductor para realizar la evaluación psicológica necesaria. La información fisiológica recabada por este trabajo incluye la señal cardiaca (ECG), señales de actividad muscular (EMG), la conductividad de la piel o actividad electrodérmica (EDA) y la señal de respiración mediante la expansión de la cavidad del pecho. La información recabada se compartió por medio de una base datos de acceso libre titulada "MIT Stress Recognition in Automobile Drivers Database", donde cabe mencionar que cada prueba tiene una duración que va desde los 50 hasta 90 minutos.

Debido a la calidad de la información recabada, dicho banco de datos ha sido utilizada en diversas investigaciones a nivel internacional con el fin de poder distinguir entre una persona (conductor) que experimenta estrés o no, desarrollando diversas metodologías basadas en análisis del dominio del tiempo, tiempo-

frecuencia o únicamente frecuencia. Cabe mencionar que se emplean ventanas de análisis de 5 minutos para cualquier señal, ya que las autoras de la base de datos indican que sería un intervalo de tiempo apropiado para tomar medidas de mitigación (Healey y Picard, 2005). Munla et al. (2015) combinaron la transformada wavelet e indicadores estadísticos (máximo, mínimo, media, por mencionar los más importantes), con una máquina de soporte vectorial (MSV) para distinguir entre persona que sufre estrés o no empleando las señales ECG. Los autores reportan que su metodología es capaz de distinguir una persona que sufre estrés con una eficiencia del 83.33%.

Keshan et al. (2015) integraron indicadores estadísticos (por ejemplo, promedio de los latidos de corazón, así como la duración de estos, entre otros indicadores) con una MSV para distinguir entre un conductor que experimenta estrés y uno que no usando las señales ECG. Los resultados obtenidos con la propuesta hecha por los autores indica que una eficiencia del 88.24% es obtenida. Por otra parte, Lee et al. (2017) examinaron diversos indicadores en el dominio del tiempo (por ejemplo, promedio, varianza y desviación estándar) y frecuencia (por ejemplo, la densidad del espectro de potencia) en conjunto con una MSV para determinar si el conductor presentaba estrés o no a través de señales de ritmo cardíaco. Los autores reportan que su método es capaz de distinguir entre ambas condiciones con una eficiencia del 95%. Chen et al. (2017) utilizaron la transformada de Wavelet y la transformada Fourier con una MSV para diferenciar entre personas con y sin estrés a través de señales ECG, de respiración y de RGP, obteniendo una eficiencia del 89%. Liu y Du (2018) integraron indicadores estadísticos (por ejemplo, promedio, mediana, desviación estándar, razón mínima y máxima, entre otros) con el análisis discriminante lineal para distinguir entre ambos grupos (conductores con y sin estrés) usando señales electro-dérmicas alcanzando una eficiencia del 81.82%. Zontone et al. (2019) combinaron indicadores estadísticos (por ejemplo., varianza, energía, valor absoluto, desviación estándar, entre otros) y redes neuronales para detectar si un conductor presentaba o no estrés mediante el análisis de señales ECG y EDA, obteniendo una eficiencia de 77.59%. Recientemente, Chui et al. (2020) investigaron el método de Tompkins en conjunto con una MSV y un algoritmo genético para la detección de estrés mediante señales

ECG, logrando una eficiencia del 96.9%. Los autores mencionan que el uso del método propuesto presenta una carga computacional muy alta, la cual limitaría su aplicación en tiempo real debido a los grandes retrasos en los diagnósticos. En ese mismo año, Cardone et al. (2020) desarrollaron un sistema no invasivo basado en imágenes térmicas para detectar estrés en conductores, comparando con un índice de estrés extraído de un ECG, para el análisis utilizaron la regresión del vector de apoyo no lineal (SVR) para estimar el índice de estrés de regiones específicas de la cara. La curva característica operativa del receptor y su área bajo la curva, fueron utilizadas para la clasificación, obteniendo un 80% de detección en personas con estrés.

A pesar de los buenos resultados obtenidos en los trabajos previamente discutidos, éstos contemplan el uso de diversas señales fisiológicas, las cuales pueden presentar desventajas o problemas durante la evaluación de un conductor si presenta o no estrés. Por ejemplo, las señales ECG se pueden ver afectadas por diferentes afecciones cardíacas tales como arritmias, lo que afectaría drásticamente la correcta clasificación (Vargas-López et al., 2019). Las señales EDA se pueden ver afectadas debido a la cantidad de sudor que se produce durante su adquisición ya que las glándulas sudoríparas varían en cantidad dependiendo de la zona de estudio (Sedano et al., 2012). Finalmente, las señales EEG se pueden ver afectadas por el cabello del sujeto de prueba, la piel o la grasa entre la piel y el cerebro (Ortiz y Reinoso, 2010). Por el contrario, las señales EMG pueden ser una alternativa adecuada para la detección del estrés, ya que éstas miden la contracción de los músculos sin que se estas se vean afectadas por otras afecciones como las mencionadas anteriormente; sin embargo, su análisis representa un gran reto debido a las características que presenta (ser no estacionarias, no lineales y estar embebidas en una gran cantidad de ruido), indicando la necesidad de un método/metodología con la capacidad de identificar patrones que permitan ser asociados a los diferentes niveles de estrés, en especial la detección de estrés de nivel bajo, el cual puede generar ligeros cambios en las señales EMG (Wang y Guo, 2020). Debe resaltarse que, dado que el propósito de las metodologías de detección de estrés, particularmente en conductores, es detectar estas condiciones lo más pronto posible, se impone la restricción a los diversos tipos procesamiento de

señales de tener una baja carga computacional sin sacrificar la eficiencia o calidad de las características extraídas. Por lo que en este trabajo de tesis a diferencia de los trabajos mencionados anteriormente se plantea utilizar un procesamiento de una mínima carga computacional y con ello permitir la obtención de resultados en tiempo real, a través del procesamiento avanzado de señales para poder identificar características relevantes en las señales EMG.

Con lo que respecta a la detección de estrés en el ámbito nacional, Dorantes-Argandar et al. (2015) estudiaron si en el tráfico de la ciudad de Cuernavaca existía un vínculo entre el número de accidentes. Esto se logra a través de encuestas y escalas ya establecidas donde los conductores evalúan sus experiencias anteriores al volante como el haber presenciado un accidente automovilístico, el provocar un accidente en los últimos 12 meses, el haber lesionado a una persona durante un accidente vial, entre otros, y con ello hacer una correlación mediante la correlación de Pearson y análisis estadístico sobre la agresividad del conductor y su nivel de estrés presente al conducir. Por otra parte, en la universidad de Colima, Montiel et al. (2016) desarrollaron una interfaz cerebro-computadora para monitorear los cambios de humor en automovilistas a través de señales de electroencefalografía (EEG) obtenida del sensor Emotiv EPOC para detectar si los conductores estaban en riesgo con su entorno al momento de conducir, aunado a eso utilizaron un sistema GPS para alertar a conductores cercanos mediante mensajes la situación de riesgo. Las bandas de frecuencia que se analizaron de las señales son la alfa (8-12 Hz) y la beta (12-30 Hz), las cuales pueden ser activadas cuando el conductor presenta estrés, ansiedad o miedo, y con ello detectar un problema emocional en el conductor y alertar a los conductores cercanos para reducir posibles situaciones riesgosas. Es muy importante mencionar que los autores no reportan los resultados de su investigación. Por lo tanto, en base a los trabajos encontrados a nivel nacional con respecto al estrés, es de suma importancia seguir indagando nuevos métodos basados en técnicas de procesamiento de señales adecuadas para determinar características en señales provenientes del cuerpo humano para la detección de estrés en conductores de manera oportuna, por ello en este trabajo de tesis a diferencia de los trabajos previos

se busca identificar característica en señales EMG a través de medidas no lineales tales como indicadores estadísticos en conjunto con máquina de soporte vectorial para la detección de estrés en conductores de manera oportuna.

Finalmente, en la Universidad Autónoma de Querétaro (UAQ) varios trabajos se han enfocado en investigar bioseñales mediante su procesamiento. Por ejemplo, Vega (2016) procesó las imágenes de una cámara termográfica para detectar cáncer de mama. Cruz (2018) desarrolló una investigación basada en la descomposición empírica de modos (DEM), fractales y redes neuronales (RN) para la clasificación automática de la imaginación motora empleando señales EEG. Igualmente, Vargas (2019) presentó una metodología basada en la DEM, entropías, fractales y RN para predecir infartos. El mismo año, Piña (2019) integró los fractales y la lógica difusa para predecir un infarto hasta 60 minutos previos de su aparición. Ruiz (2019) fusionó la transformada wavelet e indicadores estadísticos proveniente de los parámetros de Hjorth para clasificar la imaginación motora. Por su parte, Cruz (2019) desarrolló un sensor que identifica la presencia de cansancio presentado por exposición prolongada al estrés en conductores automovilísticos, el cual hace uso de procesamiento de imágenes y un árbol de decisión para alertar sobre la presencia de cansancio. Recientemente, Romo (2020) integró la transformada wavelet con indicadores estadísticos y RN para la clasificación de diversos movimientos de un miembro superior del cuerpo humano empleado señales EMG. De acuerdo con los trabajos desarrollados en la UAQ, se puede observar que se han planteado diversos trabajos enfocados al diagnóstico de diversas enfermedades, así como la clasificación de señales del músculo y del cerebro para el movimiento de un brazo robótico y la imaginación motora, respectivamente; ahora bien, en lo que se refiere a la clasificación de señales EMG para la distinción entre personas (conductores) que experimentan estrés o no hay un antecedente, por lo que es importante el proponer un método/metodología capaz de distinguir entre ambos grupos

Por otro lado, el trabajo más cercano al tema que se investigará en este trabajo de tesis es el desarrollado por Becerra et al. (2012), donde realizaron una investigación para identificar si hay un vínculo entre el estrés de los conductores de

los transportes “Escobedo” con el déficit de calidad del servicio de ésta, para ello se utilizó un examen de estrés (el cual proporciona puntos en quince áreas vinculadas con el estrés y el desarrollo de enfermedades) en 12 operadores con una edad de 18 a 49 años. Con ello pudieron señalar que el estrés se manifiesta en los conductores debido a diferentes efectos fisiológicos, emocionales, cognitivos o motores; entre los resultados más preeminentes fueron las bajas puntuaciones con en la alimentación, falta de ejercitación y descanso, consumo de sustancias para mantenerse alerta durante grandes recorridos. Sin embargo, a pesar de ser un trabajo relacionado con el estrés en conductores, este no se enfoca al procesamiento de señales fisiológicas para la detección automática de estrés en conductores; por lo tanto, este trabajo de tesis sentará las bases dentro de la UAQ con respecto al procesamiento de bioseñales que permita distinguir entre conductores que sufren o no de estrés de manera oportuna.

1.2 Descripción del problema

Las señales fisiológicas son generadas por los biopotenciales del cuerpo humano, las cuales se caracterizan por estar inmersas en ruido y artefactos propios de las ondas magnéticas, fuentes de alimentación de los sistemas de adquisición, así como del contacto de los electrodos de la piel (Treo et al., 2009); además, estas señales se caracterizan por ser no estacionarias, lo que impone un reto adicional a la técnica de procesamiento. Por estas razones, es indispensable una estrategia con procesamiento de señales avanzado para detectar y definir características de carácter importante en las señales fisiológicas que indiquen la existencia de estrés en el sujeto participante.

Los estudios para la detección de estrés en conductores se centran en el análisis de las señales ECG (Munla et al., 2015; Chen et al., 2017; Lee et al., 2017; Wang et al., 2018; Zontone et al., 2019; Chui et al., 2020), señales de la conductividad de la piel (RGP) (Liu et al., 2018; Wang et al., 2019; Zontone et al., 2019), frecuencia de respiración (Chen et al., 2017) y señales EEG (Montiel et al., 2016). Sin embargo, el uso de las señales EEG puede estar corrompida por diferentes enfermedades y el entorno de adquisición como el posicionamiento de

electrodos, el cuero cabelludo, grasa, entre otros. (Ortiz y Reinoso, 2010), lo que limita la puesta en práctica de la detección de estrés; por otra parte, el estudio del ECG se puede ver afectada por diversas enfermedades cardíacas (Vargas-López et al., 2020) y el estudio EDA se ve severamente afectada por el artefacto de movimiento de las manos (Affanni et al., 2018) ya que de otra manera, se degrada la calidad de la señal adquirida; siendo otro punto el que múltiples estudios ya han hecho uso de estas señales, demostrando que para lograr una eficiencia en la detección se deben mezclar varias de estas señales (Healey y Picard, 2005; Zontone et al., 2019). Por otra parte, el medir la actividad eléctrica generada por la tensión de los músculos resulta ser un área de oportunidad, ya que la medición de los músculos no interfiere con las maniobras requeridas para conducir; además, la medición no suele causar incomodidad en el conductor, lo que es un punto que puede potenciar su uso masivo en un futuro. A pesar de estas ventajas, analizar señales EMG es un gran reto debido a las características que presenta (baja amplitud (de 0 a 1,5 mV (Silverthorn, 2007)), no estacionariedad, no linealidad y una gran cantidad de ruido), indicando la necesidad de un metodología, con la capacidad de identificar patrones que permitan ser asociados a los diferentes niveles de estrés, en especial la detección de estrés de nivel bajo, el cual puede generar ligeros cambios en las señales EMG (Wang y Guo, 2020).

En este sentido, esta investigación, presentará un método no invasivo basada en técnicas de procesamiento de baja carga computacional tales como indicadores estadísticos para identificar oportunamente características dentro de las señales EMG y con ello detectar tempranamente conductores con estrés para que las personas tomen sus precauciones para disminuir los efectos negativos que el estrés genera en las personas que lo presentan. Es de suma importancia mencionar que este banco de datos utilizado para validar el electromiógrafo ha sido seleccionada debido a su disponibilidad y su uso en diversos estudios en la actualidad; sin embargo, es deseable desarrollar bases de datos actuales de las diferentes señales fisiológicas y su respuesta a eventos de estrés al conducir, por lo que se generará una base de señales EMG que presenten información acerca de eventos de estrés generados al conducir.

1.3 Justificación

Los conductores automovilísticos se pueden ver afectados por el estrés generado al estar al volante, lo que limita su capacidad para tomar decisiones y la concentración requerida para manejar (Healey y Picard, 2005; Useche et al., 2017), lo que puede desembocar en accidentes automovilísticos, generando que las personas involucradas en estos hechos se vean afectadas por lesiones o incluso la muerte. En la última década diversos estudios plantean que a nivel mundial las personas de 5 a 29 años se ven fuertemente afectadas por accidentes automovilísticos, afectando de peor manera a Latinoamérica (Hinojosa et al., 2012; OMS, 2018). Por esta razón, es importante desarrollar metodologías que permitan identificar y clasificar el estrés presente en conductores de manera temprana para detectarlo y alertar al conductor para que éste tome precauciones y con esto aminorar los accidentes de tráfico y, por ende, disminuir las lesiones y muertes que estos ocasionan.

Recientemente, se han realizado diferentes métodos basados en el análisis de señales fisiológicas (tales como señales de electrocardiografía, encefalografía, electromiografía, frecuencia respiratoria, de actividad electrodérmica, por mencionar algunas) para la detección de estrés en conductores en diferentes entornos, demostrando que estas señales pueden ser eficientes para la detección de estrés. Sin embargo, las señales y su procesamiento utilizados en estos trabajos resultan ser de gran carga computacional, lo cual hace que el procesamiento de estas señales no pueda ser implementado en un entorno donde se debe procesar de manera instantánea (Chui et al., 2020); adicionalmente, estos trabajos hacen uso de señales del corazón, la actividad eléctrica de la piel o las señales provenientes de la actividad cerebral, las cuales pueden ser afectadas por otros fenómenos tales como enfermedades cardíacas (Vargas-López et al., 2020), piel, cuero cabelludo, grasa (Ortiz y Reinoso, 2010), pero no empleando las señales de electromiografía. Esto constituye un área de oportunidad para detectar el estrés usando estas señales, puesto que su medición es más simple, no se ven afectadas por otras enfermedades y es menos invasiva que las señales anteriormente mencionadas (Day, 2009).

Por lo tanto, la justificación de esta investigación es el desarrollo de un método basado en el procesamiento avanzado de señales con baja carga computacional que permita realizar un análisis de señales electromiográficas (EMG) con el fin de detectar la presencia de estrés en conductores y, con esto, generar la posibilidad de que el conductor tome sus debidas medidas con el fin de disminuir los accidentes automovilísticos los cuales generan lesiones en las personas haciendo que se disminuya la calidad de vida tanto de él como de las personas involucradas.

1.4 Hipótesis y objetivos

1.4.1 Hipótesis

El análisis de señales electromiográficas basado en medidas no lineales tales como desviación estándar, varianza, kurtosis y otros parámetros estadísticos permitirá la detección de estrés en conductores automovilísticos, mejorando la precisión y carga computacional de los resultados de otros trabajos reportados actualmente.

1.4.2 Objetivo general

Proponer y desarrollar una metodología de baja carga computacional basada en indicadores estadísticos y máquina de soporte vectorial para detectar a través de señales electromiográficas aquellos conductores automovilísticos que sufren de estrés.

1.4.3 Objetivos particulares

1. Desarrollar un sistema de adquisición para la medición de señales electromiográficas en conductores automovilísticos.

2. Programar los diversos algoritmos de mediciones estadísticas tales como varianza, desviación estándar, kurtosis, y momentos, entre otros. haciendo uso de MATLAB para estimar las características en señales en el dominio del tiempo.

3. Descargar, entender y analizar las señales electromiográficas proporcionadas por Healey y Picard (2005) (base de datos de acceso libre, disponible en <https://physionet.org/content/drivedb/1.0.0/>) de personas que presentan estrés mediante los algoritmos de medidas estadísticas programados en el objetivo 2 con la finalidad de poder identificar características en las señales EMG que permitan identificar personas con estrés.

4. Realizar un análisis estadístico o análisis de varianza de los valores de los indicadores estadísticos estimados en el objetivo 3 con el fin de seleccionar los más aptos/ade cuados en la detección de personas con estrés.

5. Diseñar un clasificador tal como una máquina de soporte vectorial, usando los indicadores estadísticos más discriminantes obtenidos del análisis estadístico o de varianza (objetivo 4), para poder detectar automáticamente a personas que presentan estrés.

6. Evaluar todas las pruebas contenidas en la base datos de manera estadística con el fin de validar la precisión de la metodología propuesta buscando realizar la detección de estrés de manera eficiente.

7. Comparar los resultados obtenidos cualitativamente con otros métodos propuestos en la literatura para validar su eficiencia, manteniendo una carga computacional capaz de ser implementada en hardware.

2. REVISIÓN DE LITERATURA

En esta capítulo se presenta la consulta bibliográfica y se muestra la teoría que apoya cada paso de la metodología propuesta. Además, se abordan las señales mioeléctricas y su obtención, así como de las técnicas computacionales para procesar dichas señales, el clasificador empleado para automatizar la decisión final, y por último se realiza una breve presentación del software utilizado MATLAB®.

2.1 Sistema musculoesquelético

El cuerpo humano está conformado por diferentes sistemas tales como: respiratorio, linfático, esquelético y muscular, siendo los dos últimos los que dan funcionalidad y soporte al cuerpo. En particular, el sistema muscular responde a estímulos externos, tales como el estrés, modificando su biopotencial lo cual produce un cambio en la señal que se adquiere en los músculos (Marieb, 2008) y está conformado por aproximadamente 650 músculos dentro del cuerpo humano, y contribuye a la homeostasis (capacidad del cuerpo humano para autorregularse y cumplir con sus funciones) de este, permitiéndole su estabilización, regulación del volumen de los órganos, movimiento, el traslado de las sustancias dentro del cuerpo como sangre y orina, entre otras funciones (Tortora y Derrickson, 2009). Los músculos constituyen un 40% de la masa corporal y se dividen en tres grupos: (1) músculos cardíacos encargados de las contracciones del corazón para activar el flujo de la sangre, (2) músculos lisos o viscerales que se encuentra en las paredes de los órganos huecos, excepto en el corazón, y (3) músculos esqueléticos, como el músculo trapecio, tríceps del brazo, esternocleidomastoideo, entre otros., los que se observan en la Figura 1, cuya función es mantener erguido al cuerpo humano, caminar, saltar, nadar, entre otros y se encuentran pegados a los huesos (Marieb, 2008). Por otra parte, el sistema esquelético está conformado por el esqueleto (206 huesos), articulaciones, cartílagos y ligamentos, lo que da soporte y protección a los órganos internos. El esqueleto se divide en dos partes: el esqueleto axial (huesos del eje longitudinal del cuerpo) y el esqueleto apendicular (huesos de los miembros y la cintura). El sistema musculoesquelético es de gran importancia para la movilidad del cuerpo, ya que los músculos que están unidos a los huesos mediante tendones utilizan los huesos como palanca y con ello producir el movimiento del cuerpo humano; los músculos pertenecientes a este sistema desempeñan tres

papeles importantes en el organismo: (1) mantener las posturas corporales para mantener el cuerpo en una posición erguida al estar de pie o sentados, (2) estabilizar las articulaciones mediante los tendones y producir calor como consecuencia del movimiento y (3) ayudar a mantener la temperatura corporal (Marieb, 2008).

La actividad muscular es causada mediante un proceso de contracción-relajación, este mecanismo se basa en el deslizamiento de filamentos, en el cual el músculo se acorta (contracción) o se relaja (relajación) debido al deslizamiento de sus fibras musculares. La electricidad y su transmisión en las vías nerviosas, así como la estructura de las fibras musculares están relacionadas con el funcionamiento de los músculos (Schmidt-Nielsen, 1983), esto se produce debido a un cambio de potencial de acción producido por las unidades motoras (conformada por una neurona motora y las fibras musculares que inerva) (Tortora y Derrickson, 2008).

En este trabajo de tesis, se toman en cuenta los músculos esqueléticos ubicados en la parte posterior del cuerpo humano a nivel del cuello denotados como los músculos trapecios, se pueden observar en la Figura 2.1, los cuales por su excitabilidad (característica celular también denominada receptividad o irritabilidad) tienen la capacidad de recibir y responder a estímulos relacionados con el estrés (Marieb, 2008; Wijsman et al., 2011). Los trapecios se caracterizan por ser los músculos más superficiales del cuello, los cuales se extienden desde el hueso occipital del cráneo bajando por la columna vertebral hasta el final de las vértebras torácicas.

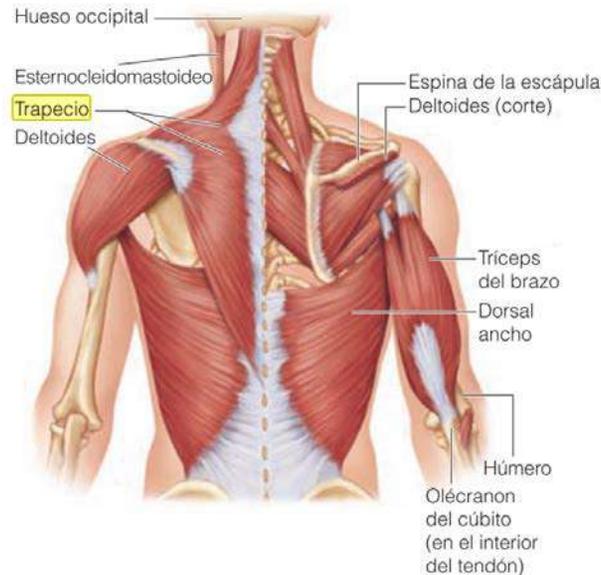


Figura 2.10 Músculos posteriores del cuerpo humano (Marieb, 2008).

2.2 Señales electromiográficas

Las fibras musculares se caracterizan por generar pequeñas corrientes eléctricas debido al intercambio de iones a través de las membranas y se mide mediante una técnica electrofisiológica llamada electromiografía (EMG). Esta técnica ha sido usada ampliamente para estudiar la influencia de diferentes condiciones en la actividad muscular (Nazami, 2019). Para poder obtener dichas señales es necesario aplicar elementos conductores o electrodos en la superficie de la piel, o, de forma invasiva, dentro del músculo. El EMG de superficie (sEMG) es el método más común de medición, ya que los electrodos no son invasivos como se muestra en la Figura 2.2 y puede ser realizado por personal que no sea médico, con un riesgo mínimo para el sujeto (Day, 2009), mientras que los electrodos invasivos ilustrados en la Figura 2.3 se insertan dentro del músculo realizar la diferencia de potencial de forma directa que se da entre la membrana celular y la piel, pueden ser pequeños como los de fibra simple o de aguja concéntrica, o macro electrodos. Los electrodos de mayor superficie tienen una mayor área de captación, provocando que las señales tengan una menor amplitud y la superficie de grabación sea lenta (Khandpur, 2004; Merletti, 2004).



Figura 2.11 Electrodo superficie (Piña, 2015)

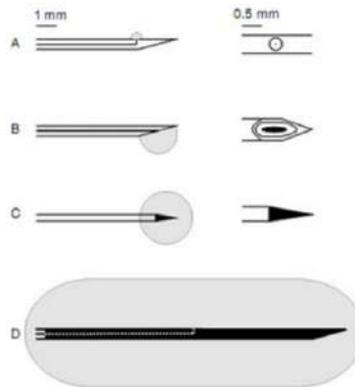


Figura 2.12 Electrodo invasivos (a) de fibra simple, (b) de aguja concéntrica, (c) electrodo mono polar y (d) macro electrodo (Merletti, 2004).

En particular, los electrodos superficiales para la medición muscular se dividen en secos y gelificados. Los secos se caracterizan por estar en contacto directo con la piel; por el contrario, los gelificados emplean un gel el cual crea una mejor química entre la piel y los electrodos; siendo estos los más empleados (80% de aplicaciones del EMG de superficie) debido a que disminuyen el ruido en la señal (Day, 2008). El cloruro de plata (AgCl) es el compuesto más común en la parte metálica de los electrodos, ya que conduce la corriente del músculo a través de la unión del gel y el electrodo (Duchene y Goubel, 1993).

Como se ha mencionado, si bien ya existe una base de datos de referencia, es de interés para este trabajo el sentar las bases para que se continúe en una etapa posterior con la adquisición de una base de datos propia que considere las nuevas tecnologías que los vehículos traen integradas. Se hace hincapié que en este trabajo se empleará la base de datos de referencia; sin embargo, se plantea el diseño de un sistema de medición de las señales EMG. Este sistema se basa en tres electrodos (2 para medir la diferencia de potencial y 1 de referencia). Estos se

posicionarán de acuerdo con la Figura 2.4, dos electrodos (denotados por círculos azules con las letras E1 y E2) para medir la diferencia en el músculo trapecio (entre la línea media formada por el acromion y la columna vertebral) y un electrodo de referencia (círculo amarillo) ubicado en el proceso espinoso de la séptima vértebra (cervical C7) (Wijsman, 2013; Luijcks, 2004).

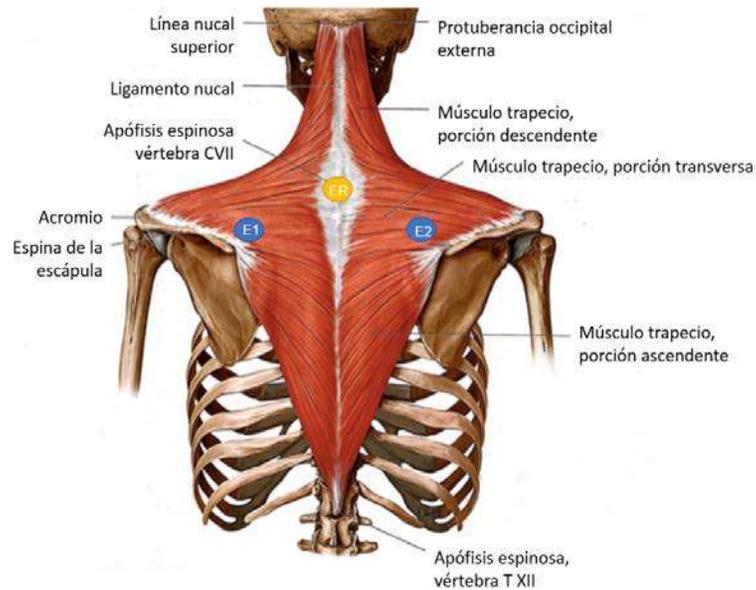


Figura 2.13 Posicionamiento de los electrodos para medir el músculo trapecio (Autoría propia).

2.3 Sistema de adquisición de datos

Un sistema de adquisición de datos (SAD) es un instrumento capaz de medir y recolectar de manera digital datos del compartimento de un fenómeno estudiado. Este se basa en diferentes etapas tales como: acondicionamiento, adquisición y envío de la señal EMG; para el diseño y desarrollo de éstas es necesario el uso de amplificadores operaciones y filtros, para adquirir las señales y extraerles ruido o artefactos que pueden interferir en el diagnóstico final. Dichas etapas serán descritas a detalle en las siguientes subsecciones.

2.4 Amplificadores operacionales

El amplificador operacional (AO) es un circuito integrado con un gran número de transistores (Sedra, 2006), los cuales pueden ser utilizados en diversas configuraciones tales como inversor, no inversor, seguidor de voltaje y el amplificador de instrumentación, entre otras, además de tener la capacidad de realizar operaciones aritméticas tales como suma, resta, integración o derivación. Un AO consta de dos entradas (inversora y no inversora, denotadas por los símbolos + y – en la Figura 2.5) y una salida como se observa en la Figura 2.5. El amplificador operacional ideal tiene como características: (1) impedancia infinita de entrada, (2) impedancia de salida es igual a cero, (3) ganancia de modo común igual a cero, (4) su ancho de banda es infinito (Cifuentes, 2010).

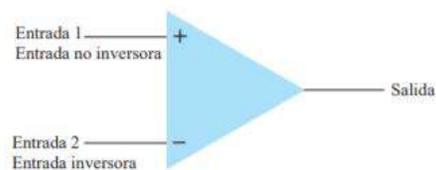
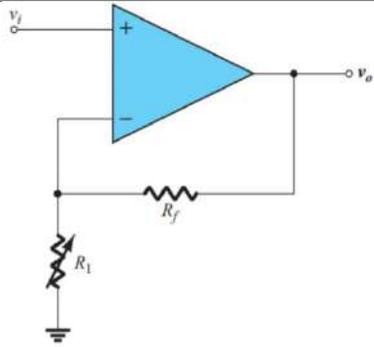
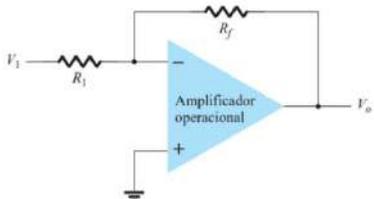
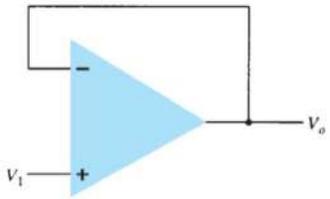
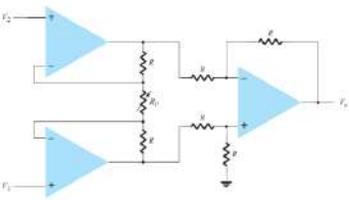


Figura 2.14 Amplificador operacional básico (Boylestad y Nashelsky, 2009).

La Tabla 1 muestra un resumen de las configuraciones más comunes hechas con amplificadores operaciones, las cuales se emplearán para el acondicionamiento de las señales EMG para su posterior adquisición. Además, sus representaciones graficas y ecuaciones en base a sus entradas (V_i, V_1, V_2) y salida (V_0), y elementos intermedios tales como resistencias que intervienen en su configuración (R_f, R_1, R, R_p) son tambien presentados en la tabla, las ecuaciones (1)-(4) se utilizan para obtener los voltajes de salida de los amplificadores respecto a las resistencias y el voltaje de entrada de éste.

Tabla 1. Amplificadores operacionales (Boylestad, 2009)

Amplificador	Configuración	Ecuaciones
NO INVERSOR		$V_0 = \left(1 + \frac{R_f}{R_1}\right)V_i \quad (1)$
INVERSOR		$V_0 = -\frac{R_f}{R_1}(V_i) \quad (2)$
SEGUIDOR DE VOLTAJE		$V_0 = V_i \quad (3)$
DE INSTRUMENTACIÓN		$V_0 = 1 + \frac{2R}{R_p}(V_1 - V_2) \quad (4)$

2.5 Filtros analógicos

Un filtro analógico es un dispositivo empleado para dejar pasar o eliminar ciertas frecuencias o rango de estas mismas (Huircán, 2012). Este tipo de filtro se clasifica en dos tipos: (1) pasivos, contruidos a partir de resistencias, bobinas y condensadores, o (2) activos, los cuales además de utilizar elementos pasivos requieren de amplificadores operacionales (Huircán, 2012). Ambos tipos de configuraciones se clasifican de acuerdo con las frecuencias que deja pasar: filtros

pasa-bajas, pasa-altas, pasa-bandas y rechaza bandas. En particular, los filtros pasa-bajas dejan pasar únicamente las frecuencias menores a una frecuencia de corte (F_c) establecida por el programador, su comportamiento ideal se observa en la Figura 2.6(a), mientras que el filtro pasa-altas dejan pasar las frecuencias mayores a la F_c , como se ilustra en la Figura 2.6(b). Por otra parte, los filtros pasa-banda y rechaza bandas están constituidos por un filtro paso altas y un pasa-bajas, este filtro requiere de dos frecuencias de corte, una mínima (F_{cmin}) y una máxima (F_{cmax}), este filtro se comporta idealmente como se muestra en la Figura 2.6(c). En este trabajo se hará uso de filtros Sallen-Key para filtrar el rango de frecuencia de las señales EMG captado por los electrodos, así como el filtro notch se utilizará para suprimir la frecuencia de 60 Hz correspondiente al ruido de inducción eléctrica.

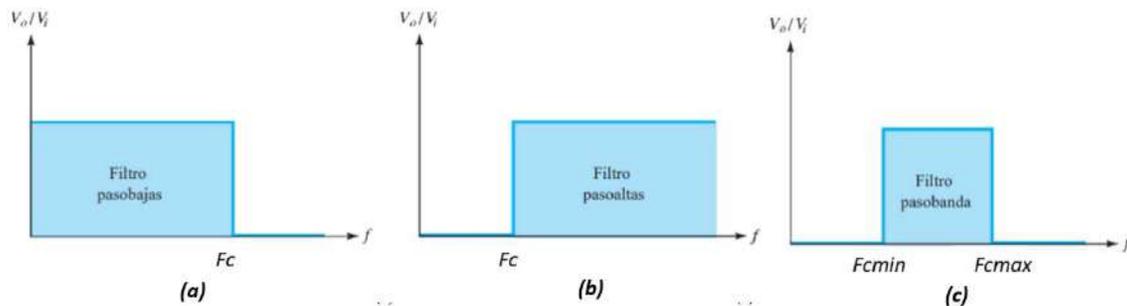


Figura 2.15 Configuraciones del filtro (a) paso bajas, (b) paso altas y (c) paso banda, donde se observan las frecuencias de corte (Autoría propia).

2.5.1 Filtros activos

Dentro de los filtros activos se encuentran los filtros llamados Sallen-Key, los cuales son eficientes y simples para dejar pasar o eliminar cierto rango de frecuencias contenidas en la señal, además, estos permiten aumentar el orden del filtro colocándolos en varias etapas en cascada (Spence, 1988). La configuración de los filtros pasa-bajas y pasa-alta del tipo Sallen-Key se observan en la Figura 2.7(a) y 2.7(b), respectivamente.

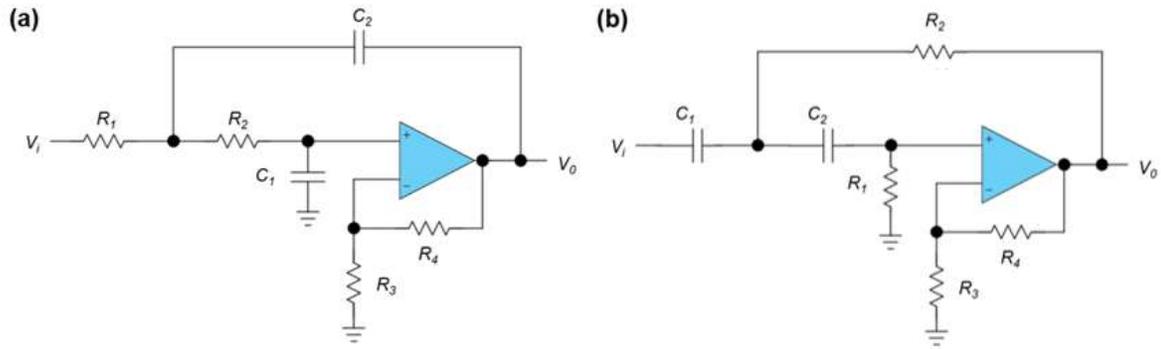


Figura 2.16 Configuración del filtro Sallen-Key (a) paso bajas y (b) paso altas ambos de segundo orden (Autoría propia).

2.5.2 Filtros Notch doble T con amplificadores

Los filtros de red doble T tienen una elevada selectividad lo que permite crear redes supresoras de banda muy estrechas tales como el filtro notch, mostrado en la Figura 2.8, su principal aplicación es la supresión de interferencias como la señal de alimentación, de 60 Hz inducida por la corriente o dispositivos electrónicos cercanos, en el registro de señales biomédicas tales como la electrocardiografía, electroencefalografía, electromiografía, entre otros (Céspedes, 2015).

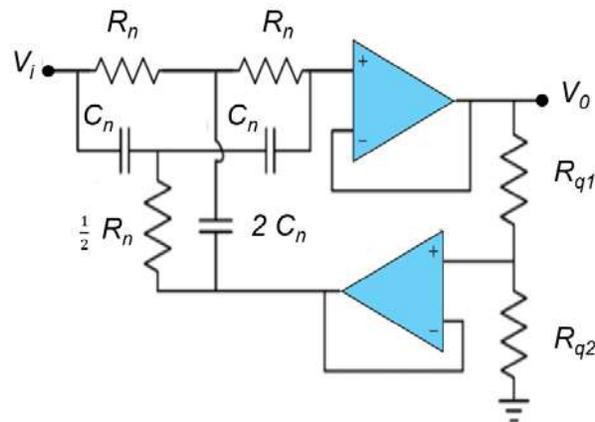


Figura 2.17 Configuración simple del filtro notch doble T (lu et al., 2011).

El comportamiento de este filtro se basa en las ecuaciones (7)-(9), donde se puede observar su función de transferencia, así como la frecuencia de corte de este.

$$H_{nf(s)} = \frac{s^2 + \frac{1}{R_n^2 C_n^2}}{s^2 + \frac{4}{R_n C_n} (1-q)s + \frac{1}{R_n^2 C_n^2}} \quad (7)$$

$$Q = \frac{R_{q2}}{R_{q1} + R_{q2}} \quad (8)$$

$$f_c = \frac{1}{2\pi R_n C_n} \quad (9)$$

2.6 Procesamiento digital de señales

El procesamiento digital de señales (PDS) es un área de la ingeniería electrónica encargada de representar, transformar y manipular la información contenida tanto en señales en tiempo como en imágenes. En base a ello, el procesamiento se encarga del manejo o identificación de las características que contienen las señales analizadas (Vignolo, 2008).

2.6.1 MATLAB®

MATLAB es una herramienta computacional, la cual integra diferentes herramientas enfocadas al procesamiento de señales e imágenes (Báez-López, 2010). Por esta razón, en este trabajo de tesis se empleará este software para procesar las señales EMG, los cuales serán analizados a través de diversos indicadores estadísticos con el fin de identificar cual permite su asociación al fenómeno estudiado.

MATLAB se basa en una ventana mostrada en la Figura 2.9, se divide en diferentes secciones lo que permite interactuar de forma sencilla con el programa, el editor de archivo permite crear la descripción del código que se requiere para el procesamiento de las señales, el espacio de trabajo permite visualizar las variables y su tamaño, mientras que en la ventana de comandos se pueden realizar operaciones, visualizar resultados o variables.

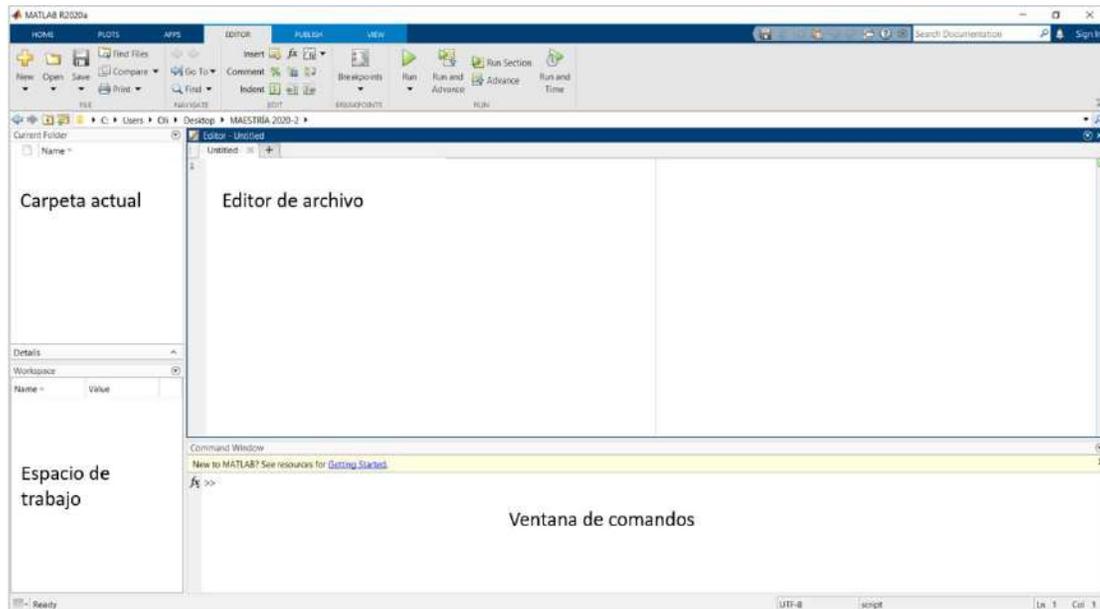


Figura 2.18 Ventana principal de MATLAB ® (Autoría propia)

2.6.2 Indicadores estadísticos

Los indicadores estadísticos son algoritmos con la capacidad de reconocer patrones o características que contienen las señales en tiempo, en especial señales que presentan características no estacionarias, como las señales fisiológicas del cuerpo humano (Xia et al., 2015; Ghorbani et al., 2016; Hassan et al., 2016; Anuragi et al., 2019). A través de los años, estos indicadores han sido empleados en diversas áreas tales como monitorear y diagnosticar fallas (Delgado-Prieto *et al.*, 2013; Igba *et al.*, 2016), identificación de convulsiones epilépticas (Jia et al., 2017), clasificación de enfermedades cardíacas (Afkhami et al., 2016; Slama et al., 2018), entre otras, arrojando buenos resultados.

Por lo tanto, en este trabajo de tesis serán investigados 17 indicadores estadísticos, descritos por las ecuaciones (10)-(26), con el fin de evaluar su viabilidad para la detección de estrés. Los indicadores estadísticos utilizados en el presente trabajo de describen en la Tabla 2, además se muestra una breve descripción del indicador, así como la ecuación para cada indicador. La variable x_i se refiere a la señal de entrada.

Tabla 2. Indicadores estadísticos (Abdullah, 2005; Phinyomark et al., 2012; Choi et al., 2020; Yanez-Borjas et al., 2020)

Indicador estadístico	Descripción	Ecuación	
MODA	Valor que más repetido en un conjunto de datos.	$X_{mo} = mode(x_i)$	(10)
MEDIANA	Valor central de un conjunto de datos.	$X_{me} = magnitud\left(\frac{N+1}{2}\right)$	(11)
VALOR CUADRÁTICO MEDIO (RMS)	Energía de la señal.	$X_{rms} = \left(\frac{1}{N} \sum_{i=1}^N x_i\right)^{\frac{1}{2}}$	(12)
VALOR MEDIO CUADRÁTICO (SMR)	Mide la energía de la señal de forma similar al RMS, pero este es más sensible al tamaño de la señal.	$X_{smr} = \left(\frac{1}{N} \sum_{i=1}^N x_i ^{\frac{1}{2}}\right)^2$	(13)
RANGO	Dispersión de los datos	$X_{rg} = \max(x_i) - \min(x_i)$	(14)
MEDIA	Valor promedio de la señal.	$X_{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$	(15)
VARIANZA	Variabilidad de una serie de datos respecto a su media	$X_{\sigma^2} = \frac{1}{N-1} \sum_{i=1}^N x_i^2$	(16)
DESVIACIÓN ESTÁNDAR	Dispersión de los datos en base su media.	$X_{\sigma} = \sqrt{X_{\sigma^2}}$	(17)
SKEWNESS	Medida de la asimetría de la probabilidad de distribución de los datos.	$X_{sk} = \frac{\sum_{i=1}^N (x_i - x_{\mu})^3}{(N-1)x_{\sigma}^3}$	(18)
KURTOSIS	Indicador del grado de nitidez de la distribución de probabilidad.	$X_{ku} = \frac{\sum_{i=1}^N (x_i - x_{\mu})^4}{(N-1)x_{\sigma}^4}$	(19)
QUINTO MOMENTO	Medida de la elongación de una distribución probabilística del conjunto de datos.	$X_{5^{th}m} = \frac{\sum_{i=1}^N (x_i - x_{\mu})^5}{(N-1)x_{\sigma}^5}$	(20)
SEXTO MOMENTO		$X_{6^{th}m} = \frac{\sum_{i=1}^N (x_i - x_{\mu})^6}{(N-1)x_{\sigma}^6}$	(21)
FACTOR DE FORMA CON RMS		$X_{sfirms} = \frac{X_{rms}}{\frac{1}{N} \sum_{i=1}^N x_i }$	(22)

FACTOR DE FORMA CON SMR	Relación de la magnitud de la forma de onda continua y el promedio de la señal.	$X_{sfsmr} = \frac{X_{smr}}{\frac{1}{N} \sum_{i=1}^N x_i }$	(23)
FACTOR CRESTA	Diferencia entre valores mínimos y máximos relativos al tamaño medio de la señal.	$X_{cf} = \frac{\max(x_i)}{X_{rms}}$	(24)
FACTOR IMPULSO	Indicador de la magnitud del mayor impulso de la señal.	$X_{if} = \frac{\max(x_i)}{\frac{1}{N} \sum_{i=1}^N x_i }$	(25)
FACTOR LATITUDE	Relación del valor máximo de la señal y el valor SMR.	$X_{lf} = \frac{\max(x_i)}{X_{smr}}$	(26)

2.6.3 Normalización

Generalmente, las señales EMG provenientes de diferentes sujetos de estudio pueden presentar una escala o magnitudes diferentes debido a que estos se ven afectados por diversas causas como la configuración y posicionamiento de los electrodos, la impedancia y preparación de la piel, la sudoración y la temperatura del sujeto, así como el número de fibras o el flujo de sangre presentes en el musculo de interés, entre otros factores [44]. Por ello, es necesario normalizar las señales de cada uno de los participantes con el fin de obtener la misma escala o magnitud para cada prueba. En este sentido, las señales EMG y los segmentos obtenidos de estas son normalizadas mediante la resta del promedio de la señal en bruto y dividiéndola entre su desviación estándar, como se muestra en la ecuación (X):

$$x_n = \frac{x - \mu}{\sigma} \quad (27)$$

2.6.4 Método de Kruskal-Wallis

El método de Kruskal-Wallis (KWM, por sus siglas en inglés) es no paramétrico empleado para estimar la independencia estadística en una base de datos con distribuciones no normales (Kruskal y Wallis, 1958), este método ha sido empleado para determinar o estimar epilepsia (Perez-Sanchez et al., 2020), niveles de somnolencia (Bajaj et al., 2020), respiración (Morales et al., 2020), entre otras

aplicaciones. El KWM evalúa la variación entre los conjuntos de datos y las variaciones dentro de cada conjunto de datos (Bashar y Bhuiyan, 2016), calculando un valor de probabilidad denominado valor p . Además, se establece un valor de hipótesis nula (generalmente 0,01), que indica que la mediana de ambos grupos es igual (Kruskal y Wallis, 1952; Hecke, 2012). En este sentido, si el valor p es más pequeño que el valor de la hipótesis nula, se rechaza la hipótesis nula, lo que indica que los conjuntos de datos seleccionados y los restantes pueden diferenciarse. Por otro lado, si el valor p es superior al rango mencionado (Hecke, 2012), la hipótesis nula es aceptada, lo que indica que los conjuntos de datos tienen información similar y, por lo tanto, no pueden utilizarse para diferenciar entre grupos (es decir, estrés y no estrés). Por lo tanto, el valor p se utiliza en este trabajo para determinar qué características son las más discriminantes para reconocer a los conductores de automóviles que presentan estrés.

2.6.5 Máquinas de soporte vectorial

Las máquinas de soporte vectorial (SVM, por sus siglas en inglés *support vector machine*) son un método de clasificación supervisado el cual asigna etiquetas por objeto de cada clase (Noble, 2012) y maximiza una función matemática basada en los datos mediante el uso de un hiperplano y márgenes máximos definidos por los vectores de soporte (los puntos de datos más cercanos al hiperplano), que son las líneas que separan los datos en dos conjuntos "positivos" o "negativos" (Ben-Hur et al., 2008), mediante el uso kernel para proyectar los datos en un espacio de mayor dimensión (Ben-Hur et al., 2008), lo que permite realizar una clasificación eficiente. Los kernels más comunes son el lineal, el gaussiano y la función de base radial (RBF) (Gupta et al., 2017). Dependiendo del kernel empleado, la SVM se puede clasificar en lineal y no lineal, la primera es cuando los datos sólo tienen dos clases y se pueden dividir linealmente (Lucas et al., 2008; Suthaharan, 2016). Por otro lado, si los datos no pueden ser separados linealmente, el dominio de los datos tiene que ser transformado a un espacio de características donde el dominio de los datos pueda ser separado linealmente (Pavlidis et al., 2004). En este sentido, la SVM se puede clasificar como en la Tabla 3, donde P es el número de predictores (Gupta et al., 2017).

Tabla 3. Tipos de SVM y sus kernels (Noble, 2006).

Tipo de SVM	Kernel utilizado
Lineal	Lineal
cuadrático	Cuadrático
Cubico	Cubico
Gaussiano fino	RBF escalado a $\sqrt{P}/4$
Gaussiano mediano	RBF escalado a \sqrt{P}
Gaussiano grueso	RBF escalado a $4\sqrt{P}$

Las SVM han sido ampliamente utilizadas para la clasificación debido a su alta precisión, robustez, la parametrización por la función kernel, y por ser capaces de analizar conjuntos de datos más grandes (Scholkopf y Smola, 2002; Pavlidis et al., 2004; Noble, 2006). Así, en las últimas décadas este algoritmo se ha aplicado en aplicaciones biomédicas como el reconocimiento de estrés (Katsis et al., 2008), la clasificación de imágenes de cáncer de mama (Gupta y Bhavsar, 2017), la clasificación de la señal EMG del codo (Virdi et al., 2016), la clasificación del estadio de artritis de la mano (Akhbardeh et al., 2019), entre otras. Teniendo en cuenta todas las ventajas mencionadas, en este trabajo se investiga su uso, incluyendo todos los kernels mencionados, para diferenciar entre conductores de automóviles con estrés y sin estrés de forma automática.

3. METODOLOGÍA

En el presente capítulo se describe la metodología utilizada para el procesamiento de las señales electromiográficas para la detección de estrés en conductores automovilísticos de forma automática. Así mismo, se describe la base de datos que sirve como validación de dicha metodología y el análisis hecho a las mismas.

3.1 Metodología propuesta

En este estudio se desarrolló la metodología mostrada en la Figura 3.1 y consta de 6 pasos: primero la segmentación de las señales, normalización de los segmentos de las señales, después se realiza la extracción de características y el análisis estadísticos mediante KWM y, por último, la clasificación con SVM para determinar la condición del conductor.

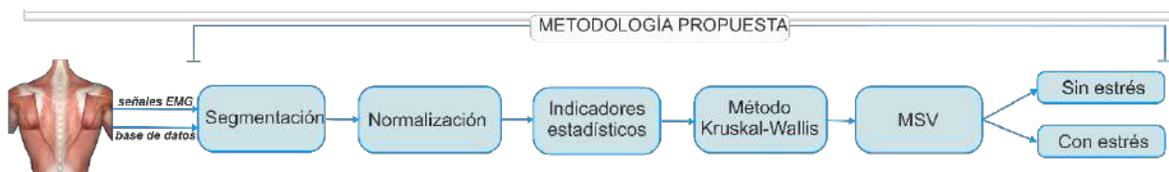


Figura 3.19 Metodología propuesta.

Mediante la base de datos se obtuvieron las señales EMG que se utilizan para la validación de la metodología propuesta en este estudio, únicamente 10 señales fueron seleccionadas debido a que las restantes no cuentan con la información completa de los 3 periodos, descanso, ciudad y carretera. Estas señales fueron tomadas de los sujetos de prueba mediante un electrodo de superficie ubicado en el musculo trapecio derecho a la altura de la vertebra C7 y el acromion. A cada señal EMG se le extraen 5 segmentos de 5 minutos cada uno, un segmento del primer periodo de descanso, dos segmentos del periodo de la ciudad y dos segmentos del periodo de carretera, esto con el objeto tener más información que pueda diferenciar entre los tres niveles de estrés mencionados por Healey y Picard (2005). Una vez obtenidos los 5 segmentos por señal EMG, se procede a realizar una normalización a toda la señal con el objetivo de tener el mismo la misma magnitud en todos los segmentos para después realizar un análisis de

características estadísticas mediante los 17 indicadores estadísticos mencionados en el capítulo anterior con el fin de obtener los mejores indicadores que sean capaces de reconocer estrés en conductores. A continuación, se evaluaron dichos indicadores mediante el método de Kruskal-Wallis para determinar cuál o cuáles eran los mejores indicadores para detectar conductores con estrés al manejar. Finalmente, las características más discriminantes seleccionadas de acuerdo con los resultados arrojados por el método Kruskal-Wallis se utilizan para entrenar al clasificador SVM para determinar la condición del conductor (si el conductor se encuentra estresado o no) de forma automática.

Como se puede observar en la Figura 3.1, las señales EMG proviene del sistema de adquisición de datos a desarrollar como parte tecnológica del trabajo, las cuales serán analizadas con la metodología antes descrita con la finalidad de realizar pruebas con señales actuales adquiridas mediante el sistema descrito a continuación.

3.2 Sistema de adquisición de datos

El sistema de adquisición de datos (SAD) planteado en este trabajo tiene como objetivo la adquisición de señales electromiográficas generadas por el cuerpo humano y consta de tres partes en general.

3.2.1 Sistema de acondicionamiento de la señal

El sistema de acondicionamiento de las señales EMG consta de cuatro etapas, las cuales se describen a continuación:

3.2.1.1 Amplificador de instrumentación

La etapa de la amplificación en la adquisición de las señales bioeléctricas es de suma importancia debido a que tiene una amplitud muy pequeña, por ello se requiere de un amplificador de instrumentación especializado en aplicaciones biomédicas como lo es el LT1920, para lograr una buena visualización en el software descrito en secciones posteriores. Este amplificador de instrumentación consta de 8 pines los cuales se muestran en la Figura 3.2, dos corresponden a las entradas negativa y positiva (-IN, +IN), dos son para la ganancia de referencia (R_G),

dos se utilizan para la alimentación del circuito ($-V_s$, $+V_s$), una es la salida de la señal del LT1920 (OUTPUT) y otra es la terminal de referencia (REF).

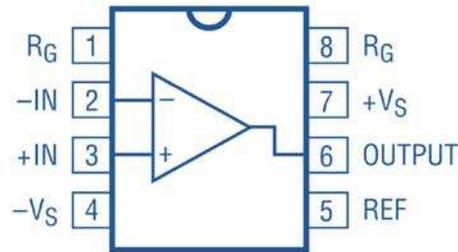


Figura 3.2 Vista superior de los 8 pines del amplificador LT1920.

3.2.1.2 Filtro Notch

La segunda etapa del SAD es un filtro utilizado para suprimir el ruido generado por la red eléctrica de 60 Hz, por lo cual se utiliza un filtro notch mediante un filtro rechaza banda activo doble T de resistencias y capacitores, para los filtros utilizados en este trabajo se utilizó el mismo amplificador operacional OPA335, el cual ha demostrado ser eficiente para instrumentación médica como lo es un electromiógrafo. Como parte del filtro rechaza banda se utilizaron únicamente resistencias y capacitores mostrados en el esquemático del filtro mostrado a continuación.

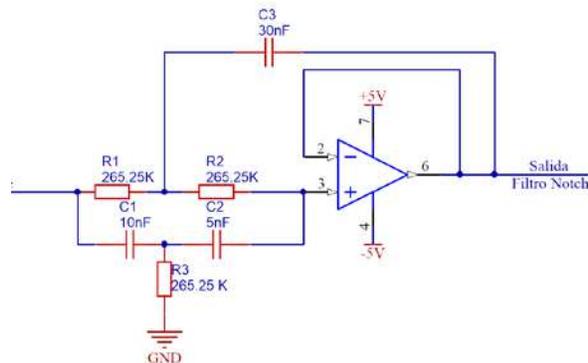


Figura 3.3 Esquemático filtro notch de 60 Hz.

3.2.1.3 Filtro Sumador no inversor

Para rectificar la señal proveniente de los electrodos se utiliza el mismo amplificador operacional que en el filtro notch, con una configuración de un filtro sumador no inversor y un divisor de voltaje.

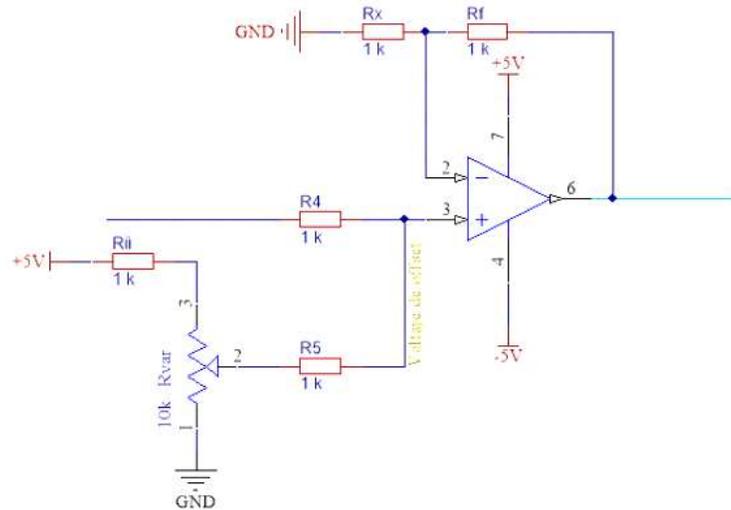


Figura 3.4 Esquemático filtro sumador no inversor.

3.2.1.4 Filtro Antialiasing

La última etapa es un filtro pasa bajas activo de segundo orden con una frecuencia de corte de 500 Hz, se utiliza como un filtro antialiasing. La frecuencia de corte utilizada se designa para asegurar que las frecuencias generadas por los músculos sean captadas, esta frecuencia es de 500 Hz, ya que estas se encuentran en un rango de 6-500 Hz, siendo de mayor interés las frecuencias entre 6-100 Hz (Merletti y Hermens, 2004; Konrad, 2006).

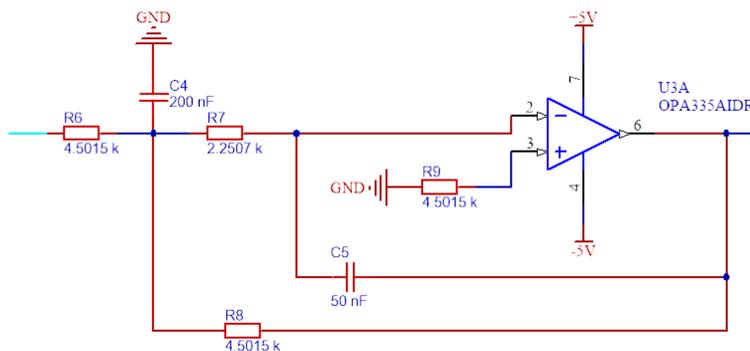


Figura 3.5 Esquemático filtro Antialiasing.

Finalmente, al juntar las etapas anteriormente descritas se obtiene el esquemático de la configuración final para la creación de la tarjeta, como se observa en la Figura 3.6, para la obtención de las señales electromiográficas donde las etapas se pueden observar claramente en el mismo orden del que fueron descritos.

La entrada de los electrodos superficiales será mediante una clema de 3 posiciones, al igual que la salida de la tarjeta, pero de 2 posiciones las cuales servirán para conectar el ARDUINO. La clema restante será la conexión para la fuente de voltaje que se prevé que sea una pila de 9 volts para que el sistema sea portable y no depender de fuentes más sofisticadas.

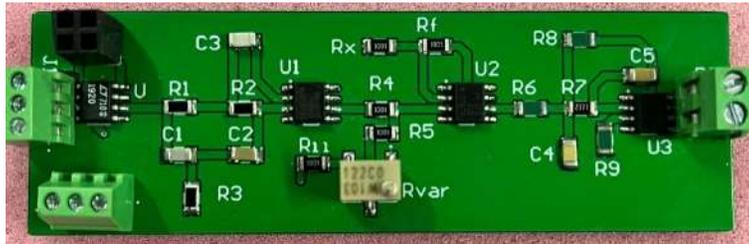


Figura 3.6 Tarjeta PCB del electromiograma.

3.2.2 Tarjeta de adquisición de datos

El sistema para la adquisición de las señales EMG requiere de un microcontrolador para realizar la adquisición y mandar los datos a una interfaz para su posible visualización, en este trabajo se utilizará una tarjeta ARDUINO MEGA® la cual consta de varias características importantes para el SAD a desarrollar entre las que se encuentran:

- Voltaje de operación de 5 volts.
- Voltaje de alimentación entre 7-12 volts.
- Velocidad de reloj de 16 MHz.
- Resolución de 10 bits en cada entrada analógica.

Es importante conocer el voltaje de alimentación y de operación debido a que se pretende desarrollar el sistema de manera portable y alimentarlo mediante pilas lo cual será posible con pilas de 9 volts.

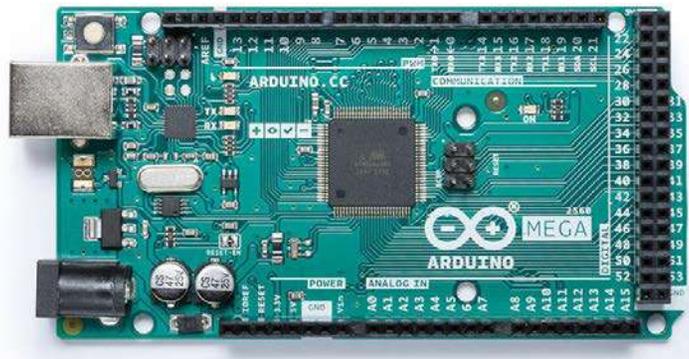


Figura 3.7 Tarjeta ARDUINO MEGA (ARDUINO, 2021).

3.2.3 Software de control para adquisición de datos

Una vez desarrollada la tarjeta de circuito impreso (PCB, por sus siglas en inglés) con el circuito del electromiograma, se procede a desarrollar la comunicación entre la tarjeta de adquisición de datos, la tarjeta Arduino® y el software MATLAB® para la visualización y guardar las señales adquiridas, posteriormente, se aplica la metodología desarrollada con el fin de detectar estrés en los conductores.

3.3 Materiales

En el presente trabajo se hace uso de un banco de datos de señales fisiológicas dentro de las cuales se encuentran las señales electromiográficas que son empleadas para la validación de la metodología propuesta anteriormente descrita. En este capítulo se describen las señales y su adquisición, así como la preparación de estas para poder probar la metodología y validarla.

3.3.1 Base de datos empleada

La base de datos creada por Healey y Picard (2005) utilizada para realizar la validación de la metodología propuesta consta de 17 conjuntos de señales obtenidas mediante cuatro distintos tipos de sensores: electrocardiógrafo, conductividad de la piel, respiración y electromiógrafo, a lo largo de 20 millas en las calles de Boston, la cual es de acceso libre. Es importante mencionar, que únicamente 13 de ellas cuentan con registro de electromiografía, pero 3 de ellas no cuentan con la información completa, por lo tanto, en el desarrollo de este trabajo se trabajará con 10 señales para dicha validación (Healey y Picard, 2005). Las

señales monitorean a los conductores durante una ruta planeada con los eventos de diferentes niveles de estrés para que estos ocurrieran en un momento específico, para ello se incluyeron tres periodos: descanso, carretera y ciudad. En el periodo de descanso se le indicó al conductor sentarse en un garaje con los ojos cerrados y poner el automóvil en neutral. Después de esto, se les indicó a los conductores que manejaran por distintas calles hasta llegar a una calle más concurrida dentro de la ciudad de Boston, con el objetivo de pasar de un nivel de estrés ligero a un estrés alto ya que se planeó que el conductor pasará por numerosas situaciones de alto y de peligro al manejar por estas calles. Finalmente, la ruta desemboca en una carretera donde la conducción no tuviera interrupciones, para después llegar a la misma calle concurrida y volver al punto de salida, y con ello terminar la ruta marcada. La duración de las señales varía entre 50 y 90 minutos, esto dependiendo del tráfico y el conductor que realiza la prueba.

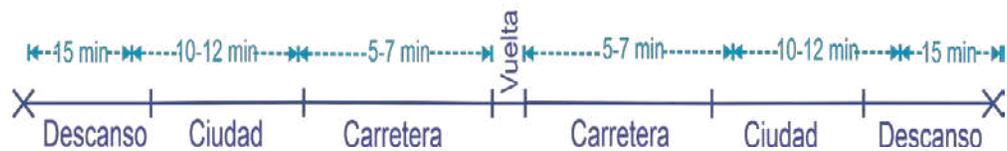


Figura 3.8 Descripción de la ruta de manejo.

3.3.2 Preparación de las señales EMG de la base de datos

Para el análisis de las señales EMG de los participantes, dos sensores fueron utilizados, uno de ellos fue colocado en el músculo trapecio derecho y otro fue utilizado como referencia como se muestra en la Figura 3.9, con una frecuencia de muestreo de 15.5 Hz (Healey y Picard, 2005). En este trabajo, se extrajeron cinco segmentos de las señales EMG: un segmento corresponde al primer periodo de descanso, que se utiliza como línea de base ya que el conductor está relajado y no ha estado expuesto a ningún tipo de estrés al volante y los dos segmentos corresponden a los periodos de ciudad y carretera, respectivamente. Por lo tanto, se extraen 50 segmentos de información (5 segmentos por cada señal EMG) para validar la metodología propuesta. Es importante mencionar que cada segmento o ventana de tiempo extraído contiene 5 minutos de información ya que esta ventana de tiempo ha demostrado ser adecuada para capturar información relevante en señales fisiológicas como las señales EMG (Healey y Picard, 2005; Zhang et al.,

2010; Wang et al., 2013). Por otro lado, en este trabajo, los segmentos obtenidos del periodo de descanso se consideran como condición de no estrés, mientras que los segmentos de ciudad y carretera se emplean como condición de estrés porque pueden comprometer una conducción normal (Cruz et al., 2020). Sólo a modo de ejemplo, la Figura 3.10 muestra el comportamiento de la señal EMG para los tres escenarios, donde se puede observar que no se pueden distinguir diferencias significativas entre los tres escenarios. Por lo tanto, un método con la capacidad de identificar características fiables en estas señales para reconocer a un conductor de automóvil con estrés o no es de importancia imperativa.

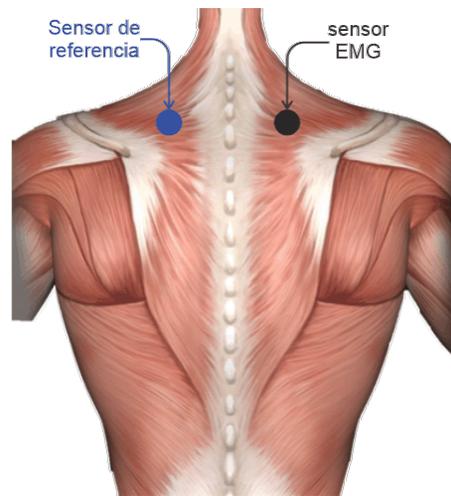


Figura 3.9 Localización de los sensores EMG.

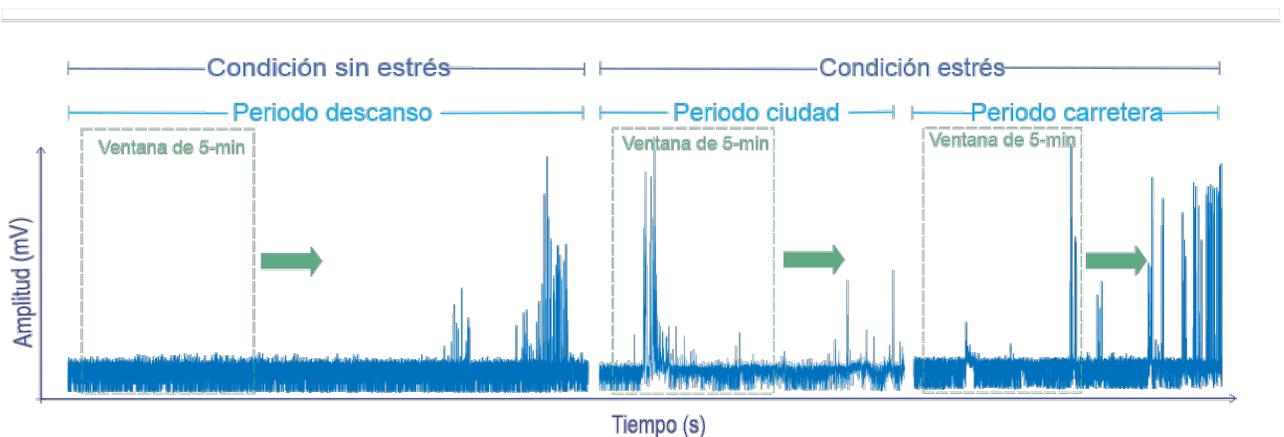


Figura 3.10 Segmentos extraídos de cada periodo de la señal.

4. RESULTADOS Y DISCUSIÓN

En este capítulo se muestran los resultados de la metodología propuesta, las tablas y las figuras creadas a partir de estos, así como la discusión de los resultados.

4.1 Presentación de resultados metodología

En total, se obtuvieron 50 segmentos, con una duración de 5 minutos cada uno, y 850 características fueron calculadas de cada una de las 10 señales disponibles en la base de datos pública anteriormente descrita. Primeramente, se utilizaron los segmentos sin realizar una normalización para realizar el análisis estadístico y determinar si es posible detectar conductores automovilísticos que presentan o no estrés sin la necesidad de un preprocesamiento de la señal, en la Figura 4.1 se observan los resultados obtenidos con el método de Kruskal-Wallis, donde se observa que en ningún indicador es posible diferenciar entre los tres periodos (descanso, ciudad y carretera) ya que los diagramas de caja se traslapan. Por esta razón, se normalizaron todos los segmentos respecto al periodo de descanso, con el fin de tener la misma magnitud, y con ello poder predecir un evento de estrés en conductores. La Figura 4.2 ilustra la distribución de datos estimada por todos los indicadores estadísticos para los tres periodos (descanso, ciudad y carretera). En esta figura se puede observar que la mayoría de los indicadores tienen zonas de solapamiento significativas entre los escenarios de descanso, ciudad y carretera, lo que limita su utilización como entradas para el clasificador; sin embargo, cabe destacar que los gráficos de caja de la varianza y la desviación estándar tienen las zonas de solapamiento más bajas entre todos los indicadores estadísticos. Este resultado indica que ambas características pueden usarse potencialmente para detectar eventos de estrés independientemente de su gravedad, lo cual es una característica muy deseable. Para confirmar este hecho, los valores de los indicadores se unen para generar dos condiciones: sin estrés (descanso) y con estrés (ciudad y carretera); a continuación, se vuelve a aplicar el KWM para representar gráficamente la distribución de los valores en los gráficos de caja que se muestran en la Figura 4.3. Es importante mencionar que este procedimiento se emplea para todos los indicadores para confirmar los resultados que se presentan en la Figura 4.2. A partir de la Figura 4.3, se muestra que, a

excepción de la varianza y la desviación estándar, todos las características estadísticas presentan zonas de solapamiento significativas, lo que confirma los resultados mostrados en la Figura 4.2; sin embargo, los gráficos de caja para la varianza y la desviación estándar no presentan un solapamiento significativo entre ellos. Este resultado permite reafirmar el potencial que tienen como entradas para el clasificador SVM, ya que no presentan zonas de solapamiento significativas. Este resultado es esperado ya que tanto la varianza como la desviación estándar miden la dispersión de las muestras con respecto a su media de forma cuadrada (desviación estándar) o sólo usando la media (varianza); por lo tanto, ambos indicadores complementan la medida de la dispersión de los datos (Thukral et al., 2019). Por estas razones, en este trabajo la varianza y la desviación estándar se definen como las características más significativas para detectar un evento de estrés en los conductores y pueden utilizarse como entradas para el clasificador.

Los resultados mencionados anteriormente se confirman con los valores p representados en la Tabla 4, donde se ve que la Varianza y la Desviación Estándar tienen los valores p más bajos de todos los indicadores, lo que confirma aún más los resultados gráficos presentados en las Figuras 4.2 y 4.3. Por otra parte, si bien se observa que el STF de SMR tiene un valor p bajo, pero presenta un ligero solapamiento entre el escenario de reposo y el de ciudad, como se observa en la Figura 4.2, que es más evidente cuando se consideran los escenarios de estrés y sin estrés (Figura 4.3). Este solapamiento podría generar una disminución del rendimiento de la clasificación, lo cual es un escenario indeseable. En consecuencia, el SMR se descarta como entrada.

Tabla 4. Valores p obtenidos mediante el método de Kruskal-Wallis.

Indicador estadístico	Valor p
Moda	0.0013
Mediana	7.04×10^{-6}
RMS	6.21×10^{-4}
SMR	6.37×10^{-5}
Rango	0.0017
Promedio	2.61×10^{-6}
Varianza	1.11×10^{-6}
Desviación estándar	1.11×10^{-6}
Skewness	0.42
Kurtosis	0.55
5to momento	0.16
6to momento	0.42
Factor de forma con RMS	0.91
Factor de forma con SMR	0.02
Factor cresta	0.23
Factor impulso	0.39
Factor latitud	0.55

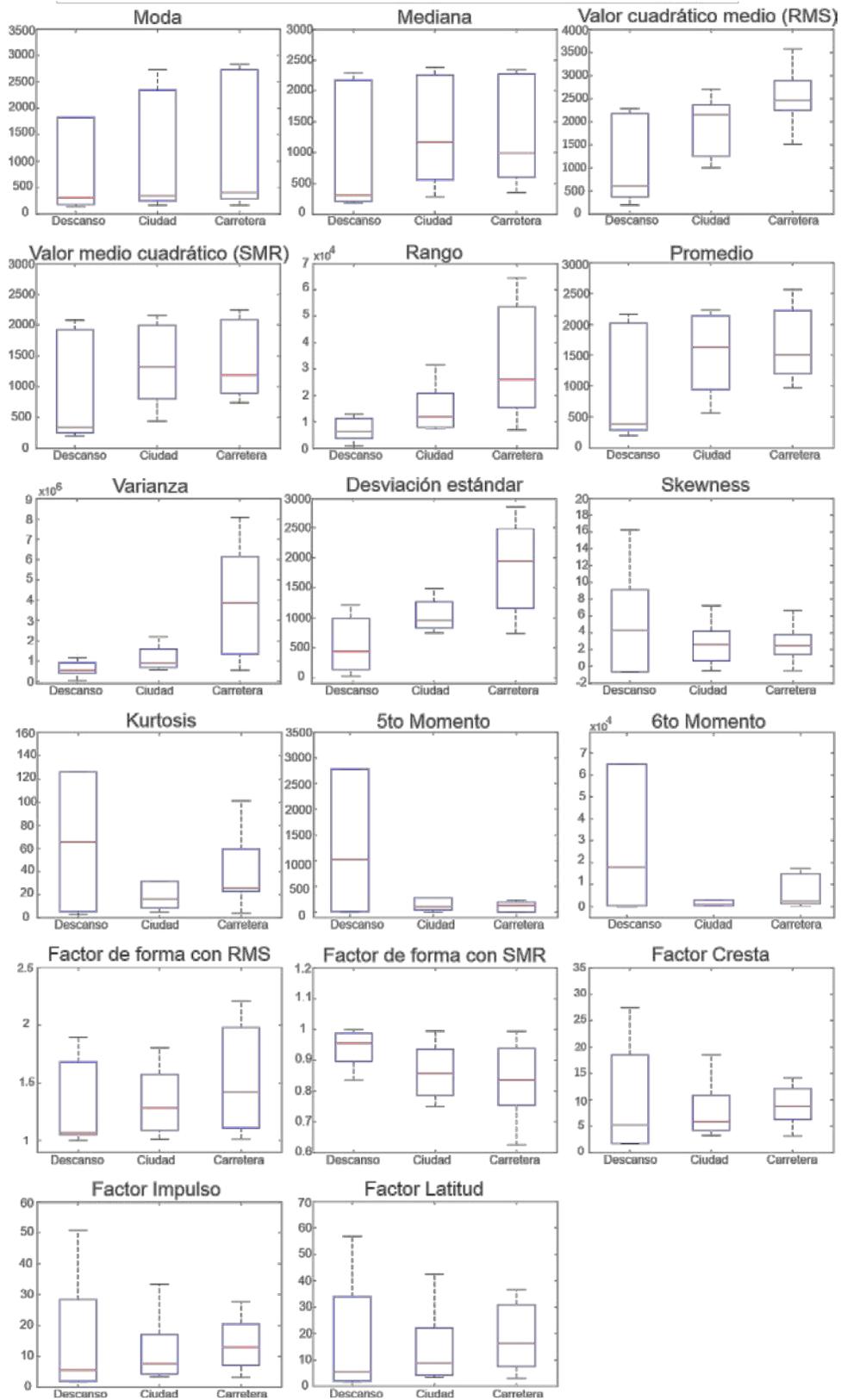


Figura 4.1 Diagramas de caja de los 17 indicadores estadísticos de los tres periodos de los segmentos sin normalizar.

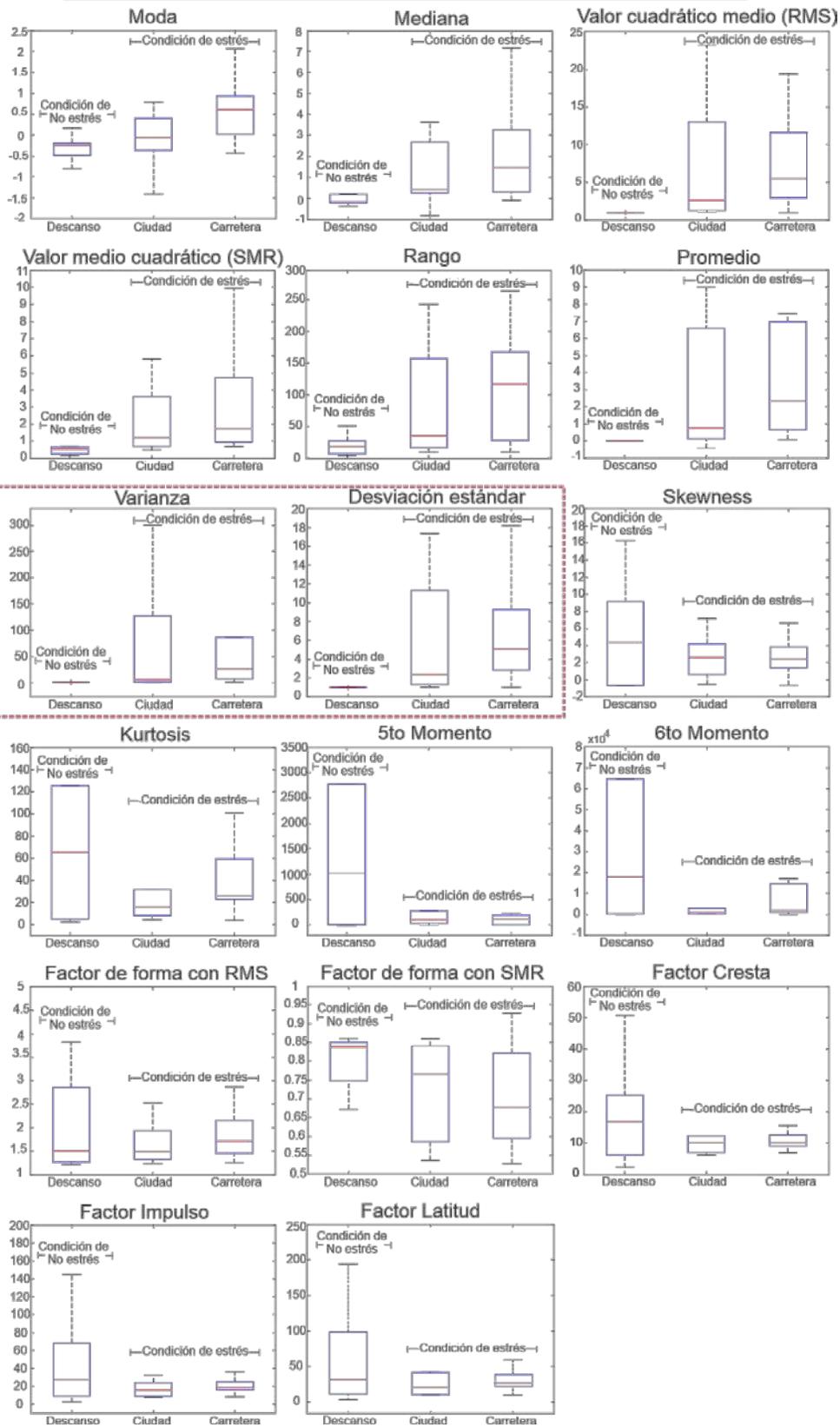


Figura 4.2 Diagramas de caja de los 17 indicadores estadísticos de los tres periodos de los segmentos normalizados.

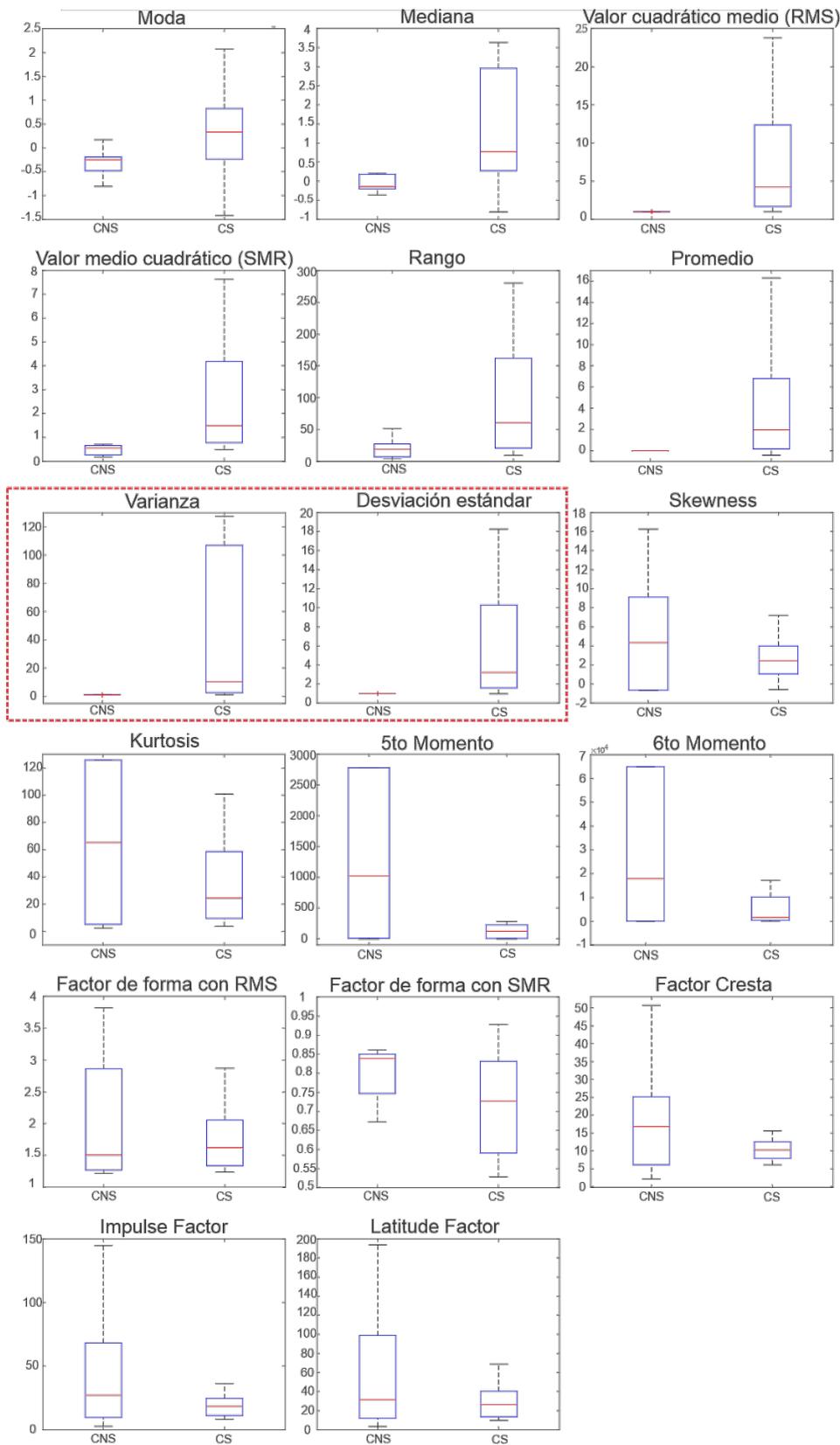


Figura 4.3 Diagramas de caja de los 17 indicadores estadísticos de ambas condiciones (con y sin estrés).

Una vez que se han seleccionado los indicadores más discriminativos, es decir, la varianza y la desviación estándar, de acuerdo con los resultados estimados por KMW y la inspección visual de las Figuras 4.2 y 4.3, se emplean para entrenar un clasificador SVM con el fin de identificar la condición de estrés en los conductores de forma automática. Para ello, se utiliza el algoritmo iterativo de datos individuales (Kecman et al., 2005) como estrategia de entrenamiento para obtener los parámetros del clasificador. Para estimar el rendimiento del clasificador, se utiliza el esquema de validación cruzada de 10- k veces. Este esquema divide aleatoriamente los datos de entrenamiento en k subgrupos con el mismo número de muestras. A continuación, de los k subgrupos, uno de ellos se utiliza como datos de validación (DV) en el modelo, y las restantes $k-1$ submuestras se emplean como datos de entrenamiento. Este procedimiento se repite k veces, ya que permite utilizar cada uno de los k subgrupos una vez como DV. Los k resultados parciales de precisión se promedian para generar la precisión global del clasificador. La ventaja de este método es que todas las muestras se utilizan por igual para el entrenamiento y la validación, pero al mismo tiempo, cada muestra se utiliza sólo una vez para validar el modelo (Rayal et al., 2019). En este trabajo se utiliza un valor de k de 10 ya que proporciona el mejor compromiso para generar el mayor número de subgrupos posibles sin comprometer la diversidad en los subgrupos (Rayal et al., 2019).

En este estudio, se analizaron todos los kernels mostrados en la Tabla 3 para conseguir la mejor precisión posible. En la Tabla 5 se presenta la precisión obtenida con los kernels antes mencionados, concluyendo que el kernel cúbico de la SVM es el que ofrece la mayor precisión, ya que este kernel capta mejor la forma geométrica de la separación entre los escenarios con y sin estrés, generando la mayor precisión posible (Kecman et al., 2005), que es del 96%; además, se obtiene una sensibilidad del 100% y una especificidad del 95%. Los resultados obtenidos permiten afirmar que los indicadores seleccionados pueden ofrecer un clasificador razonable.

Tabla 5. Efectividad obtenida para cada kernel con las diferentes ventanas.

Kernel SVM	Efectividad (%)				
	1-min	2-min	3-min	4-min	5-min
Lineal	66.7	56.7	70	73.3	80
cuadrático	63.3	63.3	70	76.7	78
Cúbico	73.3	66.7	90	90	96
Gaussiano fino	60	56.7	66.7	66.7	80
Gaussiano mediano	66.7	66.7	66.7	66.7	80
Gaussiano grueso	66.7	66.7	66.7	66.7	80

4.2 Presentación de resultados sistema de adquisición de datos

Una vez diseñado e impreso el circuito del EMG, se procede a realizar pruebas para verificar que este funcione correctamente y pueda ser utilizado para la generación de una base de datos en un futuro.

Las pruebas se realizaron en un paciente masculino, al cual se le colocan 3 electrodos de tipo superficial de la marca AMBIDERM® mostrados en la Figura 4.4, marca que es de las más utilizadas comúnmente en aplicaciones biomédicas como ECG o EMG ya que son especiales para utilizar en adultos y cuentan con la parte del cloruro de plata lo cual hace que la adquisición de las señales sea mejor y no se requiera de un gel conductor adicional. Por otra parte, el posicionamiento de los tres electrodos se puede observar en la Figura 4.5, donde el electrodo colocado en la muñeca sirve como electrodo de referencia, mientras que los otros dos electrodos, ubicados en ambos extremos del músculo bíceps, sirven para adquirir la señal eléctrica de dicho músculo, estos electrodos se conectan a la tarjeta con el EMG en la primera etapa que va dirigida al amplificador de instrumentación de uso médico, mostrado en la Figura 4.6(a), dicho amplificador es un LT1920 se requiere de una ganancia externa para fijar la ganancia diferencial del amplificador; en este trabajo se optó por trabajar con una ganancia de 5, para lo cual se requiere de una resistencia de 12.3 Komhs; esto se obtiene de la fórmula descrita en la hoja de datos del LT1920 mostrada en la ecuación 28, obteniendo una resistencia de 12.35 Komhs, para la práctica, la resistencia más cercana es de 12.3 Komhs, por lo tanto, se trabaja con dicha resistencia. Es importante recalcar, que para esta etapa la ganancia mínima es mayor 1, sin embargo, se utilizó una ganancia mayor para evitar posible problemas de adquisición de la señal.

$$R_G = \frac{49.4 K\Omega}{(G - 1)} \quad (28)$$

Una vez conectados los electrodos a la tarjeta, la Figura 4.6(b) muestra la conexión a la fuente de alimentación externa la cual alimenta el circuito con un voltaje de 5 voltios, posteriormente, la Figura 4.6(c) muestra la salida del EMG donde se conecta la punta del osciloscopio para poder visualizar en este la señal adquirida por la tarjeta como lo mostrado en la Figura 4.7(a) y 4.7(b), donde se muestra el osciloscopio funcionando mientras el paciente realizar movimientos de contracción y relajación del músculo antes mencionado. En futuras aplicaciones, la parte de salida, se conectaría a una tarjeta de adquisición de datos para poder visualizar y guardar las señales en una computadora para generar una base de datos con dichas señales. Como se puede observar en las imágenes antes mencionadas se observa que el EMG fue probado y se verificó que el desarrollo tecnológico de este trabajo funciona, es decir, procesa de manera correcta las señales EMG generadas por el músculo bíceps (en este caso).



Figura 4.4 Electrodo utilizado para las pruebas del EMG.



Figura 20.5 Ubicación de electrodos en el paciente para las pruebas del EMG.

Finalmente, como parte de este trabajo de investigación, se planteó únicamente probar y corroborar el buen funcionamiento del EMG desarrollado durante el posgrado, lo cual presenta un antecedente para la realización de base de

datos de señales EMG para la detección de estrés en conductores, abriendo la posibilidad de crear grupos de investigación en el área biomédica dentro la Facultad y de la Universidad, además de la creación de una nueva base de datos realizada con este desarrollo, la cual contemplará nuevos riesgos al conducir un automóvil debido a las nuevas tecnologías que pueden ayudar o empeorar el desempeño del conductor provocándole estrés como el uso de aparatos electrónicos o sistemas integrados de navegación en lo autos, aumento de tráfico, entre otras causas.

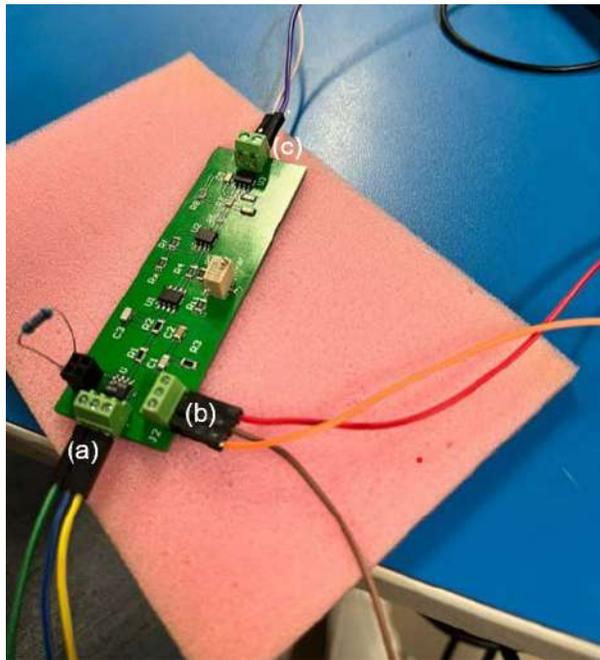


Figura 21 Electromiógrafo conectado para la realización de las pruebas.

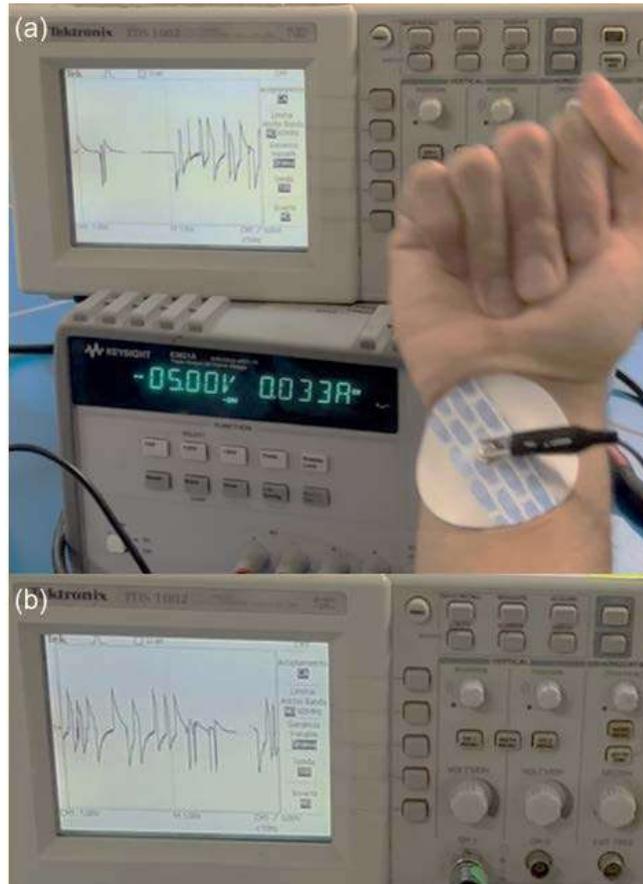


Figura 22 Señales obtenidas mediante el EMG al realizar contracciones del músculo del paciente.

4.3 Discusión de resultados

La Tabla 6 presenta los resultados de la metodología propuesta y otras metodologías que emplean el mismo conjunto de datos y que han utilizado señales EMG como parte de su metodología, donde se hace una descripción breve de los métodos utilizados, el número y tipo de señales fisiológicas, así como la precisión reportada por los autores. De esta tabla se desprende que la propuesta obtiene una precisión del 96% utilizando únicamente señales EMG acopladas a los indicadores, a diferencia de los trabajos anteriores (Katsis et al., 2008; Fu y Wang, 2014) que utilizan diferentes tipos de señales en sus metodologías. Los resultados obtenidos permiten afirmar que la propuesta puede detectar eficazmente eventos de estrés en los conductores independientemente del escenario en el que se encuentren.

Tabla 6. Comparación entre la metodología propuesta con trabajos similares.

Autor	Señales	Metodología	Eficiencia
Katsis et al. (2008) [24]	EMG, ECG, EDA y Respiración	<ul style="list-style-type: none"> • Dos indicadores estadísticos fueron extraídos como el promedio y el RMS. • Validación cruzada de 10 y un clasificador SVM. 	79.3%
Fu and Wang (2014) [23]	EMG, ECG	<ul style="list-style-type: none"> • Utilizaron un paso de preprocesamiento mediante el análisis rápido de componentes independientes a partir de ambas señales. • Se obtuvieron dos medidas no lineales de las ventanas (factor de pico y máximo de la curva de relación cruzada) • Validación cruzada de 10 veces y distancia de Mahalanobis utilizada como clasificador. 	86.7%
Wang and Guo (2020) [15]	EMG	<ul style="list-style-type: none"> • El algoritmo de aprendizaje pseudoinverso basado en el autoencoder (PILAE) se utilizó para el aprendizaje de la representación de las señales y el clasificador AdaBoost fue empleado como paso final. • La validación fue cruzada del tipo "Leave-One-Out". 	58%
Metodología propuesta	EMG	<ul style="list-style-type: none"> • Las propiedades estadísticas se utilizan como características. • La máquina de vectores de apoyo se utiliza como clasificador. • Se emplea una validación cruzada de 10 veces. 	96%

La detección del estrés en los conductores ha demostrado ser una tarea difícil, ya que las señales fisiológicas requieren metodologías robustas que puedan detectar y asociar los cambios a la condición de estrés generada en los conductores. Por estas razones, la mayoría de los autores han propuesto metodologías que utilizan más de una señal fisiológica para determinar si un conductor sufre o no un episodio de estrés (Katsis et al., 2008; Fu y Wang, 2014), logrando precisiones razonables del 79% y el 86,7%, respectivamente. Hay que señalar que estas metodologías utilizan algoritmos que requieren un tiempo importante para ofrecer un resultado, lo que impide la evaluación continua de las ventanas, limitando el funcionamiento en tiempo real, lo que es una característica deseable ya que cuanto antes se detecte el evento de estrés, el conductor puede ser alertado y tomar acciones para prevenir accidentes con las consecuencias que conllevan: los accidentes de tráfico. Por otro lado, cuando sólo se procesa la señal EMG, se emplean métodos complejos y con gran carga computacional para filtrar y clasificar

los cambios detectados (Wang y Kuo, 2020), obteniendo una eficiencia inferior al 60%. Evidentemente, este tipo de metodologías no pueden ser implementadas para estrategias de detección en tiempo real ni pueden detectar los cambios producidos en la señal EMG. Por el contrario, la propuesta obtiene una precisión del 96% utilizando una metodología sencilla pero eficaz que permite identificar cambios en la señal EMG independientemente de la gravedad del estrés a partir de la estimación del STF. La propuesta tiene algunas ventajas como: (1) sólo se emplea una señal fisiológica, (2) no se utiliza ningún paso de preprocesamiento, lo cual es una característica muy conveniente ya que se reduce la carga computacional, (3) la implementación del clasificador SVM puede realizarse de forma eficiente. En consecuencia, se puede lograr un funcionamiento en tiempo real, haciendo que la detección del estrés sea rápida y los conductores puedan tomar sus precauciones, para evitar accidentes de tráfico o reducir los problemas de salud relacionados con el estrés en la población. Sin embargo, es importante (1) crear y evaluar una nueva base de datos de señales EMG con una señal mayor para corroborar esta metodología, y (2) desarrollar nuevas metodologías capaces de detectar entre los tres niveles de estrés mencionados por los autores en Useche et al. para tener diferentes alertas dependiendo del nivel de estrés.

5. CONCLUSIONES Y PROSPECTIVAS

5.1 Conclusiones

El sistema nervioso simpático del cuerpo humano reacciona y genera cambios fisiológicos y químicos ante cualquier tipo de amenaza, lo cual genera un episodio de estrés (Alic et al., 2016). Los conductores automovilísticos tienen episodio de estrés debido a los cambios inesperados mientras manejan lo cual es una de las causas más importantes en accidentes automovilísticos; este trabajo investiga las capacidades que pueden tener los indicadores estadísticos para detectar las pequeñas diferencias en señales EMG que producen los diferentes escenarios de estrés en conductores automovilísticos. El método de Kruskal-Wallis se utilizó como el método de detección de independencia estadística de cada conjunto de valores de los indicadores estadísticos para determinar los que tienen la condición más discriminante para ser utilizados como las entradas de un clasificador que pueden diferenciar si un conductor está estresado o no; en este trabajo, se compararon diferentes algoritmos de aprendizaje automático como SVM y MLP para mejorar la eficiencia de la metodología. Los resultados presentados muestran que la normalización de las señales es necesaria para garantizar que los resultados sean mejores cuando se trata de características con diferentes rangos de variación. Además, la varianza y la desviación estándar son las características estadísticas más discriminantes, ya que no tienen zonas superpuestas entre las condiciones de estrés y de no estrés; esto nos indica que son indicadores eficaces para detectar y cuantificar los cambios sutiles que la condición de estrés generan dentro de las señales EMG, permitiendo el uso de un clasificador, tal como la máquina de soporte vectorial, con una precisión del 96%. En comparación con metodologías anteriores que utilizan una o más señales fisiológicas, la propuesta mejora significativamente los resultados presentados con una alternativa simple pero efectiva ya que no requiere de algoritmos de preprocesamiento ni clasificadores complejos. Estas características hacen que la metodología propuesta sea una opción atractiva para los sistemas que pueden operar en entornos en tiempo real, ya que se pueden generar alertas continuas, lo que puede reducir todas las consecuencias que puede tener un accidente de tráfico. Para confirmar estos resultados, la propuesta debe ser explorada con una base de datos más grande, ya

sea probando más conductores o utilizando algoritmos de aumento de datos, como el marco del autoencoder. Además, las mejoras basadas en técnicas de aprendizaje automático explicables para evaluar el nivel de estrés, ya que esto puede personalizar la severidad de la alerta que puede emitir un sistema, lo que podría reducir aún más las consecuencias personales y económicas que puede tener un accidente de tráfico (Doborjeh et al., 2020; Ieracitano et al., 2021). Además, el desarrollo de una base de datos actual que pueda considerar todos los desarrollos tecnológicos presentes en el coche y los escenarios reales de conducción es necesario, ya que la carga de trabajo del conductor podría aumentar de forma imperceptible, generando así factores de estrés adicionales, como la fatiga (Hu et al., 2020) o la sobrecarga cognitiva (Ramakrishnan et al., 2021), que pueden no estar relacionadas con la actividad de conducir, pero que pueden aumentar el estrés que siente el conductor. En este sentido, la eficacia de la metodología propuesta debe ser probada, ya que puede ser necesario un procedimiento de recalibración para obtener los mejores resultados posibles. Además, se requiere el desarrollo de sistemas portátiles de monitoreo fisiológico que no introduzcan una sensación de incomodidad, es decir, dispositivos vestibles en algunas partes del coche como el volante, los cinturones de seguridad y los asientos, ya que podrían aumentar aún más el estrés que puede desarrollar el conductor. El desarrollo de estas áreas de investigación puede conducir a la masificación de los sistemas de detección de estrés con los beneficios que este tipo de sistemas tiene.

El sistema de adquisición de datos desarrollado durante el posgrado trae consigo mejoras a los sistemas descritos en la sección de antecedentes como una mayor portabilidad debido al tamaño de la tarjeta y los componentes utilizados para su desarrollo, así como la fuente de alimentación la cual puede ser una pila de 9 voltios, sin embargo, dentro de este estudio se comprobó el funcionamiento del mismo con el fin de validar el diseño y la manufactura; como parte de las pruebas hechas a la tarjeta se obtuvieron señales que no consideran todos los aspectos de un viaje en automóvil debido al tiempo y las condiciones actuales para realizar el manejo de un automóvil y con ello produciendo una eficiencia baja en cuanto a las señales adquiridas con este sistema y la metodología propuesta.

5.2 Prospectivas

Como trabajo futuro, primeramente, sería incrementar las ventanas de tiempo de cada segmento con el fin de buscar una mejora en cuanto a la efectividad, además, se pretende realizar pruebas con una base de datos más grande para mejorar la precisión y corroborar el desempeño de la metodología planteada durante el trabajo de esta tesis.

A mediano plazo, se propone realizar una base de datos actualizada con las nuevas causas de estrés en conductores utilizando el sistema de adquisición de datos desarrollado como trabajo tecnológico en el transcurso del posgrado y con ello lograr un nuevo estudio con la misma metodología para corroborar la eficiencia reportada con la base de datos realizada por Healey y Picard (2005); esta base de datos deberá considerar nuevos factores como lo son: el tráfico y las horas pico, uso de aparatos electrónicos y sistemas de navegación, por mencionar algunos, con el fin de mejorar la efectividad del sistema desarrollado ya que las señales adquiridas considerarán un mayor número de factores adecuados a la actualidad.

6. REFERENCIAS

Abdullah, S. (2005). Wavelet bump extraction (WBE) for editing variable amplitude fatigue loadings. University of Sheffield, United Kingdom (Thesis Doctoral).

Afkhami, R. G., Azarnia, G., Tinati, M. A. (2016). Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals. *Pattern Recognition Letters*, 70, 45-51.

Akhbardeh, F., Vasefi, F., MacKinnon, N., Amini, M., Akhbardeh, A., Tavakolian, K. (2019) Classification and Assessment of Hand Arthritis Stage using Support Vector Machine. In: International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); IEEE: Berlin, Germany, 4080-4083.

Alić, B., Sejdinović, D., Gurbeta, L., Badnjevic, A. (2016). Classification of stress recognition using Artificial Neural Network. 2016 5th Mediterranean Conference on Embedded Computing, MECO 2016 - Including ECyPS 2016, BIOENG.MED 2016, MECO: Student Challenge 2016, 297–300. <https://doi.org/10.1109/MECO.2016.7525765>

Amezquita-Sanchez, J., Adeli, A., Adeli, H. (2016) A new methodology for automated diagnosis of mild cognitive impairment (MCI) using magnetoencephalography (MEG). *Behavioural Brain Research*, 305, 174-180.

Anuragi, A., Singh, D. (2019) Alcohol use disorder detection using EEG Signal features and flexible analytical wavelet transform, *Biomed. Signal Process. Control* 52, 384–393, <https://doi.org/10.1016/j.bspc.2018.10.017>

Arduino Store (2021) Arduino Mega 2560 Rev3. 23 de octubre de 2021. Sitio web: <https://store-usa.arduino.cc/products/arduino-mega-2560-rev3?selectedStore=us>

Avendaño, L.E. (2007) *Sistemas Electrónicos Analógicos: Un Enfoque Matricial*, Centro de Publicaciones Universidad Tecnológica de Pereira (1era ed), 596.

Bajaj, V., Taran, S., Khare, S.K., Sengur, A. Feature extraction method for classification of alertness and drowsiness states EEG signals, *Application Acoustics*, 163,107224.

Bashar, S.K., Bhuiyan, M.I.H. (2016) Classification of motor imagery movements using multivariate empirical mode decomposition and short time Fourier transform based hybrid method, *Eng. Sci. Technol.*, 19, 1457–1464,

Becerra, O.D., Chávez, J., Hernández, M. A. (2012). Estrés en operadores de transporte público de carga. (Tesis de licenciatura) Universidad Autónoma de Querétaro.

Ben-Hur, A., Ong, C.S., Sonnenburg, S., Scholkopf, B., Ratsch, G. (2008) Support Vector Machine and Kernels for Computational Biology, 4(10).

Bitter, R., Mohiuddin, T., Nawrocki, M. (2007) LabVIEW: Advanced programming techniques. CRC Press (2da ed), 520.

Boylestad, R.L., Nashelsky, L. (2009) *Electrónica: Teoría de Circuitos y Dispositivos Electrónicos*. Pearson (10ma ed), 912.

Cardone, D., Perpetuini, D., Filippini, C., Spadolini, E., Mancini, L., Chiarelli, A.M., Merla, A. (2020) Driver Stress State Evaluation by Means of Thermal Imaging: A Supervised Machine Learning Approach Based on ECG Signal. *Appl. Sci.*,10(16), 5673.

Carrasco, A. (2019). Desarrolló de un sistema de adquisición y procesamiento basado en la transformada de Fourier y LabView para señales eléctricas cardiacas. (Tesis de Licenciatura). Universidad Autónoma de Querétaro.

Céspedes, R.D. (2015) Evaluación mediante análisis de Monte Carlo de filtros activos basados en redes doble T. (Tesis de licenciatura). Universidad central “Marta Abreu” de las Villas.

Chen, L. Ian, Zhao, Y., Ye, P. fei, Zhang, J., Zou, J. zhong. (2017). Detecting driving stress in physiological signals based on multimodal feature analysis and

kernel classifiers. *Expert Systems with Applications*, 85, 279–291.
<https://doi.org/10.1016/j.eswa.2017.01.040>

Choi, M. G., Kim, J. Y., Jeong, I. K., Kim, Y. H., Kim, J. M. (2018). A Real-Time Monitoring System for Boiler Tube Leakage Detection. In *International Conference on Hybrid Intelligent Systems*, 106-114. Springer, Cham.

Chui, K.T., Lytras, M.D., Liu, R.W. (2020) A Generic Design of Driver Drowsiness and Stress Recognition Using MOGA Optimized Deep MKL-SVM, *Sensors*, 20(5),1474.

Cifuentes, I.A. (2010) Diseño y construcción de un sistema para la detección de señales electromiográficas. Tesis de licenciatura, UADY

Cruz Albarrán, I.A. (2019) Sensor inteligente basado en FPGA para detección de cansancio en conductores automotrices. Tesis doctorado, UAQ.

Cruz Ramirez H. (2018) Metodología basada en la descomposición empírica de modos completa y fractales para la clasificación de la imaginación motora usando señales electroencefalográficas. Tesis Licenciatura, UAQ.

Day, S. (2009). Important Factors in Surface EMG Measurement. *Measurement*, *bortec biomedical*, 1–17.

Delgado-Prieto, M., Cirrincione, G., Espinosa, A. G., Ortega, J. A., Henao, H. (2012). Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks. *IEEE Transactions on Industrial Electronics*, 60(8), 3398-3407.

DEVICE, A. (2011). Low Cost Low Power Instrumentation Amplifier AD620. *Analog Devices*.

Doborjeh, Z., Doborjeh, M., Crook-Rumsey, M., Taylor, T., Wang, G.Y., Moreau, D., Sumich, A. (2020) Interpretability of Spatiotemporal Dynamics of the Brain Processes Followed by Mindfulness Intervention in a Brain-Inspired Spiking Neural Network Architecture. *Sensors*, 20, 7354.

Dorantes-Argandar, G., Cerda-Macedo, E. A., Tortosa-Gil, F., Ferrero Berlanga, J. (2015). Accidentalidad de automóviles de uso particular en México: influencia del estrés y la agresividad. *PSIENCIA. Revista Latinoamericana de Ciencia Psicológica*, 7, 418-427.

Duchêne J., Gouble, F. (1993) Surface electromyogram during voluntary contraction: Processing tools and relation to physiological events. *Critical Reviews in Biomedical Engineering* 21(4), 313–397.

Esler, M. (2017) Mental stress and human cardiovascular disease. *Neurosci. Biobehav. Rev.*,74, 269–276.

Ghandour, R., Neji, B., El-Rifaie, A.M., Al Barakeh, Z. (2020) Driver distraction and stress detection systems: a review. *International Journal of Engineering and Applied Sciences (IJEAS)*, 7(4), 39-46.

Ghorbani, R., Azarnia, G., Tinati, M.A. (2016) Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals, *Pattern Recogn. Lett.* 70, 45–51.

Gupta, V., Bhavsar, A. (2017) Breast Cancer Histopathological Image Classification: Is Magnification Important?" In: *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*; IEEE: Honolulu, Hawaii, 769-776.

Hackett, R.A., Steptoe, A. (2017) Type 2 diabetes mellitus and psychological stress—A modifiable risk factor. *Nat. Rev. Endocrinol.* 13, 547.

Hassan, A.R., Hassan, M.I. (2016) Automatic sleep scoring using statistical features in the EMD domain and ensemble methods, *Biocybernet. Biomed. Eng.* 36, 248–255, .

Healey, J.A., Picard, R.W. (2005) Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions in Intelligent Transportation Systems* 6(2),156-166.

Hecke, T.V. Power study of ANOVA versus Kruskal-Wallis test. *Journal of statistical Managing System*, 15, 241-247.

Hinojosa, R., Jiménez, P.L., Hernández, V., Campos, J. (2012) Siniestralidad por accidentes de tránsito en México: una aproximación desde el análisis espacial. *Geografía y Sistema de Información Geográfica (GEOSIG)*, 4(4), 291-309.

Huircán, J. I. (2012). *Filtros activos, conceptos básicos y diseño*. Universidad de la Frontera, Temuco, 1-11.

Ieracitano, C., Mammone, N., Hussain, A., Morabito, F. (2021) A novel explainable machine learning approach for EEG-based braincomputer interface systems. *Neural Computing and Applications*.

Igba, J., Alemzadeh, K., Durugbo, C., Eiriksson, E. T. (2016). Analysing RMS and peak values of vibration signals for condition monitoring of wind turbine gearboxes. *Renewable Energy*, 91, 90-106.

Iu, H. H. C., Yu, D. S., Fitch, A. L., Sreeram, V., Chen, H. (2011). Controlling chaos in a memristor based circuit using a twin-T notch filter. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 58(6), 1337–1344. <https://doi.org/10.1109/TCSI.2010.2097771>

Jia, J., Goparaju, B., Song, J., Zhang, R., y Westover, M. B. (2017). Automated identification of epileptic seizures in EEG signals based on phase space representation and statistical features in the CEEMD domain. *Biomedical Signal Processing and Control*, 38, 148-157.

Katsis, C.D., Katertsidis, N., Ganiatsas, G., Fotiadis, D.I. (2008) Toward emotion Recognition in Car-Racing Drivers: A Biosignal processing approach. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, 38(3), 502-512.

Keshan, N., Parimi, P. V., Bichindaritz, I. (2015). Machine learning for stress detection from ECG signals in automobile drivers. *Proceedings - 2015 IEEE*

International Conference on Big Data, IEEE Big Data 2015, 2661–2669.
<https://doi.org/10.1109/BigData.2015.7364066>

Khandpur R.S. (2004). Biomedical instrumentations. Technology and applications, McGraw-Hill (1ra ed), 22-96.

Konrad, P. (2006) The abc of EMG, A Practical Introduction to Kinesiological Electromyography, Noraxon USA, Inc., 1(2005), 11.

Lajara, J.R., Pelegrí, J. (2011) LabVIEW: entorno gráfico de programación. Marcombo (2da ed).

Lee, D. S., Chong, T. W., Lee, B. G. (2017). Stress Events Detection of Driver by Wearable Glove System. IEEE Sensors Journal, 17(1), 194–204.
<https://doi.org/10.1109/JSEN.2016.2625323>

Liu, Y., Du, S. (2018). Psychological stress level detection based on electrodermal activity. Behavioural Brain Research, 341, 50–53.

Lucas, M.F., Gaufriau, A., Pascual, S., Doncarli, C., Farina, D. (2008) Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization. Biomedical Signal Processing and Control, 3(2), 169-174.

Rigas, G., Goletsis, Y., Fotiadis, D. I. (2012). Real-time driver's stress event detection. IEEE Transactions on intelligent transportation systems, 13(1), 221-234.

Marieb, E.N. (2008). Anatomía y Fisiología Humana (Martín-Romo, M. (PEARSON)).

Luijckx, R., Hermens, H. J., Bodar, L., Vossen, C. J., Os, J. van., Lousberg, R. (2014). Experimentally Induced Stress Validated by EMG Activity. PLoS ONE, 9(4), e95215. <https://doi.org/10.1371/journal.pone.0095215>

Marshall, G.D., Jr. (2019) Psychological stress, immunity, and asthma: Developing a paradigm for effective therapy and prevention. Curr. Opin. Behav. Sci. 28, 14–19.

Martínez-Carreño, L., Vega-López, E. M., Samano-Flores, Y., Hernández-Zavala, A., Chaparro-Cardenas, S., Jiménez-Garibay, A. A. (2018). Caracterización de esfuerzo muscular del miembro inferior mediante EMG no invasivo. In Memorias del Congreso Nacional de Ingeniería Biomédica (Vol. 5, No. 1, pp. 270-273).

Maurya, P.K., Noto, C., Rizzo, L.B., Rios, A.C., Nunes, S.O.V., Barbosa, D.S., Sethi, S., Zeni, M., Mansur, R.B., Maes, M., et al. (2016) The role of oxidative and nitrosative stress in accelerated aging and major depressive disorder. *Prog. Neuropsychopharmacol. Biol. Psychiatry*, 65, 134–144

Merletti, R. (2004) *Electromyography: Physiology, Engineering, and Noninvasive Applications* (Merletti, R., Parker, P. (Wiley-IEEE Press)).

Merletti, R., Hermens, H.J. (2004) Detection and conditioning of the Surface EMG signal en Merletti R. (Ed.), *Electromyography: Physiology, Engineering, and Noninvasive Applications*. (1ra ed.) Institute of Electrical and Electronics Engineers.

Microchip. (2010). MCP6H01/2/4. Microchip.

Munla, N., Khalil, M., Shanin, A., Mourad, A. (2015) Driver Stress Level Detection Using HRV Analysis. *International Conference on Advances in Biomedical Engineering (ICABME)*, 61-64.

Montiel, M.Y., Santana-Mancilla, P.C., Guerrero, A. (2016) Interfaz cerebro-computadora para disminuir situaciones de riesgo en un entorno de conducción vial. *Avances en Interacción Humano-Computadora*, [S.I.], 1(1), 74-76.

Morales, J.F., Willems, R., Van Huel, S., Varón, C. (2020) Evaluation of the ECG derived respiration in the presence of irregular heart beats. *Proceedings of the 11th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, Pisa, Italy, 20–21.

Namazi, H. (2019). Fractal-based classification of electromyography (EMG) signal in response to basic movements of the fingers. *Fractals*, 27(3).

Noble, W. What is a support vector machine? (2006) *Nature Biotechnology*, 24, 1565-1567.

Organización Mundial de la Salud. (2018). *Global Status Report on Road Safety 2018*. World Health Organization: Geneva, Switzerland.

Ortiz, J.P., Reinoso Avecillas, M.J. (2010) *Diseño y construcción de un prototipo de electroencefalógrafo para adquisición de señales cerebrales*. Tesis licenciatura, Universidad Politécnica Salesiana.

Pavlidis, P., Wapsinski, I., Noble, W.S. (2004) Support vector machine classification on the web. *Bioinformatics*, 20(4), 586-587

Perez-Sanchez, A.V., Perez-Ramirez, C.A., Valtierra-Rodriguez, M., Dominguez-Gonzalez, A., Amezcua -Sanchez, J.P. (2020) Wavelet Transform-Statistical Time Features-Based Methodology for Epileptic Seizure Prediction Using Electrocardiogram Signals, *Mathematics*, 8, 2125.

Phinyomark, A., Phukpattaranont, P., Limsakul, C. (2012) Feature reduction and selection for EMG signal classification. *Expert Systems with Applications*, 39(8), 7420–7431. <https://doi.org/10.1016/j.eswa.2012.01.102>

Piña, A.J. (2015) *Implementación de señales mioeléctricas en un Sistema de control para generar movimientos en una prótesis personalizada de mano*. Tesis de Posgrado, UAQ.

Piña Vega R. (2019) *Metodología basada en dimensión fractal y lógica difusa para la predicción de infartos empleando señales ECG*. Tesis Licenciatura, UAQ.

Ramakrishnan, P., Balasingam, B., Biondi, F. (2021) Cognitive load estimation for adaptive human–machine system automation. En *Learning Control* (1era edición), Elsevier: Amsterdam, pp. 35–58.

Romo Avilés M. (2020) Metodología basada en la transformada wavelet e indicadores estadísticos para la clasificación de señales electromiográficas aplicadas a un brazo robótico. Tesis Licenciatura, UAQ.

Ruiz Vazquez C. (2019) Metodología basada en la transformada wavelet y los parámetros de Hjorth para la clasificación de imaginación motora usando señales electroencefalográficas. Tesis Licenciatura, UAQ.

Schölkopf, B., Smola, A. (2002) Learning with kernels. MIT press: Cambridge, Massachusetts.

Sedano, J., Chira, C., González, J., Villar, J. R. (2012) Sistema para la medida del nivel de estrés. DYNA - Ingeniería e Industria, 87 (3): 336-344.

Sedra, A. (2006) Circuitos microelectrónicos. McGraw Hill (1era ed.), 1392.

Singh, M., Queyam, A. B. (2013). A Novel Method of Stress Detection using Physiological Measurements of Automobile Drivers. International Journal of Engineering Education, 5(2), 13–20.

Silverthorn, D. (2007) Fisiología Humana un enfoque integrado, Texas: Panamericana.

Schmidt-Nielsen, K. 1983. Fisiología Animal. Ed. Omega. Barcelona, Esp. 499pp. (ISBN: 9788428202909).

Slama, A. B., Lentka, Ł., Mouelhi, A., Diouani, M. F., Sayadi, M., Smulko, J. (2018). Application of statistical features and multilayer neural network to automatic diagnosis of arrhythmia by ECG signals. Metrology and Measurement Systems, 25(1).

Smoller, J.W. (2016) The genetics of stress-related disorders: PTSD, depression, and anxiety disorders. Neuropsychopharmacology. 41, 297.

Spence, R. (1988) Tolerance Design of Electronic Circuits. Imperial College Press (1era ed).

Suthaharan, S. (2016) Support vector machine. Machine learning models and algorithms for big data classification. Springer: Boston, MA, 2016, 207-235.

Thukral, A.K.; Bhardwaj, R.; Kumar, V.; Sharma, A. (2019) New indices regarding the dominance and diversity of communities, derived from sample variance and standard deviation. *Heliyon*, 5(10), e02606.

Useche, S. A., Ortiz, V. G., Cendales, B. E. (2017). Stress-related psychosocial factors at work, fatigue, and risky driving behavior in bus rapid transport (BRT) drivers. *Accident Analysis y Prevention*, 104, 106-114.

Useche, S.; Serge, A.; Alonso, F. (2015) Risky Behaviors and Stress Indicators between Novice and Experienced Drivers. *American Journal of Applied Psychology*, 3(1), 11-14.

Vargas López O. (2019) Metodología basada en la EMD y medidas no lineales prevalentes en muerte súbita cardíaca. Tesis Licenciatura, UAQ.

Vega Mancilla, S. Y. (2016). Procesamiento de imágenes termográficas en la detección de cáncer de mama, Tesis de Maestría, Repositorio de la Universidad Autónoma de Querétaro.

Virdi, P., Narayan, Y., Kumari, P., Mathew, L. (2016) Discrete wavelet packet-based elbow movement classification using fine Gaussian SVM. In: 2016 IEEE 1st international conference on power electronics, intelligent control, and energy systems (ICPEICES); IEEE: Delhi, India, 1-5.

Vignolo, B. J. (2008). Introducción al procesamiento digital de señales. Chile. Ediciones Universitarias de Valparaíso, pp. 19-20.

Wang, K., Guo, P. (2020). An Ensemble Classification Model with Unsupervised Representation Learning for Driving Stress Recognition Using Physiological Signals. *IEEE Transactions on Intelligent Transportation Systems*.

Wang, K., Murphey, Y. L., Zhou, Y., Hu, X., Zhang, X. (2019). Detection of driver stress in real-world driving environment using physiological signals. *IEEE*

International Conference on Industrial Informatics (INDIN), 2019-July 1807–1814.
<https://doi.org/10.1109/INDIN41052.2019.8972264>

Wijsman, J., Grundlehner, B., Liu, H., Hermens, H., Penders, J. (2011) Towards mental stress detection using wearable physiological sensors, in Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society Conference, vol. 2011, pp. 1798–801, 2011.9.

Wijsman, J., Grundlehner, B., Penders, J., Hermens, H. (2013). Trapezius muscle EMG as predictor of mental stress. Transactions on Embedded Computing Systems, 12(4). <https://doi.org/10.1145/2485984.2485987>

Xia, Y., Gao, Q., Ye, Q. (2015) Classification of gait rhythm signals between patients with neuro-degenerative diseases and normal subjects: experiments with statistical features and different classification models, Biomed. Signal Process. Control 18, 254–262, <https://doi.org/10.1016/j.bspc.2015.02.002>.

Yanez-Borjas, J. J., Valtierra-Rodriguez, M., Camarena-Martinez, D., Amezcua-Sanchez, J. P. (2020). Statistical time features for global corrosion assessment in a truss bridge from vibration signals. Measurement: Journal of the International Measurement Confederation, 160, 107858. <https://doi.org/10.1016/j.measurement.2020.107858>

Zontone, P., Affanni, A., Bernardini, R., Piras, A., Rinaldo, R. (2019). Stress detection through Electrodermal Activity (EDA) and Electrocardiogram (ECG) analysis in car drivers. European Signal Processing Conference, 2019-September 1–5. <https://doi.org/10.23919/EUSIPCO.2019.8902631>

3M ®. (2020). 3M® Red Dot™ Electrodo para Monitoreo con refuerzo de cinta médica 3M® Micropore™ 2239. 17 de septiembre de 2020, de 3M México Sitio web: https://www.3m.com.mx/3M/es_MX/inicio/todos-los-productos-3m/~/_/3M-Red-Dot-Electrodo-para-Monitoreo-con-refuerzo-de-cinta-m%C3%A9dica-3M-Micropore-2239/?N=5002385+8711017+3294796557yrt=rud

ANEXOS

En esta sección se presentan , (A.1) la portada del artículo publicado y elaborado durante el posgrado sobre la investigación de estrés en conductores automovilísticos, así como las cartas de consentimiento (A.2) y confidencialidad (A.3) que los participantes deberán firmar para ser parte del estudio, así como la responsable del proyecto donde se explica que los datos son totalmente confidenciales y por ningún motivo serán expuesto o utilizados para otro fin, así como la responsabilidad que el investigador para salvaguardar en todo momento la integridad de los voluntarios y su información.

Article

An Explainable Machine Learning Approach Based on Statistical Indexes and SVM for Stress Detection in Automobile Drivers Using Electromyographic Signals

Olivia Vargas-Lopez ¹, Carlos A. Perez-Ramirez ², Martin Valtierra-Rodriguez ¹, Jesus J. Yanez-Borjas ³ and Juan P. Amezcua-Sanchez ^{1,*}

¹ ENAP-Research Group, CA-Sistemas Dinámicos y Control, Facultad de Ingeniería, Campus San Juan del Río, Universidad Autónoma de Querétaro (UAQ), Río Moctezuma 249, San Juan del Río 76807, Mexico; ovargas08@alumnos.uaq.mx (O.V.-L.); martin.valtierra@enap-rg.org (M.V.-R.)

² ENAP-Research Group, Facultad de Ingeniería, Campus Aeropuerto, Universidad Autónoma de Querétaro (UAQ), Carretera a Chichimequillas S/N, Ejido Bolaños, Querétaro 76140, Mexico; carlos.perez@uaq.mx

³ CA Procesamiento Digital de Señales, División de Ingenierías Campus Irapuato-Salamanca (DICIS), Universidad de Guanajuato, Salamanca 36885, Mexico; jjyanezb.ugto@gmail.com

* Correspondence: juan.amezcua@enap-rg.org



Citation: Vargas-Lopez, O.; Perez-Ramirez, C.A.; Valtierra-Rodriguez, M.; Yanez-Borjas, J.J.; Amezcua-Sanchez, J.P. An Explainable Machine Learning Approach Based on Statistical Indexes and SVM for Stress Detection in Automobile Drivers Using Electromyographic Signals. *Sensors* **2021**, *21*, 3155. <https://doi.org/10.3390/s21093155>

Academic Editor: Cosimo Ieracitano

Received: 25 March 2021

Accepted: 28 April 2021

Published: 1 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The economic and personal consequences that a car accident generates for society have been increasing in recent years. One of the causes that can generate a car accident is the stress level the driver has; consequently, the detection of stress events is a highly desirable task. In this article, the efficacy that statistical time features (STFs), such as root mean square, mean, variance, and standard deviation, among others, can reach in detecting stress events using electromyographical signals in drivers is investigated, since they can measure subtle changes that a signal can have. The obtained results show that the variance and standard deviation coupled with a support vector machine classifier with a cubic kernel are effective for detecting stress events where an AUC of 0.97 is reached. In this sense, since SVM has different kernels that can be trained, they are used to find out which one has the best efficacy using the STFs as feature inputs and a training strategy; thus, information about model explain ability can be determined. The explainability of the machine learning algorithm allows generating a deeper comprehension about the model efficacy and what model should be selected depending on the features used to its development.

Keywords: stress detection; EMG signals; statistical time features; support vector machine

1. Introduction

Stress is defined as the reaction of the sympathetic nervous system to any type of threat, which generates a sudden release of hormones such as adrenaline and cortisol into the body [1]. These hormones induce to the body to a state of emergency or alert, which can cause an increase in the heart rate, muscle tension, increased blood pressure, accelerated breathing, and an increased acuity of the senses [2], among other reactions. Hence, this condition can negatively affect the daily life as well as the wellness of a person that experiences frequent stress events [3]. In particular, automobile drivers can be affected by this condition because of negative mood, lane departure, running red lights, traffic noise, congestion, heavy traffic, lack of sleep, driving phobia, impatience, curved narrow roads, and fatigue, among other causes [4–8], which can limit their concentration and ability to make reasonable decisions during any event. In consequence, stress, along with the abovementioned factors, could be an additional factor in certain car accidents that can inflict serious injuries upon those who are involved, even causing deaths [9,10]. In this regard, the World Health Organization reports that each year, about 1.35 million deaths are caused by road traffic crashes, where people from 5 to 29 years old are mainly