



Universidad de Concepción
Facultad de Ingeniería
Departamento de Ingeniería Eléctrica

Predicción de eventos tempranos de somnolencia
mediante un conjunto de datos multifactoriales
en conductores de vehículos

Tesis para optar al grado de Doctor en Ciencias de la Ingeniería
con mención en Ingeniería Eléctrica

Hermes Javier Mora Pantoja
Concepción, Chile
2024

Profesor guía: Dr. Esteban Pino Quiroga
Profesor co-guía: Dr. Tomás Echaveguren Navarro

Resumen

Durante la última década, el número de accidentes vehiculares se ha incrementado significativamente, dejando más de un millón de víctimas fatales cada año. Su principal desencadenante está relacionado con fallas humanas, donde la conducción somnolienta, la distracción y fatiga son sus tres principales causas. La somnolencia es el factor humano que más accidentes causa y es el desafío más difícil de superar al momento de mejorar la seguridad en la vía. Los principales estudios muestran que, en Chile, los accidentes de tránsito han pasado a ser la segunda causa de muerte en la población joven menor de 30 años, dejando más de 86 mil accidentes de tránsito en 2022. Además, como la somnolencia es una manifestación de una causa subyacente originada por múltiples factores como la monotonía de la tarea, el entorno del camino, el tiempo transcurrido en la tarea o la privación del sueño previo, no existe un único método para su detección.

En este sentido, un sujeto somnoliento tiene una reducida capacidad de percepción del espacio y velocidad, aumentando el tiempo de reacción y reduciendo la coordinación ojo-manos. Esto pone en grave riesgo tanto al conductor como a los demás actores de la vía. En consecuencia, se han diseñado diversos proyectos de investigación y dispositivos basados en señales fisiológicas del chofer, evaluaciones de su desempeño o datos del vehículo para detectar si el conductor está somnoliento o alerta en cada intervalo de tiempo. Estos dispositivos alertan al automovilista sobre su estado somnoliento presente para que tome acción sobre la situación. No obstante, estimar el tiempo en el cual el conductor experimentará un evento somnoliento dentro de los subsiguientes minutos es mucho más desafiante que monitorear su estado actual. Este enfoque permite alertar al conductor mientras aún tiene una adecuada capacidad de reacción ya que no está bajo efectos de la somnolencia.

Por lo tanto, en este trabajo de investigación se expone el diseño de un sistema basado en un algoritmo de regresión y un modelo de pronóstico para estimar el tiempo de ocurrencia del subsiguiente evento somnoliento. En estos modelos se utilizan datos multimodales compuestos por señales de Electroencefalografía (EEG), Electrocardiografía (ECG), Electromiografía (EMG), fuerza de agarre mediante sensores Force Sensitive Resistor (FSR) y datos técnicos del vehículo virtual, obtenidos en un experimento de conducción simulada con 36 participantes. El modelo de regresión utiliza redes CNNs mediante una arquitectura ResNet y genera una salida del nivel de somnolencia para el tiempo t_0 . El algoritmo de pronóstico está basado en redes Long Short Term Memory (LSTM) bidireccionales y emplea como entrada los resultados del modelo de regresión para estimar el nivel de somnolencia de los siguientes minutos ($t_0 \rightarrow t_{0+n}$) en intervalos de 28 segundos.

Adicionalmente, los modelos implementados también se entrenaron con diferentes combinaciones y categorías de sensores, buscando la mejor categoría y número de electrodos para definir el sistema con la precisión más alta. Como resultado, el análisis inicial de los datos reveló que al menos el 50 % de los participantes experimentó somnolencia a partir de los siete minutos de iniciado el tercer bloque experimental. Asimismo, al utilizar los datos EEG de prueba, el mejor sistema implementado obtuvo un valor *test_accuracy* de 0.887 en el rango [0 a 5] minutos ($t_0 \rightarrow t_{0+5}$) y 0.798 a los siete minutos. Además, el modelo de regresión alcanzó un valor *test_accuracy* de 0.905 al utilizar únicamente datos FSR. Por otro lado, el rendimiento de los modelos propuestos para la etapa de pronóstico se validaron utilizando la métrica Root Mean Square Error (RMSE) en las fases de entrenamiento, validación y prueba con sus respectivos data-sets.

Los resultados de este trabajo permiten concluir que es posible utilizar datos fisiológicos del conductor para pronosticar el tiempo de aparición de subsiguientes eventos somnolientos dentro de una ventana temporal definida. Asimismo, el uso de datos FSR mostró resultados prometedores en

los algoritmos de regresión y pronóstico, toda vez que estos sensores no son invasivos y permiten aplicaciones adicionales como monitorear el confort del conductor y pasajeros al ubicarse en los asientos del vehículo.

Palabras clave: FSR, EEG, ECG, EMG, Pronóstico, BiLSTM, ResNet.

En un sentido muy real, todos nosotros tenemos dos mentes,
una mente que piensa y otra mente que siente ...

Daniel Goleman

Me gustaría expresar mis más sinceros agradecimientos a mi supervisor de tesis, Dr. Esteban Pino Q. y a mi co-supervisor Dr. Tomás Echaveguren N. Su comprensión, colaboración y paciencia contribuyeron significativamente en el complejo camino de esta investigación.

Durante mis estudios en el Programa de Doctorado recibí el apoyo financiero del Programa de Becas para Estudios de Doctorado Nacional de la Agencia Nacional de Investigación y Desarrollo (ANID) Chile / Doctorado Nacional / DOCTORADO BECAS CHILE / 2020 - 21201262.

En la etapa del desarrollo experimental para esta tesis obtuve el apoyo financiero del Programa de Neurociencia, Psiquiatría y Salud Mental (NEPSAM-UdeC) de la Universidad de Concepción, Chile 2022.

En mi etapa de estudiante tesista recibí el apoyo financiero del Centro Nacional de Inteligencia Artificial CENIA FB210017, Financiamiento Basal ANID.

Tabla de contenido

	Page
Índice de figuras	IV
Índice de tablas	V
1 Introducción	1
1.1 Contextualización	1
1.2 Hipótesis	3
1.3 Objetivo general	4
1.4 Objetivos específicos	4
1.5 Temario	4
2 Estado del arte	5
2.1 Sistemas de alerta en estados somnolientos del conductor	5
2.2 Pronóstico del estado somnoliento del conductor	6
2.3 Discusión	8
3 Materiales y métodos	12
3.1 Introducción	12
3.2 Detección de la intención de frenada de emergencia	12
3.2.1 Configuración experimental y materiales	13
3.2.2 Métodos de preprocesamiento	13
3.2.2.1 Análisis por componentes independientes	14
3.2.2.2 Segmentación de los datos	14
3.2.3 Etapa de procesamiento	16
3.2.3.1 Análisis tiempo-frecuencia	16
3.2.3.2 Estimación de los grupos de electrodos	18
3.2.3.3 Tiempo de reacción en frenada de emergencia	18
3.2.4 Algoritmo implementado	19
3.3 Detección y pronóstico de eventos somnolientos	20
3.3.1 Diseño experimental y materiales	21
3.3.2 Protocolo del experimento	22
3.3.3 Preprocesamiento de los datos	24
3.3.3.1 Nivel real de somnolencia (ground truth)	24
3.3.4 Método 1: algoritmos de clasificación y pronóstico	29
3.3.4.1 Modelo de clasificación	30
3.3.4.1.1 Red neuronal residual	31
3.3.4.2 Algoritmo de pronóstico basado en señales	32
3.3.5 Método 2: algoritmos de regresión y pronóstico	36

3.3.5.1	Modelo basado en algoritmos de regresión	36
3.3.5.2	Algoritmo de pronóstico basado en niveles de somnolencia	38
3.3.5.3	Datos para los niveles de somnolencia	42
4	Resultados	45
4.1	Introducción	45
4.2	Detección de la intención de frenada de emergencia	45
4.2.1	Tiempo de reacción en frenada de emergencia	45
4.2.2	Potenciales relacionados con eventos	48
4.3	Detección y pronóstico de eventos somnolientos	50
4.3.1	Método 1: algoritmos de clasificación y pronóstico	52
4.3.1.1	Modelos de clasificación	52
4.3.1.2	Algoritmo de pronóstico basado en señales	54
4.3.2	Método 2: algoritmos de regresión y pronóstico	60
4.3.2.1	Modelos basados en algoritmos de regresión	60
4.3.2.2	Algoritmo de pronóstico basado en niveles de somnolencia	62
5	Discusión	67
5.1	Detección de la intención de frenada de emergencia	67
5.2	Detección y pronóstico de eventos somnolientos	67
5.2.1	Tamaño de los datos y desempeño de los modelos	68
5.2.1.1	Datos en los niveles de somnolencia	73
5.2.2	Ventanas: history and forecast time windows	73
5.2.3	Análisis comparativo de los resultados	74
5.2.4	Evaluación del desempeño durante el experimento	75
6	Conclusiones	78
	Anexo A: Publicaciones y Congresos	81
	Bibliografía	93
	Anexo B: Encuesta de aptitud, salud y experiencia del participante	94
	Anexo C: Encuesta post-conducción	99
	Anexo D: Encuesta sobre el estilo de conducción del participante	103
	Anexo E: Formulario de Consentimiento Informado	112

Índice de figuras

3.1	Esquema de trabajo en la segmentación de las señales para obtener las muestras de dos clases	15
3.2	Flujo de trabajo en la segmentación y el procesamiento de los datos	16
3.3	Imagen en escala de grises de un segmento de la clase objetivo.	17
3.4	Esquema de la matriz- $2D$ con variación del número de electrodos.	17
3.5	Ubicación de los sensores FSR y módulo ESP32 en el volante del simulador	21
3.6	ubicación de los sensores y demás partes del simulador	22
3.7	Descripción de las etapas que componen el experimento de conducción simulada . . .	22
3.8	Escenario virtual de un circuito implementado en el experimento	23
3.9	Puntos faciales de referencia utilizados en la detección de parpadeos	25
3.10	Esquema del conteo de parpadeos para un sujeto	25
3.11	Segmentos sintéticos de electrodos EEG generados con el método SMOTE	27
3.12	Matriz de correlación de un segmento EEG para el electrodo FP2	28
3.13	Matriz de correlación entre bloques completos para el electrodo Oz	28
3.14	Correlación cruzada entre bloques completos para el electrodo Oz	29
3.15	Matriz de correlación entre clases para el electrodo Oz	29
3.16	Matriz de correlación para el electrodo Oz y correlación cruzada entre clases para el mismo electrodo	30
3.17	Diagrama del modelo de clasificación implementado en el método de clasificación y pronóstico	31
3.18	Bloques de aprendizaje directo y bloques residuales en los modelos implementados .	32
3.19	Diagrama de flujo de operación del sistema configurado por los modelos clasificación-pronóstico.	34
3.20	Esquema del sistema configurado por los modelos clasificación-pronóstico.	35
3.21	Implementación del modelo de pronóstico en base a redes Convolutional Neural Network (CNN)s.	36
3.22	Diagrama de los modelos independientes en el método ensemble modeling.	37
3.23	Etapas iniciales de los submodelos para obtener los mapas de características independientes y configurar un Global Feature Map (GFM).	37
3.24	Diagrama de flujo de operación del sistema configurado con los modelos regresión-pronóstico.	39
3.25	Esquema del sistema basado en los algoritmos de regresión y pronóstico.	40
3.26	Esquema del modelo de pronóstico basado en capas dobles de LSTMs (BiLSTM). . .	41
3.27	Nueva serie temporal con los data-sets para entrenamiento y validación en los datos Drowsiness Levels (DLs).	43
3.28	Diagrama de flujo para obtener nueva serie de valores DLs con data-sets para entrenamiento y validación.	44

4.1	Colisiones totales en cada uno de los tres bloques del experimento.	46
4.2	Resultados de las métrica ROC(AUC) en el método matricial.	48
4.3	Resultados de la métrica ROC(AUC) en el método con imágenes.	48
4.4	Resultados de las métricas Area Under the Curve (AUC) y accuracy en los conjuntos de electrodos e imágenes.	49
4.5	Coeficientes de correlación entre los datos para el método matricial.	50
4.6	Latencias de la onda P300 en la región de la Corteza Visual Primaria.	50
4.7	Latencia de la onda P300 en la región de la Corteza Motora Primaria.	51
4.8	Promedio de la onda P300 en las regiones Primary Visual Cortex (PVC) y Primary Motor Cortex (PMC).	51
4.9	Segmento de señal ECG pronosticado y datos verdaderos para el mismo segmento.	55
4.10	Correlación cruzada entre el segmento de señal ECG pronosticado y los datos verdaderos.	55
4.11	Segmento pronosticado de señal EMG en el electrodo uno y datos verdaderos.	56
4.12	Correlación cruzada entre los datos estimados y sus valores verdaderos.	56
4.13	Segmento de pronóstico en la señal de EEG para el electrodo FP1.	57
4.14	Correlación cruzada entre el segmentos de pronóstico y los datos de prueba en EEG-FP1.	57
4.15	Resultados de precisión en los diferentes sistemas para el tiempo de pronóstico de subsecuentes eventos somnolientos.	59
4.16	Métrica Mean Square Error (MSE) y función de validación del modelo de pronóstico.	62
4.17	Resultados de la métrica MSE y función de pérdida en la validación del modelo de pronóstico.	63
4.18	Estimación del tiempo Forecasted Time of Drowsiness-1 (FTD-1) en la ventana de siete minutos.	63
4.19	Comparación de valores históricos y pronosticados en una ventana de 14 minutos.	64
4.20	Variación de los niveles de somnolencia estimados en la ventana Forecasted Time Window (FTW).	64
4.21	Valores de somnolencia en la ventana FTW sin valores de tiempo Forecasted Time of Drowsiness-1 (FTD-1).	65
5.1	Datos de entrada con segmento de pronóstico ECG y datos de prueba.	69
5.2	Correlación entre el segmento de pronóstico ECG y los datos verdaderos	70
5.3	Datos de entrada y segmento de pronóstico en la señal EEG electrodo PO7.	70
5.4	Correlación entre el segmento de pronóstico EEG electrodo PO7 y los datos verdaderos.	71
5.5	Correlación cruzada segmento de pronóstico de la señal EEG electrodo PO7.	71
5.6	Histograma del segmento pronosticado en la señal EEG electrodo PO7.	72
5.7	Histograma de los datos de prueba en la señal de EEG electrodo PO7.	72
5.8	Evaluación del desempeño de un sujeto según el número de errores e infracciones durante la conducción.	76
5.9	Estimación global del desempeño de todos los sujetos según el número de errores e infracciones.	76
5.10	Número total de muestras de las tres clases para todos los participantes en el bloque dos.	77
5.11	Cantidad total de muestras en las tres clases para todos los participantes en el bloque tres.	77

Índice de tablas

2.1	Trabajos más relevantes orientas a la detección o clasificación del estado somnoliento en conductores de vehículos.	9
2.2	Trabajos previos en la detección y predicción del estado somnoliento subsecuente en conductores de vehículos.	10
3.1	Número total de segmentos <i>objetivo</i> y <i>no-objetivo</i> según la longitud establecida en las señales para los tres bloques.	15
3.2	Configuración de los diferentes grupos de electrodos relacionados por región cerebral.	19
3.3	Ciclos y tiempo de entrenamiento por cada grupo de electrodos y conjunto de imágenes en escala de grises.	20
3.4	Parámetros de configuración en el diseño del modelo básico para los grupos de electrodos.	20
3.5	Técnica de medición subjetiva y descripción del método de medición continua diseñado para este trabajo.	26
4.1	Resultados de la medición del tiempo de reacción de frenada de emergencia en cada uno de los tres bloques.	46
4.2	Resultados de las métricas accuracy y Area Under the Curve (AUC) para cada uno de los grupos de imágenes y electrodos.	47
4.3	Resultados en el entrenamiento del primer algoritmo para la detección del estado del conductor utilizando la base de datos balanceada mediante el método SMOTE.	53
4.4	Resultados en el entrenamiento del primer algoritmo en la detección del estado del conductor al utilizar la base de datos balanceada con el método SMOTETomek.	53
4.5	Resultados del entrenamiento de los modelos para la estimación del estado subsecuente del conductor utilizando la base de datos balanceada con el método SMOTE.	54
4.6	Estimación del estado somnoliento del conductor mediante la base de datos balanceada con el método SMOTETomek.	58
4.7	Resultados de precisión en los diferentes sistemas para el tiempo de pronóstico de subsecuentes eventos somnolientos.	59
4.8	Resultados de la detección del estado del conductor mediante la base de datos balanceada con el método SMOTE.	61
4.9	Resultados en la detección del estado del conductor mediante la base de datos balanceada con el método SMOTETomek.	61
5.1	Número de electrodos ubicados por región cerebral en la detección y predicción de eventos somnolientos.	73

Lista de acrónimos

EEG	Electroencefalografía
ECG	Electrocardiografía
EMG	Electromiografía
FSR	Force Sensitive Resistor
RMSE	Root Mean Square Error
ISM	Independent Sub-Model
GFM	Global Feature Map
MSE	Mean Square Error
FTD-1	Forecasted Time of Drowsiness-1
FTW	Forecasted Time Window
HTW	History Time Window
DLs	Drowsiness Levels
BRT	Braking Reaction Time
std	standard deviation
IQR	Inter-Quartile Range
CNN	Convolutional Neural Network
AUC	Area Under the Curve
ROC	Receiver Operating Characteristic
PVC	Primary Visual Cortex
VAC	Visual Association Cortex
PAA	Posterior Association Area
PMC	Primary Motor Cortex
FC	Frontal Cortex
MAC	Motor Association Cortex
PmC	Premotor Cortex
PfC	Prefrontal Cortex

ICA Independent Component Analysis
BSS Blind Source Separation
ICs Independent Components
STFT Short-Time Fourier Transform
BS-V blurry short-videos
BET Blink Elapsed Time
PERCLOS PERcentage of eyelid CLOSure
KSS Karolinska Sleepiness Scale
EAR Eye Aspect Ratio
SMOTE Synthetic Minority Over-Sampling Technique
Tomek-Links Under-Sampling Technique
CNNU Condensed Nearest Neighbors Undersampling
ResNet Residual Network
LSTM Long Short Term Memory
BiLSTM Bidirectional Long Short Term Memory
RNN Recurrent Neural Network
ADAS Advanced Driver-Assistance Systems
DL Deep Learning
VAR Vector Autoregression
ARIMA Autoregressive Integrated Moving Average
ANN Artificial Neural Network
PPG Pulse Plethysmography
EDA Electrodermal Activity
ERP Event Related Potential

INTRODUCCIÓN

1.1. Contextualización

Conducir bajo efectos de la somnolencia, distracción o fatiga es la principal causa de accidentes automovilísticos fatales en todo el mundo. Aproximadamente el 20 % del total de accidentes vehiculares con víctimas mortales son causados por conductores somnolientos [1, 2]. Según la Administración Nacional de Seguridad del Tráfico en Carreteras (NHTSA) [3], de 2020 a 2021 se presentó un incremento aproximado del 10 % en el número de muertes por esta clase de accidentes en los EE. UU., donde el 1,6 % involucró a conductores somnolientos. Además, los accidentes causados por conductores distraídos representaron el 8,2 % del total en 2021, mientras que la conducción con sueño aumentó un 8,2 % en 2020. En Chile, anualmente se presentan más de 86000 siniestros viales, de los cuales, entre el 15 y 30 % están asociados a la somnolencia, dejando cerca de 480 víctimas fatales y 2400 personas con lesiones graves [4]. En Europa, alrededor del 16 % de las muertes en carretera para el año 2016 fueron causadas por conductores somnolientos en vehículos pesados y un 2 % por conductores fatigados de autobuses [5]. Según un estudio realizado en esta misma región, el 70 % de los conductores de autobuses y el 59 % de los conductores de camiones afirmaron que los viajes largos les provoca fatiga y somnolencia [6], lo que aumenta la posibilidad de sufrir un accidente mortal.

Adicionalmente, la cantidad de automovilistas fatigados y somnolientos en las carreteras es considerablemente mayor en países con un Producto Interior Bruto (PIB) per cápita medio o bajo. En estos territorios, por cada cien mil habitantes, hay una elevada tasa de vehículos que generalmente carecen de sistemas avanzados para la asistencia al conductor o Advanced Driver-Assistance Systems (ADAS) que pueden ayudar a prevenir accidentes mientras se conduce un vehículo; esto suele estar directamente relacionado con el costo final de los sistemas en automóviles nuevos. En términos económicos, esta clase de accidentes costaron en Chile más de UD\$ 5485 millones (2.01 % del PIB) y el daño a vehículos por colisión costó más del 0.3 % del PIB en el año 2020 [7, 8]. En los sectores de la minería, energético y de la construcción, los costos por accidentes causados por somnolencia en los operarios fueron mayores a USD\$ 2.100 millones en reparación a las víctimas en 2016 [4]. Adicionalmente, en el sector de la minería, evitar un solo accidente fatal por trayecto en un operario somnoliento puede llegar a representar un ahorro aproximado de USD\$ 2.75 millones por concepto de reparación a la víctima, reparación de la maquinaria y demás actores involucrados.

Por otro lado, en los últimos años se ha realizado un considerable número de trabajos sobre el comportamiento del conductor al volante. Estos permiten afirmar que los automovilistas

somnolientos degradan significativamente su capacidad para controlar el vehículo y reaccionar ante posibles accidentes [3]. La fatiga y somnolencia afectan principalmente la capacidad del conductor para mantenerse alerta en la carretera y responder adecuadamente a los eventos que suceden en el entorno. Esto genera un mayor tiempo de reacción durante las emergencias, un nivel de vigilancia reducido, una reducida capacidad para realizar múltiples tareas, una insuficiente percepción de velocidad, una reducción en la capacidad de tomar decisiones y en la coordinación extremidades-ojos [9, 10]. El estado somnoliento de un conductor también pone en peligro a otros usuarios de la vía; conducir requiere una serie de habilidades físicas, psicológicas y fisiológicas complejas [11, 12], que varían de un conductor a otro.

Además, si la somnolencia es una necesidad subjetiva y excesiva de dormir como resultado de múltiples factores provocadores o de causas subyacentes que suceden en eventos donde la persona debería permanecer despierta, puede derivar en eventos de microsueño, definidos como períodos de más de cuatro segundos de falta de atención e inconsciencia involuntaria. En este tiempo, por ejemplo, un vehículo que vaya a velocidad de autopista puede seguir avanzando en línea recta en lugar de tomar una curva o cruzar una distancia equivalente a un campo de fútbol, provocando colisiones frontales o entrar en el carril contrario sin percatarse del evento de microsueño ni poder reaccionar adecuadamente. En consecuencia, dadas las diferencias intrínsecas de cada persona, algunos conductores son más susceptibles que otros a determinados factores provocadores de somnolencia. Según [5], los conductores jóvenes se encuentran en un grupo de alto riesgo de sufrir accidentes automovilísticos, especialmente aquellos causados por conductores somnolientos y con falta de experiencia al volante.

La conducción monótona en viajes nocturnos provoca más somnolencia y fatiga en conductores jóvenes comparado con los conductores de mayor edad. De este modo, los conductores mayores y experimentados tienen mejores resultados al conducir cuando descansan regularmente en comparación con aquellos que no duermen bien o conducen por largo tiempo sin pausas intermedias [13]. Además, las personas de mayor edad tienden a regular voluntariamente su conducción para evitar algunas situaciones peligrosas específicas durante su tarea [9, 14, 15], pero cuando no regulan la ejecución de su tarea, enfrentan un grado de somnolencia mucho mayor. No obstante, tanto los conductores jóvenes como los mayores experimentan niveles similares de fatiga y somnolencia en condiciones de tráfico vehicular intenso o monotonía del entorno, pero la respuesta a este factor depende de la edad y experiencia al volante. Por esta razón, en la mayoría de los accidentes vehiculares de conductores jóvenes se involucra la somnolencia como factor provocador, mientras que en los conductores mayores y poco experimentados se asocia el microsueño como la principal causa.

En este sentido, diferentes trabajos de investigación se han centrado en detectar la somnolencia en los conductores. Algunos enfoques se basan en un único tipo de señales fisiológicas [16–20], mientras que otros han utilizado más de una fuente fisiológica de información del conductor. Por ejemplo, en los trabajos de [21–24], emplearon al menos dos modalidades diferentes de datos para clasificar el estado del conductor. Un sistema multimodal puede mejorar el rendimiento de un sistema de detección al ampliar su capacidad para detectar más indicadores de somnolencia o fatiga. Las señales de EEG proporcionan la información más precisa sobre la somnolencia porque miden de forma directa y segura la actividad cortical del conductor [25–27]. No obstante, estas señales son altamente susceptibles al ruido de fuentes y artefactos biológicos [28–30]. Por lo tanto, combinar múltiples fuentes de información como EMG, ECG y Electrodermal Activity (EDA) compensa las desventajas de las señales de EEG y mejora el rendimiento del modelo [31–33]. Además, las medidas del vehículo y los datos de rendimiento del conductor también pueden utilizarse en combinación

con señales fisiológicas [34,35], resultando en una mejora de la capacidad del modelo para evaluar el estado del conductor.

En la actualidad, la mayoría de estos trabajos están centrados en la detección del estado somnoliento del conductor mediante algoritmos de Machine Learning. Los modelos más utilizados se basan en redes convolucionales o CNNs como método de clasificación y en redes neuronales o Artificial Neural Network (ANN) para algoritmos de regresión no lineal, donde el grado de somnolencia del conductor se mide como un valor continuo [36–39]. Por otro lado, a diferencia de clasificar el estado alerta y somnoliento del conductor en el tiempo t_0 , predecir la condición subsecuente del conductor para los minutos siguientes ($t_0 \rightarrow t_{0+n}$) es un enfoque altamente desafiante [40,41]. Este método involucra pronosticar o estimar el momento en el cual el conductor estará bajo efectos de la somnolencia en los siguientes minutos. En este método, si t_{0+n} es muy cercano a t_0 el modelo tiene una mejor precisión porque el número de datos a pronosticar son mínimos y pueden aproximarse por métodos probabilísticos menos complejos. Sin embargo, si n está en el orden de minutos, tal que $t_n \gg t_0$, este método se torna mucho más complejo y menos preciso, donde los métodos basados en probabilidad no son efectivos.

Una técnica para abordar el complejo problema de pronóstico es emplear un modelo basado en LSTM que permiten analizar datos secuenciales y encontrar relaciones no lineales entre la entrada y salida [42,43]. En consecuencia, en este trabajo se aborda el problema de predecir el tiempo de aparición de subsiguientes eventos de somnolencia mediante un modelo de pronóstico basado en LSTM. Para ello, se combina un algoritmo de regresión para obtener los niveles de somnolencia (modelo 1) con un algoritmo de pronóstico (modelo 2). Además, se implementaron varias combinaciones de fuentes de datos (EEG, ECG, EMG, FSR, Telemetría del vehículo) para determinar la mejor combinación y número de sensores. Con estos data-sets también se configuró una base de datos multimodal que permite evaluar el desempeño de los modelos con un enfoque multivariable.

En relación con la base de datos multimodal, esta se obtuvo mediante un experimento de conducción simulada, donde a los participantes se les indujo a presentar eventos de somnolencia por desactivación. Este tipo de somnolencia no surge por falta de sueño, sino por la reducción de estímulos externos que mantienen al cerebro alerta. Durante la conducción, la atención y el estado de alerta disminuyen cuando el entorno es predecible y repetitivo, como en carreteras rectas y poco transitadas. En estas situaciones, el sistema nervioso entra en un estado de menor actividad cognitiva, lo que provoca que el conductor no perciba los primeros signos de somnolencia y tenga una mayor predisposición a quedarse dormido. En el experimento, los participantes se expusieron a condiciones de baja estimulación física y cognitiva, obteniendo los datos fisiológicos de (EEG, ECG, EMG, FSR) que reflejan la progresión del estado de alerta al de somnolencia.

1.2. Hipótesis

Los datos fisiológicos de EEG, ECG, EMG, FSR y datos técnicos de aceleración, posición del volante, desviación del pedal de los frenos y acelerador de un vehículo virtual, obtenidos en un experimento de conducción simulada, permiten identificar patrones cuantificables de somnolencia y entrenar un algoritmo de pronóstico para predecir eventos subsecuentes de somnolencia en conductores de vehículos.

1.3. Objetivo general

Predecir eventos subsecuentes de somnolencia en conductores de vehículos mediante un algoritmo de pronóstico, partiendo de un conjunto de datos multimodales del conductor y técnicos del vehículo obtenidos en un experimento de conducción simulada.

1.4. Objetivos específicos

- Diseñar un experimento de conducción simulada para inducir somnolencia por desactivación en los participantes del experimento.
- Conformar una base de datos multimodal compuesta de señales fisiológicas de EEG, ECG, EMG, FSR de un grupo de conductores de vehículos y datos técnicos de aceleración, posición del volante, desviación del pedal de los frenos y acelerador de un vehículo virtual obtenidos en una conducción simulada.
- Entrenar un algoritmo de regresión mediante el conjunto de datos multimodal para clasificar los niveles de somnolencia en el evento t_0 .
- Entrenar un algoritmo de pronóstico a través de diferentes combinaciones de los datos obtenidos para estimar los subsecuentes eventos somnolientos.
- Identificar el subconjunto de sensores más eficaces y adecuados para detectar somnolencia en conductores de vehículos.

1.5. Temario

Este trabajo se describe en el siguiente orden:

- Capítulo 2. Presenta una revisión y discusión del estado del arte relacionado con la detección y predicción de eventos de somnolencia en conductores de vehículos
- Capítulo 3. Se describen todos los materiales utilizados en el experimento de laboratorio para obtener la base de datos, los métodos implementados para el procesamiento de las señales obtenidas y el desarrollo de los modelos implementados.
- Capítulo 4. Presenta los resultados obtenidos con los diferentes algoritmos y métodos implementados.
- Capítulo 5. Presenta una discusión y análisis de los resultados obtenidos en el desarrollo de esta tesis.
- Capítulo 6. Recopila las diferentes conclusiones a las que se llegó durante el desarrollo de esta investigación.

ESTADO DEL ARTE

2.1. Sistemas de alerta en estados somnolientos del conductor

La somnolencia en el conductor conlleva una necesidad involuntaria de dormir, donde requiere pasar de una fase de alerta a un estado de dormido, involucrando el primer estado non-rapid eye movement sleep (NREM) [44]. Por lo tanto, los sistemas de monitorización están orientados a detectar el estado somnoliento del conductor. Este proceso también ha encontrado diversas aplicaciones en otras áreas de la industria, ya que la somnolencia y sobrecarga cognitiva en la ejecución de una tarea afecta indistintamente a los operarios. El enfoque principal de estos sistemas es detectar los eventos somnolientos del conductor para activar un sistema de alerta, buscando que el conductor reaccione o detenga la marcha del vehículo. En este sentido, básicamente existen dos métodos de medición del nivel de somnolencia: *a)* Los métodos subjetivos de auto evaluación, generalmente basados en la escala de Karolinska Sleepiness Scale (KSS) [45], aunque puede haber otras escalas, esta es la más utilizada. *b)* Los métodos objetivos basados en datos fisiológicos del conductor, datos del vehículo o datos híbridos (conductor-vehículo).

Los sistemas subjetivos dependen de la auto-percepción de somnolencia que el conductor tenga en cada momento de evaluación. Por ello, no son utilizados en entornos reales como métodos de alerta, pero pueden utilizarse a nivel de investigación para discriminar los diferentes niveles de somnolencia entre sujetos en un experimento. El método de KSS está dividido en nueve escalas con anclajes verbales para cada nivel [46], pasando del estado de máxima alerta al de máxima somnolencia. Las calificaciones subjetivas no siempre coinciden con las mediciones fisiológicas o de comportamiento basadas en datos objetivos, esto es debido a que el nivel de somnolencia se mide en intervalos de tiempo fijo que no permiten detectar variaciones intermedias según las medidas subjetivas. En relación con las mediciones objetivas, estas siempre se basan en datos continuos obtenidos en periodos de varios minutos, facilitando una evaluación permanente en cualquier instante de tiempo.

Por otro lado, los sistemas orientados a monitorear el estado de somnolencia del conductor están basados en alguno de los siguientes métodos: *a)* Sistemas centrados en el vehículo. Constantemente se hace una monitorización de la desviación de la posición del vehículo sobre el carril, el movimiento del volante o la presión sobre el pedal de aceleración. Cualquier cambio en estas métricas que sobrepase de un umbral predefinido indica una alta probabilidad de que el conductor está somnoliento [34, 47]. *b)* Sistemas centrados en medidas de comportamiento. El comportamiento del conductor se determina mediante algunos datos como postura de la cabeza o postura corporal al conducir. Esto

incluye cualquier dato que no sea fisiológico, pero que identifique el comportamiento del conductor [48]. *c)* Sistemas centrados en el conductor. En este punto se utilizan datos fisiológicos como EEG, ECG, EMG, conteo de bostezos y parpadeos o porcentaje de cierre ocular. Este es el método más utilizado y preciso, donde los datos permiten una monitorización constante del conductor [49–51]. *d)* Sistemas híbridos. Estos sistemas están basados en una combinación de algunos de los métodos anteriores. Por ejemplo, el uso de datos del vehículo y mediciones fisiológicas, lo cual aumenta la precisión del sistema.

Por lo tanto, algunos vehículos integran sistemas conocidos como ADAS [52–54], los cuales están centrados en el vehículo. Estos dispositivos utilizan una gran variedad de sensores para monitorear condiciones del entorno y del vehículo, pudiendo incluso tener control sobre el automóvil. Sin embargo, son sistemas cada vez más costosos, lo cual hace que estén presentes mayoritariamente en vehículos lujosos, súper equipados y de alto costo. Esto deja a sistemas ADAS más básicos y menos precisos para vehículos compactos o de costo reducido. No obstante, todos estos sistemas se orientan a detectar el estado somnoliento del conductor y alertarlo en caso de manifestar señales de somnolencia. En consecuencia, la mayoría de los trabajos actuales se han centrado en utilizar diversos tipos de datos fisiológicos del conductor en combinación con nuevos algoritmos para detectar el estado del conductor [55–60]. Los primeros algoritmos con este enfoque utilizaban un análisis manual de las diversas características presentes en las señales. Por ejemplo, el análisis de la variabilidad del ritmo cardíaco en las señales de ECG como indicador de somnolencia o estado de alerta del conductor.

Sin embargo, en la actualidad existen diversas técnicas para el análisis de esta clase de datos a fin de extraer las mejores características para implementar algoritmos más precisos: *a)* Modelos matemáticos, los cuales ofrecen un análisis cuantitativo del efecto de la somnolencia en el desempeño del conductor [61]. *b)* Análisis basado en reglas. Este es uno de los enfoques menos desafiante en su implementación, aunque menos precisos. Pero los sistemas de inferencia difusa (FIS) son más precisos que los sistemas simples basados en reglas [62,63]. *c)* Análisis basado en Machine Learning [64–68]. Constituyen el estado del arte en el procesamiento de las señales fisiológicas y algoritmos para la detección de somnolencia. Existe una gran variedad de algoritmos. Por ejemplo, modelos poco profundos, máquinas de soporte vectorial o modelos Deep Learning (DL). Puede considerarse que los modelos DL basados en redes neuronales convolucionales o CNN son los más utilizados [36,37]. Las redes CNN son ampliamente utilizadas en visión por computadora y en el análisis de diferentes clases de imágenes donde destacan por su alta precisión y porque requieren de grandes cantidades de datos para su entrenamiento.

2.2. Pronóstico del estado somnoliento del conductor

Un conductor somnoliento tiene una reducida capacidad de reacción, percepción, decisión y un tiempo de reacción mucho mayor al de un chofer alerta [47,69]. Esto deja al conductor en un estado de alto riesgo de sufrir un accidente vehicular. Además, cuando el nivel de carga cognitiva y el tiempo transcurrido en la tarea son altos, activar una alarma de alerta no tiene una efectividad del 100%; la reacción del motorista depende del nivel de somnolencia, fatiga, carga cognitiva y experiencia. Por lo tanto, dado que los sistemas de monitorización están centrados en clasificar o inferir el estado somnoliento del conductor en el tiempo t_0 , resulta de mayor utilidad poder estimar el estado del conductor para un intervalo ($t_0 \rightarrow t_{0+n}$); n -subsecuentes minutos. Un sistema de estas características puede alertar al conductor mientras tiene una adecuada capacidad de reacción; aún

no está bajo efectos de la somnolencia. En este sentido, un sistema de pronóstico o forecasting se refiere a encontrar un valor o estado para los subsecuentes minutos en base a un conjunto de datos históricos.

De esta manera, el algoritmo toma como entrada la secuencia $(\mathbf{x}_{t_0-m}, \dots, \mathbf{x}_{t_0-2}, \mathbf{x}_{t_0-1}, \mathbf{x}_{t_0})$ y estima $(\mathbf{y}_{t_0+1}, \mathbf{y}_{t_0+2}, \dots, \mathbf{y}_{t_0+n})$ valores, donde $m > n$; $m, n \in \mathbb{Z}$. Los algoritmos de pronóstico son utilizados en diversas áreas tales como la previsión de la demanda del mercado, pronóstico del precio de acciones bursátiles, estimación del stock de almacenes o del consumo de servicios como electricidad o agua [70–75]. En estas aplicaciones, los datos pueden ser periódicos, poseer una tendencia definida, ser estacionarios o asumen un proceso lineal en las variables temporales, lo cual facilita implementar algoritmos de pronóstico con métodos como Vector Autoregression (VAR) o Autoregressive Integrated Moving Average (ARIMA) [76, 77]. No obstante, las señales fisiológicas como EEG no son periódicas ni estacionarias, lo cual dificulta la implementación de un modelo de pronóstico basado en esta clase de algoritmos. Bajo estas condiciones, los algoritmos basados en Machine Learning han mostrado ser más eficaces en procesos de pronóstico. Los primeros modelos en utilizarse están basados en redes neuronales o ANN que aprenden de datos muestrales y capturan relaciones funcionales entre las muestras siempre que exista suficientes datos u observaciones de entrada [78, 79], donde la variabilidad en los datos facilita el aprendizaje del algoritmo.

Por consiguiente, las ANNs pueden inferir correctamente la parte no visible de los datos, incluso si contienen información moderadamente ruidosa. Como los modelos de pronóstico suponen una relación subyacente desconocida entre los datos históricos de la serie temporal (variables relevantes) y los datos pronosticados o valores futuros, las ANNs pueden inferir esa relación. No obstante, esta clase de algoritmos tienen una desventaja significativa al utilizarse con datos complejos; no son capaces de extraer representaciones significativas de largo plazo. Por otro lado, los modelos basados en CNNs son ampliamente utilizados porque el proceso de convolución se focaliza en regiones locales de los datos de entrada, capturando características efectivas de corta duración y fundamentales en series temporales [43, 80]. Además, sus diferentes capas permiten captar características profundas a corto plazo, aprendiendo representaciones jerárquicas. Sin embargo, las CNNs no pueden obtener características globales y complejas debido a su proceso de convolución, lo cual es más complicado incluso en datos temporales simples. En principio, esto genera una menor precisión con un mayor tiempo de procesamiento en el modelo implementado.

Por otra parte, los algoritmos basados en LSTMs son el punto más significativo en el estado del arte con modelos de pronóstico [42, 81, 82]. Sus principales ventajas radican en que tienen una mayor capacidad de extraer y clasificar características significativas de diferentes categorías de eventos. Tienen una capacidad de procesar secuencias de datos con dependencias de largo alcance, capturando información de varios pasos temporales anteriores y retenerla por más tiempo en base a series temporales. También pueden aceptar entradas con longitudes variables, lo cual es especialmente útil en modelos de pronóstico. Además, la estructura de celdas de las LSTMs aborda el problema del gradiente de fuga presente en las CNNs al evitar que su gradiente se vuelva demasiado pequeño durante el entrenamiento del modelo, lo cual permite una mayor capacidad de aprendizaje. Una mejora de las LSTMs es implementar una arquitectura Bidirectional Long Short Term Memory (BiLSTM) [42, 83], la cual tiene la capacidad de analizar los datos en dirección hacia adelante y luego hacia atrás (bidireccionales). Esto les da la capacidad de utilizar información que fluye desde una ventana temporal del pasado y de marcas temporales futuras (en esencia es una capa LSTM adicional que funciona en dirección inversa).

Un algoritmo BiLSTM divide la entrada en una secuencia de componentes que son tratados individualmente, generando una salida por cada componente presente en la secuencia para luego

generar la salida requerida según los datos históricos de entrada; su salida no se basa en la probabilidad de cada componente. No obstante, el inconveniente de estos algoritmos está en su compleja estructura, lo cual requiere mayor tiempo de entrenamiento frente a modelos CNNs con un mayor número de capas.

2.3. Discusión

En base a que la somnolencia en el conductor es una expresión de un proceso subyacente que tiene múltiples factores de origen, su detección es un proceso bastante complejo que puede requerir más de una tecnología de sensores a fin de captar el mayor número posible de expresiones de somnolencia [84–86]. Esto sucede dado la variabilidad interindividual e intraindividual del conductor, llevando a tener diferentes niveles de susceptibilidad a los factores de exposición. Por ello, a pesar del avance en equipos para obtener señales e imágenes fisiológicas, no existe un método único estandarizado que permita el diagnóstico y estimación global de la somnolencia y fatiga en el conductor. Por el contrario, la combinación de diferentes categorías de datos como la dinámica del vehículo (patrón de dirección, sistema de frenado, posición del vehículo en el carril), datos de expresión física y fisiológicos del conductor, obtenidos en tiempos prolongados de exposición, posibilitan un análisis más robusto y preciso. No obstante, los datos utilizados siempre deben estar en balance entre la eficacia, el número de sensores utilizados y el confort del conductor durante la ejecución de su tarea.

En este punto, las señales de EEG representan el estado del arte en la detección de somnolencia porque reflejan directamente la respuesta de la actividad neuronal del sujeto [87–89]. Además, los datos neurofisiológicos permiten estimar la cantidad de máximo esfuerzo cerebral (estrés mental) que el conductor puede soportar antes de haber una degradación en su rendimiento (desempeño). Por otro lado, las señales de EMG y ECG también son utilizadas para monitorear el estado del conductor [90–92]. Por ejemplo, una disminución de la variabilidad del ritmo cardíaco es generalmente un patrón asociado a estados de somnolencia que puede detectarse con una o más derivaciones durante experimentos de laboratorio o mientras se conduce en entornos reales. No obstante, algunos de estos sensores pueden resultar algo intrusivos e interferir con el confort del conductor durante la conducción. Adicionalmente, el uso de los datos del vehículo es un método ampliamente utilizado para detectar estados de somnolencia. Esto a pesar de que los sensores en el vehículo suelen verse afectados por condiciones atmosféricas, el entorno y condiciones del camino.

En este sentido, los sensores como el GPS, radar y sonar son menos precisos en entornos de carretera, que es donde la mayoría de los vehículos de carga pesada y de pasajeros circula. Por lo tanto, esta técnica es la menos precisa comparada con utilizar datos fisiológicos que no se ven afectados por condiciones externas al chofer. En este contexto, los datos del vehículo reflejan el estado general del conductor como consecuencia de la fatiga física, sobrecarga de trabajo, somnolencia o tiempo de acción en la tarea; no detectan señales directas de somnolencia en el sujeto. Por otra parte, los sensores FSR no son intrusivos ya que están implementados directamente en los pedales y el volante del vehículo [93, 94]. Tienen amplias aplicaciones en sistemas como robótica, estimación de peso, medición de fuerza de agarre, al igual que pueden utilizarse fácilmente para monitorear el confort de los pasajeros y el conductor. Finalmente, la mayoría de los trabajos suelen utilizar hasta dos tipos diferentes de sensores para monitorear el estado del conductor [10, 57, 95, 96], donde el número de sensores y su eficacia son factores clave a considerar.

Por otra parte, el algoritmo PERCLOS (Percentage of Eye Closure) es ampliamente utilizado en la detección y clasificación del estado del conductor en el tiempo (t_0) [101–103]. Este método

Tabla 2.1: Trabajos más relevantes orientas a la detección o clasificación del estado somnoliento en conductores de vehículos.

Ref.	Método	Enfoque	Descripción	Métrica
[97]	SVM KNN RFA	Clasificación	Señales de EEG: los algoritmos son entrenados con segmentos de las señales. En este trabajo también estiman la falta de atención del conductor durante su tarea. A pesar de su alta precisión, los datos fueron obtenidos de un experimento en simulador de conducción. Por lo tanto, la generalización del algoritmo no presentará el mismo valor de precisión reportado.	100 %
[98]	CNN	Clasificación	Señales de EEG: el algoritmo implementado es entrenado directamente con segmentos de las señales, pero únicamente detecta el estado del conductor mediante una categoría de señales.	90.4 %
[99]	SVM	Detección	Análisis de video: la duración del parpadeo fue utilizada como un parámetro de somnolencia mediante Eye Aspect Ratio (EAR). El algoritmo basado en SVM presenta los mejores resultados.	94.9 %
[100]	Tres modelos de Regresión Logística Multinomial	Detección	Señales de EEG, autoevaluaciones y medidas conductuales: el objetivo fue evaluar la efectividad de las medidas de evaluación conductual para predecir la somnolencia. Sin embargo, las evaluaciones subjetivas no son una herramienta independiente de la psicología y percepción del sujeto, lo que las hace menos precisas y eficientes en entornos reales. Este método imposibilita poder implementar un modelo escalable y generalizable. Esto es evidente cuando los autores reportan que el mejor valor de precisión (90.2 %) se da cuando utilizan las señales EEG para predecir el instante de somnolencia reportado por el usuario.	EEG: 90 %

La mayoría de trabajos están orientados a la detección del estado somnoliento del conductor. Este proceso se dispone de las manifestaciones de somnolencia (bostezar, pestañeo lento o frecuente, orientación del cuello, entre otros.

se basa en la monitorización permanente del comportamiento ocular del conductor, donde se mide el porcentaje de tiempo de cada minuto durante el cual los ojos están al menos un 80 % cerrados. Se considera que este parámetro es un indicador fiable de somnolencia y fatiga bajo condiciones específicas. Un valor mayor al 20 % indica una alta probabilidad de somnolencia o fatiga y valores menores indican que el conductor está alerta. Aunque esta técnica no necesita tener contacto físico con el conductor, si requiere un análisis permanente de imágenes o vídeos del rostro del conductor, particularmente de los ojos. Para ello se requiere de cámaras con alta definición o sistemas de visión por computador para capturar las imágenes, presentando dificultades en presencia de poca iluminación, el ángulo de la cámara, condiciones climáticas adversas y el uso de lentes oscuros. Además, se pueden presentar serias dificultades ante el uso de cascos para protección y movimientos bruscos o continuos de la cabeza, reduciendo la efectividad del sistema.

Tabla 2.2: Trabajos previos en la detección y predicción del estado somnoliento subsecuente en conductores de vehículos.

Ref.	Método	Enfoque	Descripción	Métrica
[104]	Método pNN50	Predicción	Señales de ECG: el trabajo consiste en utilizar la tasa de intervalos RR que tienen >50 milisegundos de diferencia respecto a un intervalo RR contiguo, calculada en relación a todos los intervalos RR evaluados anteriormente (pNN50). No obstante, un evento de microsueño es un estado avanzado de somnolencia, donde el conductor está en alto riesgo de sufrir un accidente.	30-seg.: 83 %
[105]	ANN	Detección Predicción	Señales de EEG, respiración, información del sujeto y datos del vehículo: entrenan los algoritmos con parte de los datos de 20 participantes, luego el modelo se adapta para el participante 21 y se prueba con los datos restantes de los demás participantes. Una entrada mayor a 15 minutos implicaría que en ese lapso no es posible predecir el estado somnoliento. Además, variar la longitud de los datos de entrada al modelo está en relación lineal con el factor de exposición: a mayor tiempo conduciendo, hay una mayor probabilidad de somnolencia en el sujeto.	Detección: 90 % Predicción: 6.15 min.
[11]	ANN	Detección Predicción	Señales de EEG, respiración, información del sujeto y datos del vehículo: los modelos utilizan una entrada con diferentes tipos de información y una variación de 1 a 30 minutos de longitud. El incremento de la precisión del modelo se da porque la entrada contienen más información histórica de la evolución somnolienta del conductor. En la detección, para el 80 % de los datos de prueba (todas las categorías de datos), el error obtenido es máximo la mitad (0.5) de un nivel de somnolencia. En la predicción, para la combinación de tiempo en conducción, información del participante y comportamiento del conductor, el error es de 2 minutos en el 95 % de los datos. Sin embargo, bajo estas condiciones, el modelo tendrá un error de ± 1.33 minutos para el mejor de los casos.	Detección: 80 % Predicción: 4.6 min.

El proceso de predicción del estado subsecuente (predecir el próximo evento somnoliento) dentro de un intervalo $[t_0, t_{0+n}]$ es más complejo que detectar o clasificar el estado del conductor en el instante t_0 . Para que el algoritmo sea preciso, se requiere evaluar su nivel de alerta en t_0 y que esta condición se mantenga o evolucione a un estado superior, sin que intervengan otras variables.

Un ejemplo claro de este problema se presenta al girar la cabeza para mirar a los espejos laterales o traseros, causando falsos positivos o negativos en la detección de eventos somnolientos. Por lo demás, esta clase de sistemas pueden plantear preocupaciones de privacidad, ya que realizan grabación y análisis continuo de las imágenes del rostro de los conductores. En relación a la detección de eventos somnolientos en un intervalo temporal ($t_0 \rightarrow t_{0+n}$), el análisis de la literatura científica revela que existe un menor número de trabajos enfocados en estimar el estado subsecuente del conductor.

En el experimento de [104], los autores utilizaron sensores ECG para predecir eventos subsecuentes de microsueño en conductores de vehículos. No obstante, estos eventos son un estado avanzado de somnolencia que dura al menos cuatro segundos con una pérdida involuntaria de consciencia, llevando al conductor a estar en alto riesgo de sufrir un accidente vehicular. Este estado se refleja fuertemente en las señales de EEG al variar el índice de las bandas teta y alfa asociado a fases de alerta-somnolencia. En señales de ECG hay una disminución de la variabilidad del ritmo cardíaco, lo cual puede validarse mediante algoritmos basados en el análisis de la onda R del ECG.

En las Tablas 2.1 y 2.2 se relacionan los trabajos más relevantes orientados a detectar, clasificar o predecir el estado somnoliento del conductor. En este sentido, los trabajos de [11, 105], utilizaron sensores ECG, Pulse Plethysmography (PPG) y respiración para detectar y predecir los estados somnolientos. Es claro resaltar que entre más tiempo pasa el participante en la ejecución de la tarea, mayor es la probabilidad de somnolencia en el sujeto. Así, a los cinco minutos de iniciado el experimento hay un mínimo número de participantes somnolientos, pero a los 60 minutos todos experimentaron somnolencia. Los autores emplearon un algoritmo basado en ANNs como predictor del tiempo de próximos eventos de somnolencia. En este punto, los autores utilizaron una variación ascendente de (1 a 30) minutos como longitud de los datos de entrada a este algoritmo. Sin embargo, se presenta una relación cuasilineal ya que a mayor longitud en los datos de entrada, el tiempo pronosticado también se incrementa. Así, reportan un tiempo de hasta 60 minutos y valores *accuracy* del 70 % con 1.97 minutos en el mejor de los casos para el 95 % de los datos de validación. Por lo tanto, una mayor longitud de los datos de entrada genera un incremento de la variabilidad de somnolencia en las variables de salida, incrementando proporcionalmente el tiempo de predicción.

MATERIALES Y MÉTODOS

3.1. Introducción

En la primera sección de este capítulo se describen los materiales y métodos implementados en el análisis de los datos utilizados en el desarrollo de la etapa preliminar de esta tesis. Los algoritmos implementados inicialmente predicen los eventos de frenada de emergencia. Por consiguiente, se utilizan los mismos métodos para el análisis de los datos y entrenamiento de los modelos para la predicción de eventos subsecuentes de somnolencia. Por ejemplo, el método de análisis en las señales de EEG, basado en Independent Component Analysis (ICA) para eliminar artefactos no deseados, es el mismo en las dos etapas de preprocesamiento de los datos. En este orden de ideas, el algoritmo utilizado para segmentar las señales de la fase preliminar también se reutilizó en la segmentación de los datos obtenidas del experimento. Por consiguiente, se hace necesario reportar en un mismo informe el análisis y resultados obtenidos en las dos etapas de esta investigación.

De esta manera, en este capítulo también se presentan los datos resultantes del experimento de laboratorio, métodos y algoritmos implementados en el desarrollo final de la investigación. Los datos utilizados en la etapa para la detección de la intención de frenada de emergencia, obtenidos del trabajo de [106], son completamente diferentes a los utilizados en la detección y predicción de eventos subsecuentes de somnolencia. Esta última base de datos se obtuvo mediante un experimento de conducción simulada, donde se indujo a los participantes a presentar somnolencia por desactivación mediante la variación de las condiciones experimentales.

3.2. Detección de la intención de frenada de emergencia

En esta sección del proyecto de investigación se busca detectar la intención de la frenada de emergencia. Para ello se utilizó un algoritmo de clasificación basado en redes neuronales convolucionales o CNN, entrenado con señales EEG sin transformar a imágenes en escala de grises. Este data-set se configuró como un tensor de matrices $2D$ (N -electrodos \times muestras), facilitando manipular bases de datos grandes. Por facilidad y recursos computacionales, se utilizó la plataforma Google Colaboratory (Colab) para entrenar el modelo implementado. El rendimiento del modelo también se evaluó al variar el número de electrodos de cada matiz, formando seis grupos (4, 8, 13, 18, 33, 59-electrodos). Además, para evaluar el efecto de la forma de los datos en el entrenamiento del modelo, este se entrenó con seis grupos diferentes de imágenes en escala de

grises obtenidas mediante el espectrograma de cada segmento de señal en cada uno de los grupos de electrodos.

3.2.1. Configuración experimental y materiales

Los datos de la fase preliminar de esta tesis, se obtuvieron de 18 participantes sanos (4 mujeres y 14 hombres) con edades entre 25 y 36 años, los cuales condujeron un automóvil virtual provisto de volante, pedal de acelerador y frenado [106]. La tarea de los sujetos consistía en conducir un coche virtual de carreras siguiendo a un automóvil líder a una velocidad de 100 km/h para mantener una distancia máxima de 20 metros entre ambos vehículos. Luego, en intervalos aleatorios de 20 y 40 segundos, el coche líder desaceleraba abruptamente desde los 100 km/h a una velocidad de entre 60 y 80 km/h. Este proceso activó las luces traseras del coche líder, constituyendo un estímulo para el participante en el simulador. Tres segundos más tarde, el vehículo líder aceleraba de nuevo hasta los 100 km/h para alcanzar la distancia máxima de 20 metros entre ambos vehículos.

Como consecuencia, el proceso de conducción, frenado y aceleración realizado por los participantes se desencadenó como un único evento de emergencia. Estos eventos se repitieron durante tres bloques de 45 minutos cada uno con períodos intermedios de descanso de 10 a 15 minutos. El estímulo al participante corresponde al parpadeo de las luces de frenado del vehículo líder, la respuesta está representada por la desviación del pedal del freno del vehículo del conductor (inmediatamente después del estímulo). Los participantes no tuvieron la oportunidad de cambiar al carril opuesto porque el software Open Racing Car Simulator (TORCS) estaba configurado para que hubiese tráfico vehicular en sentido contrario.

Los investigadores utilizaron 59 electrodos EEG bajo el sistema internacional 10-20 [55]. El simulador estaba configurado con tres monitores de 19 pulgadas y corría bajo el software TORCS con una configuración personalizada. Esta base de datos proporciona información técnica sobre la distancia entre vehículos, los tiempos de inicio del estímulo, la deflexión de los pedales del freno y acelerador y la aceleración de los vehículos. Los tiempos de respuesta de frenado se basaron en la primera deflexión perceptible por encima del nivel de ruido del pedal del freno después de un evento de frenado inducido por el vehículo líder. Tanto las señales de EEG como los datos técnicos proporcionados por el simulador se volvieron a muestrear a una frecuencia común de 200 Hz. Este conjunto de datos constituye un análisis multivariado multimodal por cada sujeto del estudio.

3.2.2. Métodos de preprocesamiento

Para preprocesar las señales de EEG se utilizó el toolbox EEGLAB en Matlab [107, 108], el cual proporciona un entorno gráfico interactivo (GUI) para procesar datos electrofisiológicos. En este toolbox, es posible realizar directamente un análisis tiempo frecuencia o mediante el método de ICA. Los datos de EEG se filtraron previamente para atenuar el ruido de la línea eléctrica en 50Hz. EEGLAB implementa funciones de filtrado a través de la función (filtfilt) de Matlab [108], la cual realiza un filtrado de fase cero, procesando los datos de entrada tanto en dirección directa como inversa. Además, los datos se sincronizaron en un conjunto estímulo-respuesta, estableciendo un análisis multivariable (MVA).

3.2.2.1. Análisis por componentes independientes

El análisis por componentes independientes ICA es un método estadístico para revelar artefactos ocultos que subyacen en las señales de EEG [60]. Esta técnica permite encontrar una representación lineal de las señales no gaussianas de forma que las componentes sean estadísticamente independientes o lo más independientes posible. Esto permite obtener la estructura fundamental de las señales para la extracción de características y la separación de sus componentes. Por lo tanto, es posible separar las fuentes independientes mezcladas linealmente en los electrodos de EEG como una serie temporal con amplitud en μV . En cada registro de EEG no sólo se codifica una mezcla de señales sincronizadas por compuestos neuronales en regiones cerebrales específicas, sino que también se mezclan señales externas como los artefactos de movimiento, latidos cardíacos, movimientos oculares y musculares. Como consecuencia, cada muestra $\mathbf{x}_i(\mathbf{t})$ de los datos de EEG representa señales mixtas, definidas como combinación de muchas fuentes independientes.

Por lo tanto, la separación de las señales mixtas se basa en el método de separación ciega de fuentes o Blind Source Separation (BSS), permitiendo separar las señales de origen aún cuando se tiene poca información de ellas. Considerando $\mathbf{X}_{2D} = [\mathbf{x}_1(\mathbf{t}), \mathbf{x}_2(\mathbf{t}), \dots, \mathbf{x}_n(\mathbf{t})]^T$ como un conjunto de datos $2D$ (*electrodes* \times *samples*), entonces $\mathbf{x}_i(\mathbf{t})$ representa el i^n electrodo del conjunto EEG que está determinado por la ecuación 3.1.

$$\mathbf{x}_i(\mathbf{t}) = \mathbf{a}_{(i,1)}s_1(\mathbf{t}) + \dots + \mathbf{a}_{(i,k)}s_k(\mathbf{t}) + \dots + \mathbf{a}_{(i,m)}s_m(\mathbf{t}) \quad (3.1)$$

Cada muestra de $\mathbf{x}_i(\mathbf{t})$ se define por el producto de una fuente s_k , $k = 1, 2, \dots, [m \text{ sources}]$ y un peso $\mathbf{a}_{i,k} = (\mathbf{a}_{1,k}, \dots, \mathbf{a}_{m,k})$ que corresponden a los elementos de la matriz de mezcla \mathbf{A} . Finalmente, las componentes independientes o Independent Components (ICs) de la señal están determinadas por las ecuaciones 3.2 y 3.3

$$\mathbf{X}_i(\mathbf{t}) = \mathbf{A}\mathbf{S}(\mathbf{t}) = \sum_{i=1}^m \mathbf{a}_{(i,k)}s_i(\mathbf{t}); \quad k = 1, \dots, m \quad (3.2)$$

$$\mathbf{S}(\mathbf{t}) = \mathbf{W}\mathbf{X}(\mathbf{t}) = [s_1(\mathbf{t}), s_2(\mathbf{t}), \dots, s_m(\mathbf{t})]^T \quad (3.3)$$

Donde los pesos $\mathbf{a}_{i,k}$ se desconocen, pero permiten transponer \mathbf{A} . Como resultado, $\mathbf{W}_{m \times n} = \mathbf{A}_{m \times n}^{-1}$ representa la separación matriz-ICA donde sus coeficientes $\mathbf{w}_{i,k}$ permiten separar las fuentes $\mathbf{S}_i(\mathbf{t}) = (s_1, \dots, s_m)^T$ como señales máximamente independientes. Por otro lado, en el análisis de estas señales, el toolbox EEGLAB permite detectar artefactos automáticamente en las señales, pero se requiere verificar los artefactos mediante inspección visual, análisis de tiempo y por mapas de distribución en las ICs. Finalmente, la eliminación de los artefactos en las señales de EEG se realizó manualmente, conservando así las ICs de activación neuronal de las señales con la menor cantidad posible de artefactos.

3.2.2.2. Segmentación de los datos

La detección de la intención de frenada de emergencia se configuró como un proceso de clasificación binaria entre las clases conducción normal (*no-objetivo*) y eventos de emergencia (*objetivo*). Por lo tanto, las señales fueron segmentadas en dos clases, donde la longitud de

los segmentos *objetivo* y *no-objetivo* de las señales EEG se definió en función de los datos técnicos obtenidos del simulador. En primer lugar, los segmentos *objetivo* se establecieron en 2000 milisegundos de longitud y se definieron como aquellos en los que el encendido de las luces de los frenos del vehículo delantero (estímulo) se daba no antes de 500 milisegundos ni después de 1500 milisegundos tras el inicio de este estímulo. En segundo lugar, los segmentos *no-objetivo*, con 2000 milisegundos de longitud, se definieron considerando un intervalo de 500 milisegundos entre cada segmento *no-objetivo* y al menos una longitud de 3000 milisegundos de diferencia desde el inicio de cada estímulo, Figura 3.1. De esta manera, los segmentos *objetivo* representan un estudio alineado con estímulos, permitiendo analizar la respuesta neuronal de los participantes ante una situación estresante como un eventual accidente automovilístico. Este procedimiento de segmentación resulta en una cantidad de segmentos *no-objetivo* mucho mayor que el conjunto de segmentos *objetivo*.

En consecuencia, se procede a equilibrar el conjunto de datos (muestras de las dos clases), considerando una cantidad menor de muestras *no-objetivo*, Tabla 3.1. Para equilibrar este conjunto de datos se empleó la técnica de random-undersampling, lo cual reduce una determinada cantidad de muestras en la clase *no-objetivo*. El proceso random-undersampling consiste en eliminar muestras de la clase que tiene más observaciones en comparación con las otras clases. Cada bloque en la clase *no-objetivo* tiene 91 ± 20 , mientras que la clase *objetivo* tiene 79 ± 12 segmentos por sujeto. Es decir, hay 79 ± 12 estímulos (eventos) de frenado de emergencia por bloque durante la conducción, mientras que el resto de la señal corresponde a conducción normal. El número de muestras por sujeto varía porque se rechazan segmentos que no cumplen con los requisitos del esquema de segmentación, Figura 3.2; los eventos de frenado de emergencia ocurren en intervalos de tiempo aleatorios.

De esta manera, en el conjunto de clases (*no-objetivo*) se eliminó un segmento antes y un segmento después del intervalo *objetivo*. Luego, de los segmentos restantes en el conjunto de clases *no-objetivo*, se eliminó segmentos aleatorios hasta que la diferencia entre las dos clases fue inferior al 1% (42,56 muestras). La Figura 3.1 y la Tabla 3.1 resumen los criterios de segmentación y datos de las clases.

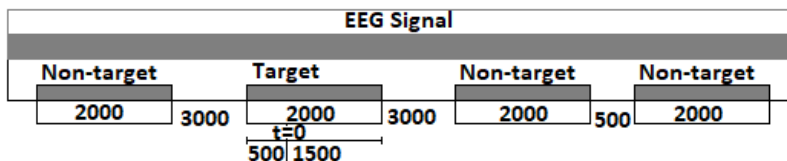


Figura 3.1: ($t = 0$) marca el inicio del estímulo. (i) La duración de los segmentos (clases) *objetivo* y *no-objetivo* es de 2000 milisegundos y están separados por más de 3 segundos del siguiente estímulo. (ii) Con el fin de obtener el mayor número de características, se consideró un intervalo de 500 milisegundos entre cada segmento *no-objetivo*.

Tabla 3.1: Número total de segmentos *objetivo* y *no-objetivo* según la longitud establecida en las señales para los tres bloques.

Name	Target	Non-target
Segments per block	79 ± 12	91 ± 20
Length (ms)	2000	2000
Total amount of balanced segments per block	4266	4296

(i) Se obtuvieron tantos segmentos *objetivo* (target) y *no-objetivo* (non-target) como fue posible a partir de las señales EEG. (ii) El tiempo de conducción normal es mucho mayor en comparación con el total de eventos en frenada de emergencia, por lo que se obtuvo un data-set desequilibrado.

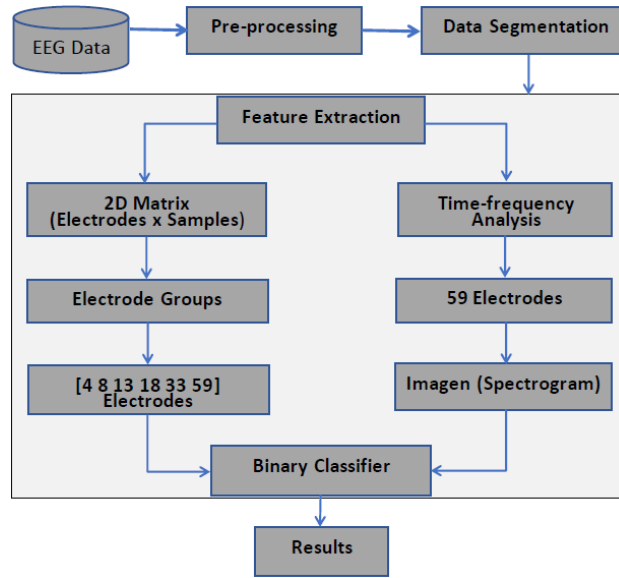


Figura 3.2: Diagrama de trabajo implementado en el análisis y procesamiento de los datos para esta sección del proyecto de investigación.

3.2.3. Etapa de procesamiento

El grupo de segmentos *objetivo* y *no-objetivo* se configuró en una disposición tensorial por matrices (*electrodos* \times *muestras*). El número de electrodos definió la altura de las matrices y la longitud de cada segmento corresponde al ancho de la matriz. Por lo tanto, hay un número de n matrices *objetivo* y un número de m matrices *no-objetivo* configuradas en un tensor. Este esquema se utilizó para entrenar el algoritmo implementado, clasificando la intención de frenada de emergencia y la conducción normal. Además, se hicieron diversas modificaciones en la altura de las matrices para evaluar la precisión del algoritmo con un menor número de electrodos. Por esta razón, se configuró seis grupos de electrodos en los que el más pequeño tiene solo cuatro electrodos (4, 8, 13, 18, 33, 59). También se comparó la precisión del modelo cuando este se entrenó utilizando las señales directamente y cuando se entrenó mediante imágenes en escala de grises obtenidas de las señales. Para esta comparación, cada segmento de los grupos de electrodos se convirtió en imágenes en escala de grises mediante su espectrograma, generando seis conjuntos diferentes de imágenes (un conjunto de imágenes por grupo de electrodos).

3.2.3.1. Análisis tiempo-frecuencia

El espectrograma es una representación visual en forma de imagen RGB o en escala de grises de la intensidad de una señal no estacionaria en la que su contenido de frecuencia varía en el tiempo [109]. Esta técnica utilizada para extracción de características se basa en el análisis del dominio tiempo-frecuencia y se realiza mediante la transformada de Fourier de tiempo corto o Short-Time Fourier Transform (STFT) [110], Ecuación 3.4. Cada segmento obtenido se transforma del dominio tiempo $\mathbf{x}(t) \in \mathbb{R}^{t \times N}$ al dominio tiempo-frecuencia $\mathbf{X}(\tau, \omega) \in \mathbb{R}^{\tau \times \omega}$, lo cual se interpreta como una matriz $\mathbf{X}(\tau, \omega)$. Dada la naturaleza de la señal, la STFT se calcula al deslizar una ventana de análisis $\mathbf{w}(n)$ de longitud M sobre la señal para calcular la Transformada Discreta de Fourier (DFT) de los datos en cada nueva ventana, Ecuación 3.5. Esta ventana recorre la señal $\mathbf{x}(t)$ a intervalos de R muestras, generando $L = M - R$ muestras de solapamiento entre cada segmento contiguo

de ventana. De este modo, $\mathbf{X}_{mth}(\tau, \omega) = [X_1(m, \omega) \ X_2(m, \omega) \ \dots \ X_k(m, \omega)]$ representa los valores (DFT) con datos de ventana centrados en el tiempo mR de la mth columna de la matriz STFT para cada segmento y (τ, ω) corresponde a los dominios de tiempo y frecuencia, respectivamente.

$$STFT(x(t)) \equiv X(\tau, \omega) = \int_{-\infty}^{\infty} x(t)w(t - \tau)e^{-j\omega t} dt \quad (3.4)$$

$$STFT(x[n]) \equiv X_m(m, \omega) = \sum_{n=-\infty}^{\infty} x[n]w[n - mR]e^{-j\omega n} \quad (3.5)$$

La matriz \mathbf{X} simboliza el espectrograma de cada segmento *objetivo* y *no-objetivo* y w representa la función de ventana deslizante cuyo tamaño está dado por M . Las filas de \mathbf{X} están indexadas por frecuencia, mientras que las columnas se indexan por tiempo, representando \mathbf{X}_m columnas de valores (DFT). En este trabajo, el espectrograma obtenido es una matriz bidimensional representada como una imagen en escala de grises.

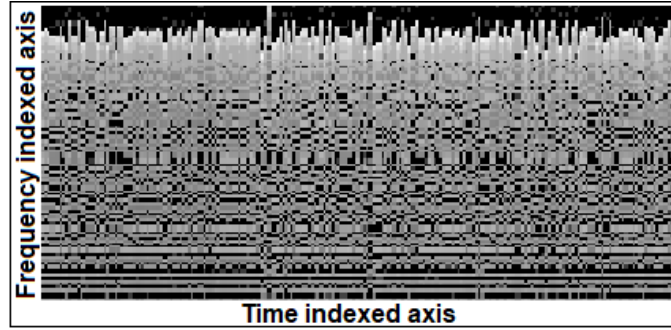


Figura 3.3: Resultado del espectrograma de un segmento *objetivo* donde el eje horizontal representa los cambios presentes en la señal en el dominio del tiempo. (i) La magnitud de cada imagen del espectrograma fue configurada en $200\text{-width} \times 100\text{-height}$ (pixels).

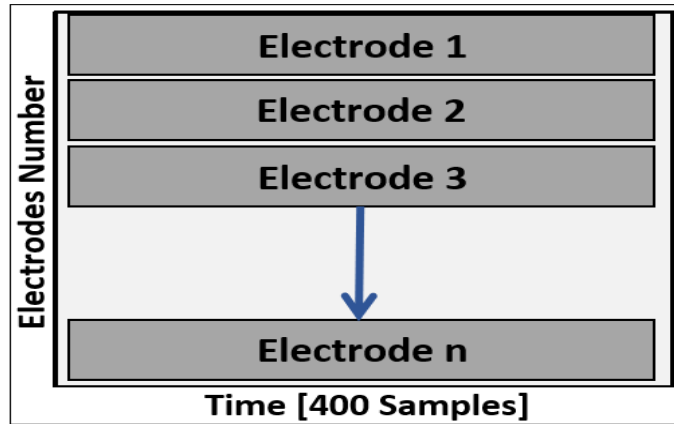


Figura 3.4: La longitud de cada segmento en las dos clases es de 2 segundos (400 *timesamples*). (ii) La altura de cada matriz varía de 4 a 59 electrodos según el grupo correspondiente.

Para generar el espectrograma de los segmentos, se estableció un tamaño de ventana de 25 milisegundos con una superposición del 75 %. Si se conoce un valor de frecuencia que se desea

analizar en una imagen, se puede utilizar dicho parámetro para definir el tamaño apropiado de la ventana. Sin embargo, más allá de este punto, sólo es posible ejecutar múltiples pruebas y observar los resultados de la imagen del espectrograma, dejando las que se considera más apropiadas. En consecuencia, se estableció un tamaño de ventana Hanning de 25 milisegundos equivalente a 40Hz puesto que está asociado a la banda gamma de 40Hz, la cual se relaciona con eventos de concentración y alerta máxima del conductor [59]. Además, a esta frecuencia hay una resolución tiempo-frecuencia óptima en la imagen del espectrograma. La superposición del 75 % intenta reducir el efecto de las funciones de ventana que afectan el proceso del espectrograma perdiendo información importante.

Finalmente, la magnitud de cada gráfico de espectrograma se configuró en 200×100 (*frecuencia* \times *tiempo*). La Figura 3.3 muestra el espectrograma en escala de grises de un segmento y la Figura 3.4 representa la variación de la matriz en $2D$ según el número de electrodos.

3.2.3.2. Estimación de los grupos de electrodos

La mayoría de los artículos analizados para este trabajo utilizaron al menos 10 electrodos en sus estudios [111–119]. Sin embargo, en el proceso de predicción de la intención de frenada de emergencia, la corteza visual o PVC, el área Frontal Cortex (FC) y la región motora o PMC son las más activas en la recepción de estímulos, la toma de decisiones y activación de las extremidades, respectivamente. Por este motivo, para analizar la precisión del clasificador, se consideró un número reducido de electrodos colocados en estas regiones cerebrales.

Por otro lado, los datos de entrenamiento del modelo implementado se configuran como un tensor matricial (N -dimensional) sin los datos de imágenes. Este procedimiento reduce tanto el tiempo de procesamiento como los recursos de computación utilizados al entrenar el algoritmo. La Tabla 3.2 resume el conjunto de electrodos más activos utilizados en cada una de las regiones cerebrales antes mencionadas, configurando un tensor de matrices con altura variable, Figura 3.4. Con el fin de evaluar el desempeño del modelo al entrenarse directamente con los segmentos de señales y mediante imágenes en escala de grises, estas se convirtieron a un formato de imagen común (*.png*), Figura 3.3. Se implementaron seis grupos de electrodos (4, 8, 13, 18, 33, 59), donde la altura de la matriz varía según el número de electrodos, pero su ancho es siempre de 400 muestras temporales (2 segundos).

3.2.3.3. Tiempo de reacción en frenada de emergencia

El tiempo de reacción (Braking Reaction Time (BRT)) mide el tiempo transcurrido desde que la señal de frenado del vehículo líder se activa hasta que el participante pisa el pedal del freno en el simulador [120]. El BRT representa la suma del tiempo de reacción, procesamiento cognitivo, decisión y activación muscular en cada evento de frenado de emergencia. Además, el BRT final calcula la diferencia entre el tiempo de respuesta (frenado del vehículo líder) y el tiempo de estímulo. Las condiciones establecidas en el experimento generaron diferentes niveles de estrés en los participantes que afectan el BRT final. Asimismo, el rendimiento de un conductor se ve fuertemente afectado por la somnolencia, fatiga física y mental que se acumula durante la ejecución de la tarea. La fatiga física aumenta los tiempos de reacción y la precisión de las respuestas. La fatiga cerebral también se manifiesta como un estado fisiológico que disminuye las capacidades de rendimiento cognitivo. Por lo tanto, el nivel de alerta del conductor se ve reducido por la fatiga (atención y toma de decisiones), resultando en un incremento de la probabilidad de un accidente vehicular con un mayor BRT.

Tabla 3.2: Configuración de los diferentes grupos de electrodos relacionados por región cerebral.

Electrode groups	Brain regions	Electrode names
4	PVC, PMC	O1, O2, C3, C4
8	PVC, PAA, PMC, Pfc	O1, O2, P3, P4, C3, C4, F3, F4
13	PVC, PMC, PmC	O1, Oz, O2, C3, C1, Cz, C2, C4 FC3, FC1, FCz, FC2, FC4
18	PVC, PAA, PMC, Pfc	O1, Oz, O2, P3, P1, Pz, P2, P4, C3 C1, Cz, C2, C4, F3, F1, Fz, F2, F4
33	PVC, VAC, PAA PMC, MAC, Pfc	O1, Oz, O2, PO3, POz, PO4, P3, P1 P2, P4, CP3, CP1, CPz, CP2, CP4, C3 C1, Cz, C2, C4, FC3, FC1, FCz, FC2 FC4, F3, F1, Fz, F2, F4, AF3, AF4, Pz
33	Occipital, Parietal Temporal, Frontal	O1, Oz, O2, PO7, PO3, POz, PO4, PO8, P9, P7 P5, P3, P1, Pz, P2, P4, P6, P8, P10, TP7, CP5 CP3, CP1, CPz, CP2, CP4, CP6, TP8, T7, C5 C3, C1, Cz, C2, C4, C6, T8, FT7, FC5, FC3 FC1, FCz, FC2, FC4, FC6, FT8, F7, F5, F3, F1 Fz, F2, F4, F6, F8, FP1, FP2, AF3, AF4

(i) Primary Visual Cortex (PVC), Visual Association Cortex (VAC), Posterior Association Area (PAA), Primary Motor Cortex (PMC), Motor Association Cortex (MAC), Premotor Cortex (PmC), Prefrontal Cortex (Pfc). (ii) Para el grupo con cuatro electrodos, se consideró las regiones PVC y PMC porque son las más activas al recibir un estímulo y en el movimiento de las extremidades. (iii) Se obtuvo un grupo de imágenes de cada grupo de electrodos como medio de comparación.

3.2.4. Algoritmo implementado

Como en esta sección del trabajo se requiere asignar una etiqueta a dos clases, se implementó un clasificador mediante la arquitectura de redes CNNs; se compone de múltiples capas ocultas distribuidas en la sección de extracción de características, incluso en la etapa de clasificación [121, 122]. En la extracción de características hay al menos una capa de convolución + *ReLU* y una de Max-pooling. Generalmente, la fase de clasificación consta de capas completamente conectadas y una función como softmax. La primera capa de neuronas convolucionales hace que la entrada \mathbf{x}_p convolucione con cada filtro \mathbf{f}_q^L , lo que da como resultado un mapa altamente característico que preserva las dependencias detectadas en los datos de entrada de la primera capa, Ecuación 3.6. En este trabajo σ es la función de activación lineal rectificadora (ReLU) definida como $\sigma(\mathbf{z}) = \mathbf{max}(\mathbf{0}, \mathbf{z})$, ampliamente utilizada en modelos basados en redes CNNs debido su mejor rendimiento en clasificaciones binarias.

$$\mathbf{y}_q^{(L)} = \sigma \left(\sum_{p=1}^{L(p-1)} \left[\mathbf{x}_p^{(L-1)} * \mathbf{w}_{pq}^{(L)} + \mathbf{b}_q^{(L)} \right] \right) \quad (3.6)$$

De esta manera, se diseñó un primer modelo de clasificación y se entrenó con el data-set de 59 electrodos. Este modelo se adaptó al resto de grupos de electrodos en los que se buscó las mejores métricas de evaluación al modificar los hiperparámetros del modelo. Sin embargo, para que sea posible una comparación de los resultados, el modelo entrenado en cada grupo de electrodos se entrenó también con imágenes en escala de grises, manteniendo los mismos hiperparámetros, Tablas

3.3 y 3.4. Como resultado, la modificación más recurrente ocurrió en la primera capa de la red, donde el modelo para el grupo de 13 electrodos tiene 45 filtros, mientras que el modelo para el grupo de 59 electrodos tiene 32 filtros. Esta modificación, el tamaño de la imagen, la cantidad de capas y los filtros en cada capa aumentan considerablemente el tiempo de entrenamiento de los modelos.

Tabla 3.3: Ciclos y tiempo de entrenamiento por cada grupo de electrodos y conjunto de imágenes en escala de grises.

Data		Epochs		Training time	
Electrode groups	Image groups	Electrode groups	Image groups	Electrode groups	Image groups
4	Set 1	90	90	40	167
8	Set 2	70	70	80	275
13	Set 3	100	100	300	312
18	Set 4	80	80	240	370
33	Set 5	80	80	240	455
59	Set 6	100	100	300	500

(i) El modelo en cada grupo de electrodos y set de imágenes tiene asociado los mismos parámetros de entrenamiento, facilitando una comparación de sus resultados. (ii) El tiempo de entrenamiento en el set-6 es mucho mayor debido a la cantidad de imágenes a procesar. (iii) La primera capa del modelo en el grupo de 13 electrodos tiene 45 filtros, lo cual puede explicar que tenga el mismo tiempo de entrenamiento que el modelo entrenado con el grupo de 59 electrodos que tiene menos filtros.

Tabla 3.4: Parámetros de configuración en el diseño del modelo básico para los grupos de electrodos.

EEG dataset	Conv. 32, (1x3)	Conv. 32, (3x3)	MaxPool. (2x2)	BatchNorma.	Conv. 32, (3x3)	Conv. 45, (3x3)	MaxPool. (2x2)	BatchNorma.	Conv. 75, (3x3)	BatchNorma.	Flatten ()	FC dense 32	FC dense 2	Output (0 - 1)
-------------	-----------------	-----------------	----------------	-------------	-----------------	-----------------	----------------	-------------	-----------------	-------------	------------	-------------	------------	----------------

(i) Modelo básico implementado para los grupos de 59 electrodos en datos con matrices e imágenes en escala de grises. (ii) El modelo utiliza un mínimo de 32 y no más de 75 filtros en las capas de convolución. (iii) El paso (stride) (1×3) en la primera capa de convolución permite que el modelo extraiga más características de los datos de entrada. (iv) Las capas de Batch Normalization y Max-Pooling reducen significativamente los recursos computacionales y el tiempo de entrenamiento de los modelos. (v) Las imágenes y el tensor de matrices son las entradas al modelo, mientras que su salida es $(0, 1)$.

3.3. Detección y pronóstico de eventos somnolientos

En esta etapa de la tesis se busca monitorear continuamente el estado del conductor y predecir el tiempo de ocurrencia de subsecuentes eventos de somnolencia dentro de una ventana temporal previamente definida. Esto se realiza mediante un sistema conformado por un algoritmo de detección de los eventos somnolientos en el tiempo t_0 y un algoritmo de pronóstico para la ventana $(t_0 \rightarrow t_{0+n})$. Los resultados de este trabajo permitieron la edición de tres artículos, donde los dos últimos comparten la misma base de datos y métodos de preprocesamiento de los data-sets. Adicionalmente, el método ICA, implementado en la sección de detección de la intención de frenada de emergencia, es el mismo que se utiliza en la segunda sección de esta investigación. Por lo tanto, los datos utilizados en este apartado son el resultado del experimento de laboratorio llevado a cabo con el fin de obtener una base de datos multisensor de conductores voluntarios en un simulador de conducción.

3.3.1. Diseño experimental y materiales

En el experimento de conducción virtual participaron 36 sujetos sanos (28 hombres, edad: $24,5 \pm 2$ y 8 mujeres, edad: $21,5 \pm 1$). El ensayo duró 105 minutos y se llevó a cabo en el Laboratorio de Gestión de Sistemas de Transporte (GESITRAN) [123]. Cada participante condujo durante cinco minutos para adaptarse al simulador y luego descansó por 10 minutos para evitar posibles molestias físicas (enfermedad del simulador). Antes de iniciar el experimento, cada participante debía leer y aprobar el formulario de consentimiento informado (2022-01-V3). Asimismo, todos los sujetos respondieron a una encuesta sobre su experiencia de conducción y su estado de salud. Los participantes tenían visión normal o corregida a normal, audición normal y no tenían antecedentes de enfermedades psiquiátricas o neurológicas que pudieran interferir con el experimento. De igual manera, a todos los participantes se les solicitó abstenerse de consumir café, alcohol y sustancias psicoactivas durante las últimas 24 horas previas al experimento. Es de notar que, del tamaño muestral ($n = 36$), los datos de 6 participantes tuvieron que ser descartados por el alto nivel de ruido en las señales de EEG, quedando con los datos finales de 30 participantes (23 hombres y 7 mujeres).

El entorno experimental contempló un simulador de conducción configurado con tres grandes pantallas de 32 pulgadas, una cámara IP y una sala de control independiente equipada con dos pantallas de 24 y una de 40 pulgadas. Esta sala es utilizada para vigilar continuamente, a través de la cámara, el desempeño y el estado físico de los participantes. Además, en cada pedal del simulador se ubicó un sensor FSR (5×5 centímetros) y el volante se cubrió con dos sensores FSR Flex ($45 \times 1,5$ centímetros). Estos sensores permiten detectar la actividad sobre los pedales del simulador y la presión de agarre del volante de dirección. También se ubicó un sensor de giroscopio y acelerómetro detrás del centro del volante, que registró la posición y velocidad de rotación del volante mientras cada participante conducía. Se implementó un módulo de bajo costo ESP32 en el simulador para registrar los datos de todos esos sensores. Este sistema envió todos los datos a través de su protocolo WiFi a un segundo módulo ESP32 situado en la sala de control, donde se almacenaron los datos de todos los sensores en una tarjeta SD, Figura 3.5.

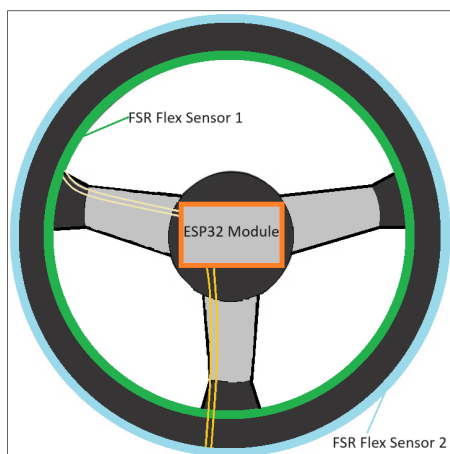


Figura 3.5: Los sensores implementados en el volante cubren las secciones interna y externa de este. (i) Los sensores de giroscopio y acelerómetro se colocaron detrás de la parte central del volante. (ii) El módulo ESP32, situado en la sección posterior al centro del volante, envía datos de los sensores ubicados en los pedales del simulador y volante a un segundo ESP32 ubicado en la sala de control.

Por otro lado, se utilizó el sistema g.Nautilus Pro de g.tec para registrar las señales de EEG con un total de 16 electrodos, según el sistema internacional 10-20, Figura 3.6. Las señales de ECG



Figura 3.6: Se utilizó la cámara IP para monitorear el estado del participante y una cámara-USB para registrar el comportamiento facial de los voluntarios mientras conducían.

y EMG se adquirieron utilizando el equipo Biopac MP35. Todos estos datos se registraron a una frecuencia de muestreo de 500Hz. En los datos de ECG se implementó la *II* derivación utilizando tres electrodos conectados a la pared torácica de acuerdo con la configuración estándar de tres derivaciones. Los electrodos EMG se colocaron sobre los músculos tibia anterior (Tibialis Anterior (TA)) y gastrocnemio (Gastrocnemius (Gn)) de la pierna derecha. Los equipos g.Nutilus, Biopac y FSR se sincronizaron mediante el segundo módulo ESP32 ubicado en la sala de control. Los algoritmos de grabación y control de los sistemas ESP32 se desarrollaron utilizando Arduino IDE 2.2.1. Además, el sistema de control del simulador entregó los datos de Telemetría (datos técnicos del vehículo). Este data-set representa indicadores indirectos de rendimiento y comportamiento de conducción y se compone de lo siguiente: coordenadas (x, y, z) de la posición y velocidad del vehículo, ángulo y velocidad del volante, posición de los pedales del acelerador, embrague y frenos, distancia respecto a la línea media y lateral del vehículo en la carretera.

3.3.2. Protocolo del experimento

El experimento fue diseñado para inducir somnolencia por desactivación en los participantes mientras conducían en tres diferentes escenarios virtuales. Los voluntarios condujeron durante un período de al menos 85 minutos, Figura 3.7. Los participantes podían abandonar el experimento en cualquier momento si así lo deseaban.

Adjustment Period Test Circuit	Work Block - 1 Circuit → Obstacles	Work Block - 2 Interurban Road → No Vehicles	Work Block - 3 Highway → No Vehicles
5	From 20 to 30	30	30
From 85 to 95 Minutes			

Figura 3.7: (i) En el circuito de prueba, los participantes tenían libertad de conducir a la velocidad que quisieran, buscando que el voluntario se adapte al simulador. (ii) Este paso evita una carga de trabajo adicional en el conductor durante el primer bloque por no estar adaptados al simulador.

- **Bloque 1:** En este primer bloque, el escenario se configuró como un circuito compuesto por dos

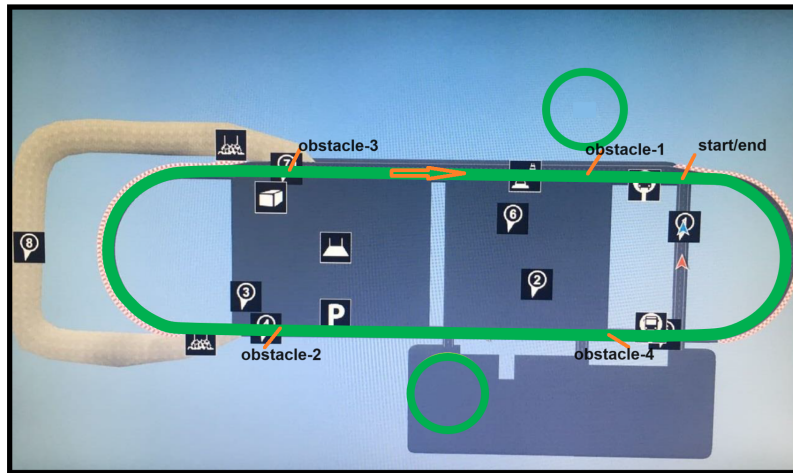


Figura 3.8: (i) En el escenario del circuito virtual para el bloque de trabajo uno se dispone de cuatro obstáculos que aparecieron en diferente orden y distancia. (ii) En cada una de las vueltas, el obstáculo aparece a una distancia mínima necesaria para que fuera visible por el participante en el vehículo virtual y que este pueda ser esquivado a la velocidad establecida.

semicírculos iguales con un radio de 80 metros y dos rectas paralelas, cada una de 250 metros de largo. Además, se agregó cuatro obstáculos diferentes al escenario, ubicados en diferentes puntos del circuito, Figura 3.8. Los participantes tuvieron que completar cuatro vueltas por obstáculo, 16 vueltas en total y aparecieron en el siguiente orden y distancia: 1) conos de tráfico a 10 metros y aparecieron en la primera vuelta. 2) un peatón a 13 metros y fue visible en el quinto giro. 3) un perro a 17 metros y apareció en el noveno giro. 4) un caballo a 22 metros y apareció en la decimotercera vuelta. Cada cuatro vueltas aparecían un nuevo obstáculo, mientras que los anteriores permanecían dentro del circuito. Como resultado, los primeros cuatro giros tuvieron un solo obstáculo, pero desde la decimotercera vuelta hubo cuatro obstáculos. Además, para cada obstáculo se estableció una velocidad diferente que se debía mantener por cuatro vueltas, empezando por 60km/h para los conos, 70km/h para el peatón, 80km/h para el perro y 90km/h para el último obstáculo. La solicitud de aumentar la velocidad se dio inmediatamente terminó la cuarta vuelta en su respectivo obstáculo, garantizando que este se esquite a la velocidad indicada.

- **Bloque 2:** En este bloque se configuró un escenario interurbano asfaltado de 15 kilómetros con un carril por sentido vehicular. La ruta presentaba 20 curvas horizontales y rectas de entre 25 y 1000 metros de longitud. Esta calzada virtual se configuró sin tráfico de coches, buscando que fuera el primer bloque de conducción monótona para inducir somnolencia. Los participantes tuvieron que conducir durante 29 minutos a una velocidad constante de 30 km/h, haciendo que los participantes pudieran experimentar somnolencia por desactivación de la carga de trabajo. Se esperaba obtener varios eventos somnolientos después de los 15 minutos de iniciado este bloque debido al efecto de la primera sección en los participantes.
- **Bloque 3:** Se configuró el bloque tres como un escenario que simulaba un recorrido de 14 km por autopista. En este trayecto se dispuso dos carriles por sentido, 15 curvas horizontales y rectas de al menos 50 metros de longitud. Los participantes tuvieron que conducir durante 29 minutos a una velocidad constante de 30 km/h sin encontrar tráfico en ninguna dirección. Esta etapa del experimento permitió a los participantes experimentar varios eventos de somnolencia debido a la mínima carga de trabajo involucrada y su conducción monótona. Se esperaba encontrar eventos de somnolencia aproximadamente a los ocho minutos de iniciado este bloque,

los cuales serían visibles en los vídeos y señales fisiológicas del participante.

3.3.3. Preprocesamiento de los datos

Las señales de EEG, ECG y EMG fueron filtradas para atenuar la interferencia de línea eléctrica utilizando un filtro notch a 50 Hz. Además, en las señales de EEG y ECG se aplicó un filtro pasa banda de 0,01 a 60 Hz para reducir el ruido de alta frecuencia. Luego, se analizaron las señales de EEG mediante el algoritmo ICA [60], reduciendo significativamente el ruido por artefactos de movimiento, movimientos oculares, cardíacos y musculares [124]. La definición de este algoritmo, el proceso de aplicación y eliminación de las componentes ICs se describe en la sección 3.2.2.1 correspondiente a la detección de la intención de frenada de emergencia.

3.3.3.1. Nivel real de somnolencia (ground truth)

El estado fundamental (estado de referencia) de somnolencia corresponde al nivel real de somnolencia del conductor. No obstante, no se puede acceder directamente al estado del conductor para clasificar su nivel de somnolencia y alerta. Por ello, en este trabajo se evalúa el nivel de somnolencia mediante un método de análisis de vídeo, el cual muestra las expresiones faciales de los participantes. Este vídeo se obtuvo mediante una cámara con resolución 2K ubicada frente a cada participante durante la etapa experimental. Desafortunadamente, durante el experimento, los voluntarios no miraban constantemente las pantallas del simulador, obstruían la lente de la cámara o cubrían su rostro con las manos. Por tanto, no fue posible grabar permanentemente el rostro de los participantes. Debido a esto, hubo segmentos de vídeo muy borrosos o blurry short-videos (BS-V) donde no fue posible aplicar el algoritmo de análisis (tal como ocurre en la realidad).

Por lo tanto, al considerar que las señales y videos están sincronizados, se rechazó estos segmentos BS-V de las dos categorías de datos, obteniendo una base de datos con un menor tamaño, pero con segmentos de vídeo limpios, los cuales se segmentaron en lotes de un minuto de duración al igual que las señales. El algoritmo de análisis de vídeo evalúa las expresiones individuales de somnolencia, como (80 %) del cierre parcial de los párpados y el tiempo transcurrido de parpadeo o Blink Elapsed Time (BET). Esto se hace mediante el método de medición del porcentaje de cierre ocular o PERcentage of eyelid CLOSure (PERCLOS) [101–103, 125]. Este es el método más utilizado, siempre que la calidad del vídeo sea lo suficientemente buena. La figura 3.9 representa los puntos faciales de referencia utilizados en el método PERCLOS. Adicionalmente, la Figura 3.10 representa una muestra del conteo de parpadeos al igual que se estima la postura de la cabeza como señal adicional de somnolencia.

Al analizar las expresiones faciales es posible implementar un procedimiento continuo de evaluación de la somnolencia del conductor. Por lo tanto, este método puede generar los niveles de somnolencia que se estime convenientes. En base a este método, se implementó la escala **DLs** que comprende únicamente tres niveles, Tabla 3.5: un nivel de alerta, un nivel de somnolencia leve y un nivel de somnolencia máxima. Estos niveles están definidos en base a la escala de KSS [45], la cual define nueve niveles diferentes: 1 para estados plenamente conscientes, 9 para una fase de mucho sueño. No obstante, el método KSS no es una medida continua del estado del conductor y se basa en la autoevaluación del sujeto evaluado, asociando una percepción individual del conductor según su experiencia y psicología. Como resultado, la escala DLs contempla únicamente un estado por cada tres niveles del método KSS.

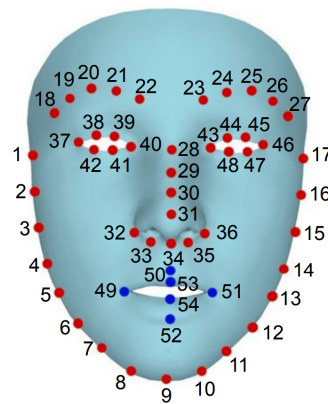


Figura 3.9: Se utilizaron los puntos de referencia 37 a 42 para estimar la relación de aspecto ocular en el conteo de parpadeos [126].

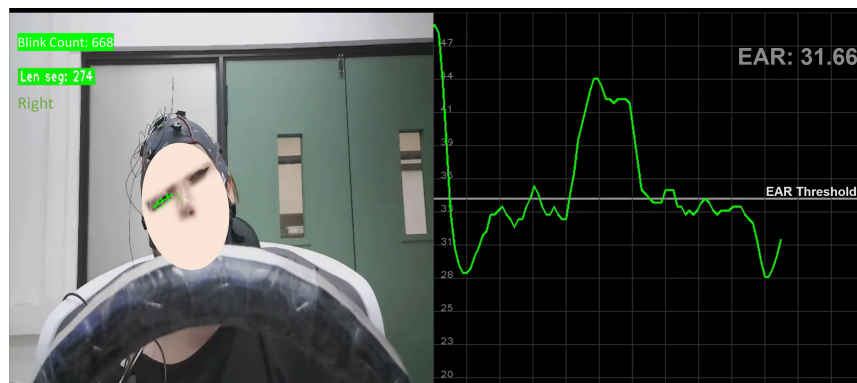


Figura 3.10: (i) El estado de este participante es somnoliento y tienen inclinado la cabeza a la derecha. (ii) Se considera que hay un parpadeo cuando el parámetro EAR (Eye Aspect Ratio) es menor que el valor fijado (34) de EAR Threshold. (iii) La línea en color verde representa el comportamiento del estado ocular o variación del parámetro EAR.

- Tiempo transcurrido de parpadeo (BET): Indica cuánto tiempo dura un parpadeo. Según [127, 128], la duración media del parpadeo es de unos 240 milisegundos y oscila entre 170 y 450 milisegundos para conductores en estado de alerta en escenarios virtuales. Si el cierre del párpado dura más de 500 milisegundos y el párpado está ocluido por más del 80 %, generalmente se define como eventos de somnolencia o fatiga [127, 128]. En este trabajo se consideró que los ojos están cerrados sólo cuando el párpado ocluye el 80 % de la pupila. En la Tabla 3.5, el parámetro BET es una referencia del tiempo de parpadeo según el análisis del estado del arte, pero se mide indirectamente mediante el método PERCLOS.
- Escala de somnolencia (DLs): Como se consideró solo dos niveles de somnolencia y un nivel de alerta según la escala KSS [45], los niveles extremadamente alerta, muy alerta y alerta de KSS se representan como nivel de alerta en el método DLs. Además, el primer nivel de somnolencia (drowsiness-1) representa el cambio de estado del conductor de alerta a somnoliento, donde su rendimiento comienza a deteriorarse.
- PERCLOS: Representa el porcentaje de cierre de los párpados con al menos 80 % de oclusión de la pupila [129], el cual se utiliza como indicador del nivel de alerta [101]. Se emplea esta métrica para indicar cuánto tiempo están cerrados los párpados en un minuto. El umbral de 0,20 (20 % de un minuto) indica el primer nivel de la escala DLs. Este método se basó en la puntuación del método EAR que evalúa si los ojos están cerrados.

Tabla 3.5: Técnica de medición subjetiva y descripción del método de medición continua diseñado para este trabajo.

Karolinska Sleepiness Scale		Custom Continuous Measurement Method		
Description	Level	PERCLOS (P %)	Blink Elapsed Time (BET[ms])	Drowsiness Scale
Extremely Alert	1			
Very Alert	2	$0 \leq P \% < 0.20$	$90 \leq BET < 500$	Alertness
Alert	3			
Rather Alert	4			
Neither Alert nor Sleep	5	$0.20 \leq P \% < 0.30$	$500 \geq BET < 800$	Drowsiness-1
Some Signs of Sleepiness	6			
Sleepy, But no Effort to Keep Alert	7			
Sleepy, Some Effort to Keep Alert	8	$P \% \geq 0.30$	$BET \geq 800$	Drowsiness-2
Extremely Sleepy, Fighting Sleep	9			

(i) El método KSS tiene nueve niveles, mientras que en la métrica de somnolencia propuesta hay tres. (ii) En un método de medición subjetivo es más complejo distinguir entre niveles de alerta extrema y niveles de muy alerta según el método KSS. (iii) El parámetro Blink Elapsed Time (BET) no se mide individualmente, sino que resulta del método PERCLOS y se registra en esta Tabla como dato de referencia.

Después de clasificar cada minuto (lote) dentro de la escala DLs, estos se dividieron en segmentos de cuatro segundos con un traslape de tres segundos, conformando la base de datos definitiva (original). Finalmente, se dividió esta base de datos en subconjuntos (70 % para entrenamiento, 20 % para validación y 10 % para probar la capacidad de los modelos al trabajar de manera eficiente con datos no utilizados). En comparación con otros estudios [11, 21], que extraen datos de características en la etapa de preprocesamiento, en este trabajo se entrenan los modelos implementados directamente con los segmentos sin transformar las señales a imágenes en escala de grises.

Como resultado de la segmentación final, se obtiene un conjunto de data-sets que no están balanceados. La clase “alertness” tiene la mayoría de las muestras, seguido de la clase “drowsiness-2” que constituye aproximadamente el 12,5 % del total de muestras en cada data-set. Entrenar un modelo de regresión o clasificación con un conjunto de datos desequilibrado puede hacer que el modelo aprenda de la cantidad de datos en lugar de las características intrínsecas de los datos. Para equilibrar los data-sets, se consideró hacer oversampling y undersampling mediante los métodos Synthetic Minority Over-Sampling Technique (SMOTE) y Under-Sampling Technique (Tomek-Links). Estos son los métodos de generación sintética y eliminación de datos más populares y precisos [130–132]. El concepto Tomek-Links es una modificación de la técnica undersampling condensado mediante el vecino más cercano o Condensed Nearest Neighbors Undersampling (CNNU) desarrollada por [131].

El método SMOTE utiliza la distancia (normalmente distancia euclidiana) de cada punto de datos y los vecinos minoritarios más cercanos de la clase minoritaria en lugar de la duplicación aleatoria de muestras. Esto genera nuevas muestras sintéticas sin tener duplicación en su espacio de características basado en un cierto porcentaje de la clase minoritaria. El método Tomek-Links fue diseñado originalmente para dos clases diferentes (una mayoritaria y otra minoritaria), es decir \mathbf{x}_a y \mathbf{x}_b , respectivamente. La distancia entre estas clases es $d(\mathbf{x}_a, \mathbf{x}_b)$ y es conocido como el enlace de Tomek, siempre que no haya otra clase \mathbf{x}_z tal que $d(\mathbf{x}_a, \mathbf{x}_z) < d(\mathbf{x}_a, \mathbf{x}_b)$ o $d(\mathbf{x}_b, \mathbf{x}_z) < d(\mathbf{x}_a, \mathbf{x}_b)$ [133]. Es decir, este método encuentra las muestras de la clase mayoritaria \mathbf{x}_a con la distancia euclidiana más corta a los datos de la clase minoritaria \mathbf{x}_b . La distancia se define como $d(\mathbf{x}_a, \mathbf{x}_b)$ y permite definir qué puntos están más cerca de la clase minoritaria para

eliminarlos de la clase mayoritaria [134]. El algoritmo SMOTETomek, introducido por [132], es un generador de datos sintéticos para la clase minoritaria y un eliminador de datos en la clase mayoritaria (SMOTE y Tomek-Links, respectivamente). SMOTETomek es una combinación de oversampling y undersampling que balancea conjuntos de datos desequilibrados.

Para balancear los data-sets, se aplicó el método antes definido para incrementar el número de muestras de las clases “drowsiness-1” y “drowsiness-2”, buscando que las tres clases tengan el mismo número de muestras. Por otro lado, para construir una segunda base de datos, también se redujo el número de muestras de la clase “alertness” y se generaron nuevas muestras sintéticas en la clase “drowsiness-2”, buscando que tuvieran el mismo número de muestras que la clase “drowsiness-1”. Con este procedimiento se obtuvieron dos bases de datos balanceadas que difieren en un 50 % en el número de muestras de sus clases en comparación con la base de datos uno. La Figura 3.11 corresponde a los segmentos de electrodos EEG obtenidos al balancear la base de datos mediante el método SMOTE. La Figura 3.12 representa la matriz de correlación de un segmento entre las tres clases (alertness, drowsiness-1, drowsiness-2) para el electrodo FP2, donde el valor de correlación máximo (0.16) indica que no hay una fuerza de relación lineal significativa entre dos o más clases.

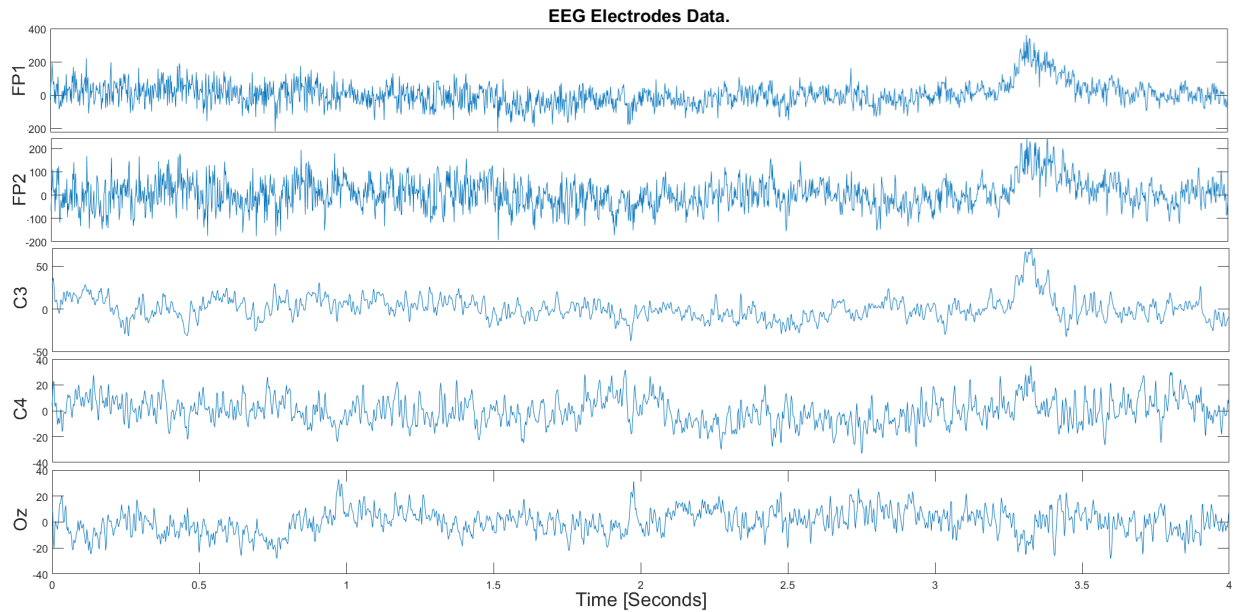


Figura 3.11: Señales de segmentos en diferentes electrodos de EEG con datos sintéticos, los cuales se diferencian significativamente entre cada segmento.

Adicionalmente, la Figura 3.13 corresponde a la matriz de correlación de los datos completos del electrodo Oz entre los tres bloques. Se evidencia que no existe una fuerza de relación significativa mayor o menor a (0.02) entre las combinaciones de los bloques. Además, entre los bloques dos y tres no existe ningún grado de relación, esto considerando que en el bloque tres se obtuvo una mayor cantidad de muestras de la clase “drowsiness-2”. La Figura 3.14 demuestra el grado de similitud entre los datos obtenidos para el electrodo Oz entre los tres bloques. Por otro lado, la Figura 3.15 muestra los valores de correlación para un segmento del electrodo Oz entre las tres clases. El máximo valor de relación es 0.21 entre los datos de las clases “alertness” y “drowsiness-1”. En cuanto a los valores de correlación cruzada de la Figura 3.16 para este mismo segmento, se evidencia que existe un valor de máxima similitud (0.36) localizado a los 0.018 segundos entre las clases “alertness” y “drowsiness-2”. Estos resultados demuestran que las muestras sintéticas generadas por el algoritmo SMOTE y los datos originales se pueden diferenciar completamente de una clase a otra.

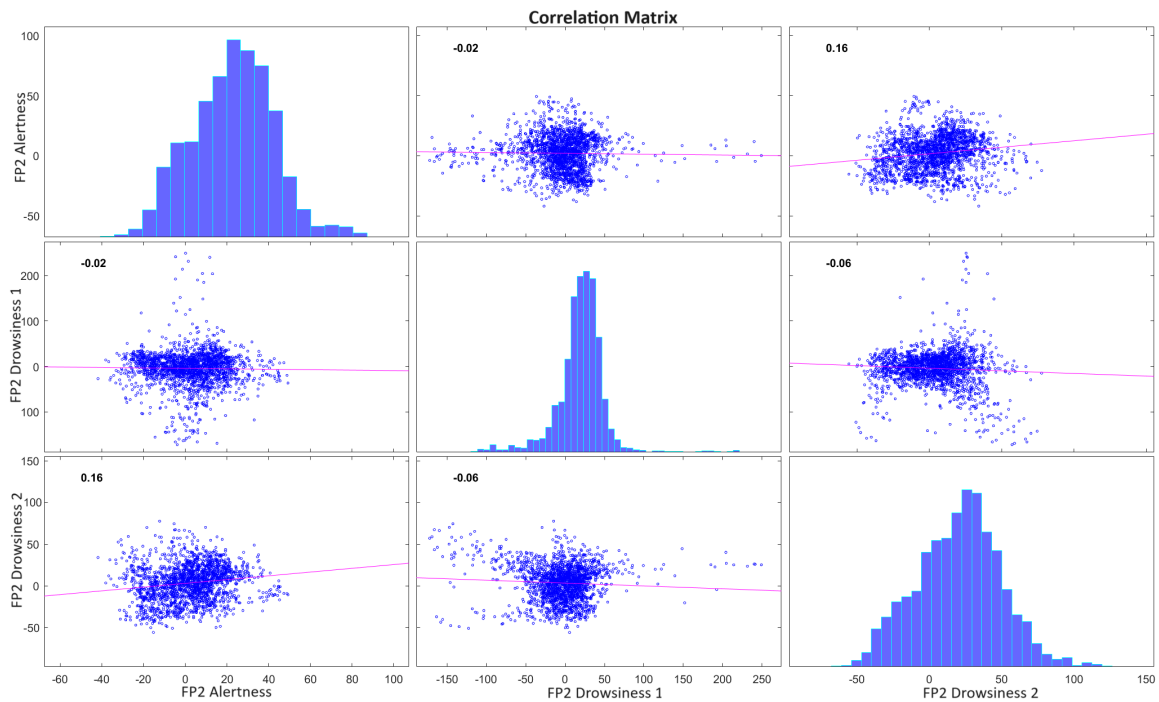


Figura 3.12: Matriz de correlación entre las clases para el electrodo FP2, donde también se aprecia diferencia significativa entre las tres clases. (i) Se demuestra que el método SMOTE implementado para generar datos sintéticos sobre las clases “drowsiness-1” y “drowsiness-2” funciona adecuadamente.

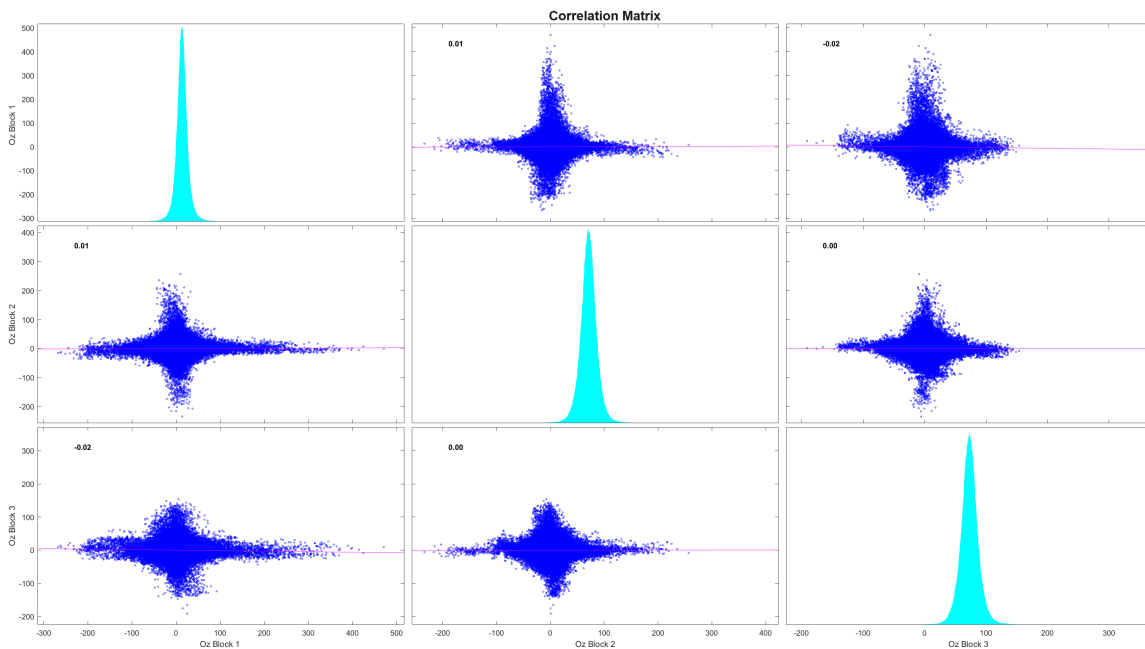


Figura 3.13: Los resultados de la matriz de correlación para para los datos del electrodo Oz entre los tres bloque completos muestran una fuerza de relación mínimamente representativa de -0.02 entre los bloques tres y uno y de 0.00 entre los bloque dos y tres.

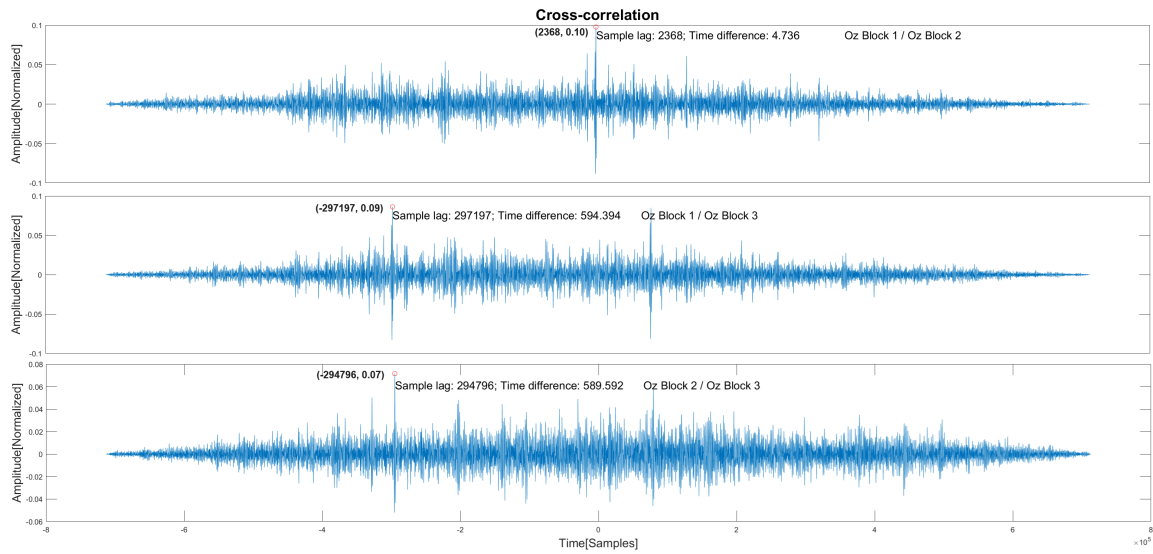


Figura 3.14: Estos resultados muestran el grado de similitud entre los datos de los bloques completos, donde el valor máximo de relación es 0.07 entre el bloque dos y el bloque tres.

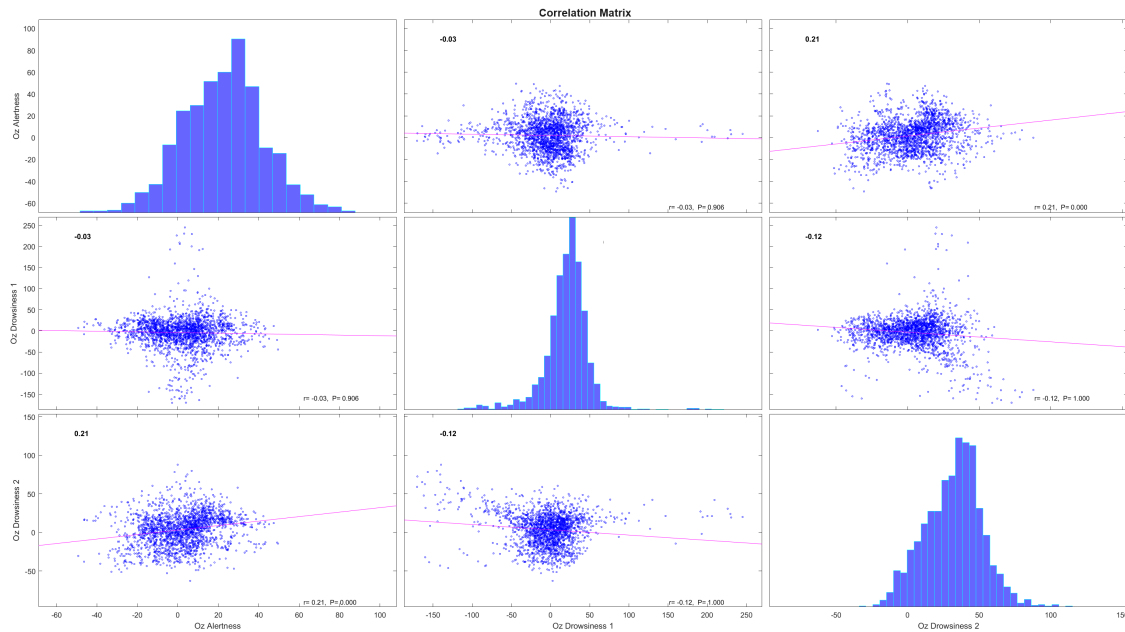


Figura 3.15: Los resultados de la matriz de correlación para un segmento del electrodo Oz en el método SMOTE muestran una fuerza de relación máxima de 0.21 entre las clases “alerteness” y “drowsiness-2”.

3.3.4. Método 1: algoritmos de clasificación y pronóstico

En esta etapa del trabajo, el primer método implementado está basado en la combinación de un modelo de clasificación en el tiempo t_0 y un modelo de pronóstico para estimar el estado somnoliento durante la ventana temporal ($t_0 \rightarrow t_0+n$). Los dos modelos utilizan directamente los segmentos de las señales; el algoritmo de pronóstico genera una señal de salida con la misma dimensión que la señal de entrada.

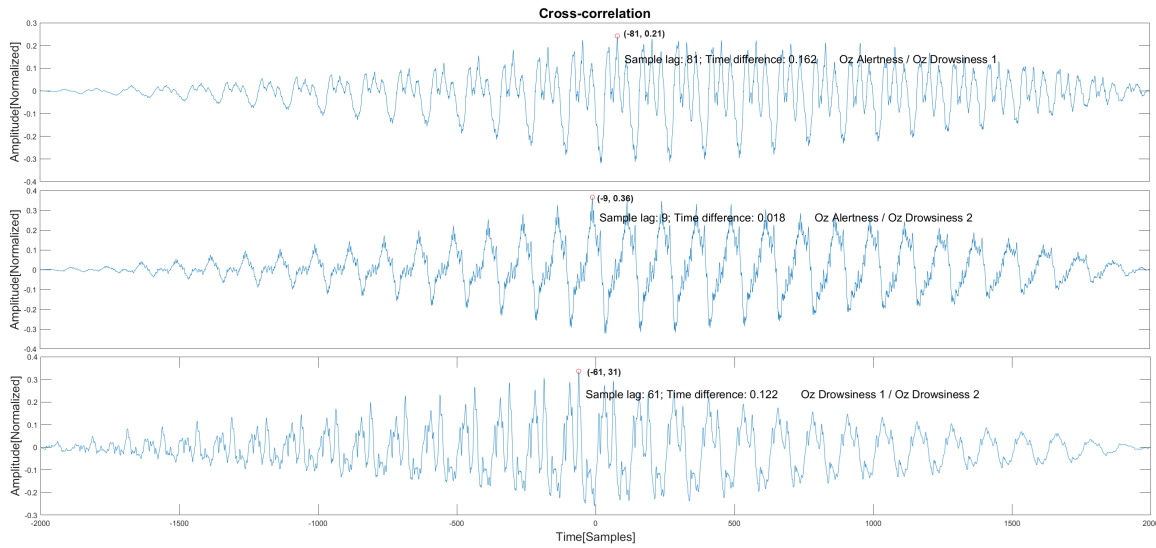


Figura 3.16: La correlación cruzada destaca que el valor de máxima similitud del segmentos entre las clases es de 0.36 entre las clases “alertness” y “drowsiness-1”.

3.3.4.1. Modelo de clasificación

Para determinar el estado del conductor en el tiempo t_0 , se desarrolló un modelo de clasificación que asigna una etiqueta a cada segmento de los datos multimodales de EEG, ECG, EMG, FSR y Telemetría (datos de medición del vehículo). Es decir, cada segmento de cuatro segundos se asocia a una de las tres etiquetas definidas como “alertnes”, “drowsiness-1”, “drowsiness-2” o (0, 1, 2), respectivamente. Este modelo está basado en redes CNNs mediante una técnica de redes residuales o Residual Network (ResNet), Figura 3.17. Una capa de una CNN aplica una gran cantidad de procedimientos de convolución en un enfoque de avance (feedforward) para trabajar de manera eficiente. Sin embargo, en una implementación tradicional con numerosas capas de profundidad, su principal desventaja es que enfrenta el problema del gradiente de fuga (vanishing gradient problem). Es decir, la derivada parcial de la salida de una función con respecto a sus entradas se vuelve cero, impidiendo que se mida el cambio de la salida de una función con respecto a los cambios ocurridos en sus entradas. Si el gradiente descendente se vuelve muy pequeño, impide que la red aprenda durante la etapa de retro-propagación del error (backpropagation) [135], haciendo que el aprendizaje sea menos eficiente.

Como se dispone de una base de datos multi-sensor, se implementa una etapa de submodelos para cada modalidad de datos, los cuales se entrenan directamente con los respectivos segmentos de las señales, Figura 3.17. De esta manera, el modelo implementado concatena los feature maps en diferentes niveles según la categoría de datos de entrada. Esto permite obtener un mapa de características general que es la entrada a la etapa de clasificación. Los mapas de las señales fisiológicas son concatenados en una primera etapa dado que comparten la misma frecuencia de muestreo, resultando en mapas con la misma dimensión. Tras una modificación del mapa de FSR se concatena con el mapa de características fisiológico y este resultado con el mapa de los datos de Telemetría, resultando en un mapa de características global (GFM). De acuerdo a la Figura 3.17, en la sección (feature extractor), cada etapa ResNet-N está provista únicamente de la fase de extracción de características, no hay etapa de clasificación.

No obstante, para verificar el efecto de la modalidad de los datos utilizados, cada submodelo

ResNet-N también se complementa con una etapa de clasificación y se entrena como un modelo independiente de una única categoría de datos. Tanto el modelo basado en los GFM y los modelos de comparación entregan una salida en el formato $[0, 1 \text{ o } 2]$ para cada segmento de entrada.

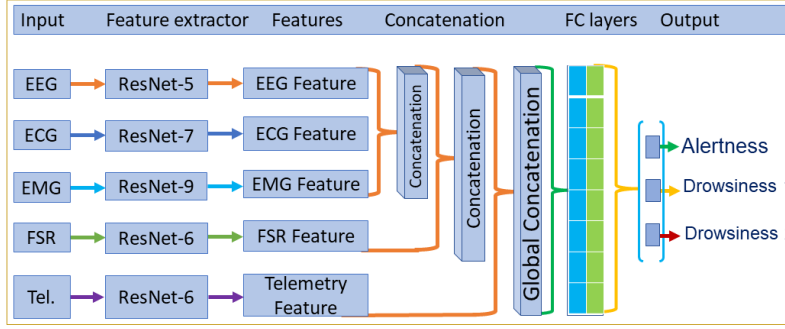


Figura 3.17: (i) Los datos de EEG, ECG y EMG varían solo en el número de electrodos utilizados para obtener estas señales, lo cual permite concatenar los mapas de características según los hiperparámetros establecidos en cada submodelo ResNet-N. (ii) Por otro lado, los datos de FSR y Telemetría varían tanto en su número de canales y muestras.

3.3.4.1.1. Red neuronal residual Una red ResNet es un tipo de arquitectura de red neuronal convolucional (CNN) capaz de manejar una gran cantidad de capas convolucionales con un mínimo impacto por el problema del gradiente de fuga [136, 137]. Este tipo de red utiliza conexiones de salto (skip connections) conocidas como bloques residuales, los cuales constan de capas convolucionales [138, 139]. Estos bloques saltan ciertas capas al principio y luego reutilizan las activaciones anteriores para dar una salida. De este modo, una ResNet aborda el problema del gradiente de fuga, generando un mayor gradiente de bloque. Esto mejora significativamente el aprendizaje, rendimiento general y reduce el problema de sobreajuste de un modelo. La Figura 3.18a representa un bloque de aprendizaje directo en el vector de entrada ($\mathbf{x} \rightarrow \mathbf{f}(\mathbf{x})$) y relaciona el problema de gradiente de fuga, donde $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{W} \in \mathbb{R}^{n \times n}$ es la matriz de los pesos y $\mathbf{y} \in \mathbb{R}^n$ es la salida de la capa que representa el mapa de características obtenido después de la convolución representada como $\mathbf{y} = \mathbf{W}^T \mathbf{x}$. Es decir, si la salida del i -ésimo bloque es la entrada del $(l + 1)$ -ésimo bloque, entonces $\mathbf{x}_{l+1} = \mathbf{G}(\mathbf{x}_l, \mathbf{W}_l^i) = \mathbf{W}_l^2 \beta(\mathbf{BN}(\mathbf{W}_l^1 \beta(\mathbf{BN}(\mathbf{x}_l))))$, siempre que \mathbf{W}_l^i sea el i -ésimo peso en el l -ésimo bloque, $i = 1, 2$ y β sea la función de activación o ReLU [140, 141].

Por otro lado, tomando como base el bloque de aprendizaje directo se deduce el bloque residual de la Figura 3.18b. Este bloque aprende de su mapa de características residual ($\mathbf{x} \rightarrow \mathbf{G}(\mathbf{x}) = \mathbf{G}(\mathbf{x}) - \mathbf{x}$). Considerando que \mathbf{x}_l es la entrada vectorial de ResNet, su vector de salida es \mathbf{x}_{l+1} en el l -ésimo bloque. Por lo tanto, la salida del módulo ResNet está definida como $\mathbf{x}_{l+1} = \mathbf{G}(\mathbf{x}_l) + \mathbf{x}_l = \beta[\mathbf{G}(\mathbf{x}_l, \mathbf{W}_l^i)] + \mathbf{x}_l$; $\mathbf{G}(\mathbf{x}_l, \mathbf{W}_l^i) = (\mathbf{W}_l^2)^T \beta((\mathbf{W}_l^1)^T \mathbf{x}_l)$. Así la representación del l -ésimo módulo se da como $\mathbf{y}_{l+1} = (\mathbf{W}_l)^T \beta(\mathbf{x}_l) + \mathbf{x}_l$ y su función de pérdida o (loss function) es $\xi = \frac{1}{2}(\mathbf{Y}_{l+1} - \mathbf{y}_{l+1})^2$. Considerando la etapa de normalización \mathbf{BN} como en la Figura 3.18c, \mathbf{x}_l es la entrada vectorial del l -ésimo bloque. La salida en este bloque se define como $\mathbf{x}_{l+1} = \mathbf{G}(\mathbf{x}_l, \mathbf{W}_l^i) + \mathbf{x}_l$. Donde $\mathbf{G}(\mathbf{x}_l, \mathbf{W}_l^i) = (\mathbf{W}_l^2)^T \beta(\mathbf{BN}((\mathbf{W}_l^1)^T) \beta \mathbf{BN}(\mathbf{x}_l))$; \mathbf{BN} representa la operación de normalización. Finalmente, la representación de este bloque es $\mathbf{y}_{l+1} = (\mathbf{W}_l)^T \beta \mathbf{BN}(\mathbf{x}_l) + \mathbf{x}_l$ y su función de pérdida está definida mediante $\xi_{\mathbf{BN}} = \frac{1}{2}(\mathbf{Y}_{l+1} - \mathbf{y}_{l+1})^2$ [140–142]. Esto permite deducir que la derivada del gradiente de bloque residual tiene un resultado más alto en comparación a la derivada de un bloque de aprendizaje directo, reduciendo al máximo la posibilidad de tener un gradiente de valor cero en la fase de backpropagation.

De esta manera, al considerar que la entrada del bloque (\mathbf{x}_{l+2}) está dada por la Ecuación 3.7,

la generalización de la entrada a un siguiente bloque se define por la Ecuación 3.8, donde L es el índice de cualquier bloque residual posterior y l es el índice de cualquier bloque anterior. Esta Ecuación demuestra que siempre hay una entrada que fluye hacia adelante directamente desde un bloque menos profundo l a un bloque más profundo L . Adicionalmente, la derivación de x_l en estado (forward propagation) permite deducir el comportamiento de la función de pérdida ξ en la etapa de (backward propagation), Ecuación 3.9 y ξ es la función a ser minimizada. Esto demuestra que cada término $\partial\xi/\partial x_l$ de una capa tiene agregado una fracción $\partial\xi/\partial x_L$, sugiriendo que el gradiente tiene agregado un término que incrementa su magnitud, incluso cuando el gradiente de $G(x_l)$ es realmente pequeño. Este comportamiento es lo que ayuda a controlar el problema del gradiente de fuga y hace que una ResNet sea más eficiente que una CNN tradicional.

$$x_{l+2} = x_l + G(x_{l+1}) + x_{l+1} = G(x_{l+1}) + G(x_l) + x_l \quad (3.7)$$

$$x_L = x_l + \sum_{i=l}^{L-1} G(x_i) \quad (3.8)$$

$$\frac{\partial\xi}{\partial x_l} = \frac{\partial\xi}{\partial x_L} \frac{\partial x_L}{\partial x_l} = \frac{\partial\xi}{\partial x_L} \left(1 + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} G(x_i) \right) = \frac{\partial\xi}{\partial x_L} + \frac{\partial\xi}{\partial x_L} \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} G(x_i) \quad (3.9)$$

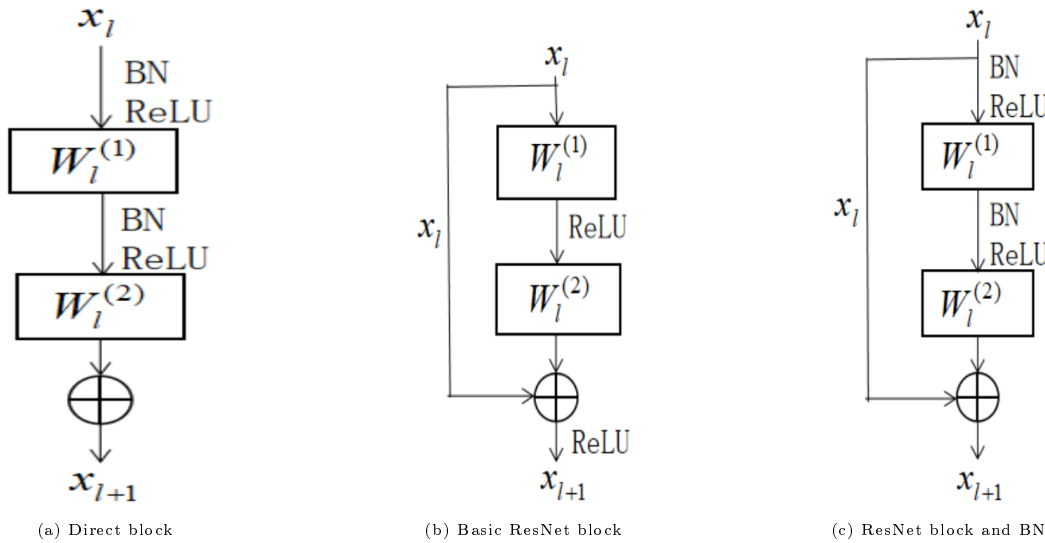


Figura 3.18: a) Representa un bloque de aprendizaje directo que mapea la entrada ($x \rightarrow f(x)$) y presenta problemas en modelos realmente profundos, donde solo tiene funciones de convolución y funciones de activación lineal. b) El modelo aprende del mapeo residual ($x \rightarrow g(x) = f(x) - x$) y dispone de la función de activación ReLU entre dos bloques, por lo tanto, el gradiente de salida de un bloque es mayor en comparación al gradiente de salida en un bloque de aprendizaje directo. c) Bloque ResNet donde el vector de entrada realiza la normalización (BN), luego pasa por la función de activación (ReLU) y finalmente se realiza la convolución [140, 141].

3.3.4.2. Algoritmo de pronóstico basado en señales

Este modelo de pronóstico emplea una secuencia de segmentos para generar una señal de salida del posible estado somnoliento subsecuente, la cual tiene el mismo formato y dimensión de la señal

de entrada en la etapa de clasificación. Los datos de entrada están definidos según una ventana de valores históricos o History Time Window (HTW) de tres minutos (180 segundos), equivalente a 45 segmentos de cuatro segundos cada uno. Este número corresponde al total de muestras que el algoritmo de pronóstico utiliza como entrada para generar la señal de salida con cuatro segundos de longitud. Las muestras resultantes se almacenan dentro de la ventana temporal definida como FTW de siete minutos (420 segundos), correspondiente a 105 segmentos subsecuentes de cuatro segundos de longitud cada uno y generados mediante el algoritmo de pronóstico.

De esta manera, el tamaño de las ventanas se fundamenta en: **a)** Durante el bloque dos, hay al menos un participante que experimentó somnolencia a los tres minutos de iniciado este bloque experimental. **b)** Por lo menos el 16 % de los participantes muestran señales de somnolencia de grado uno a los tres minutos de iniciado el bloque tres. **c)** En una ventana temporal de siete minutos es posible alertar al conductor de su condición subsecuente de somnolencia. **d)** El algoritmo de pronóstico requiere una GPU de mayor capacidad y más memoria RAM, imposibilitando definir una señal de datos históricos de mayor longitud. **e)** Si un evento subsecuente de somnolencia aparece a más de siete minutos, puede generar en el conductor una falsa sensación de tener demasiado tiempo para reaccionar y posiblemente continuará conduciendo. Por lo tanto, como el algoritmo de clasificación y pronóstico deben operar de manera alternada, Figura 3.19, siguen un orden de operación:

1. Cada segmento de señal de entrada clasificado como “0” se almacena en la ventana HTW.
2. Cuando la ventana HTW tiene 45 segmentos de cuatro segundos cada uno, el primer algoritmo deja de operar y el segundo modelo pronostica (genera) únicamente un segmento del subsecuente estado del conductor. Esta salida tiene las mismas dimensiones que la señal de entrada ($800 \times k$); k electrodos.
3. El segmento anteriormente pronosticado es clasificado por el primer algoritmo. Si esta nueva salida tiene la etiqueta “1”, se considera que el conductor estará somnoliento y se estima el tiempo del subsecuente evento de somnolencia. De lo contrario, el segmento pronosticado se almacena al final de la venta HTW, se elimina el primer segmento de esta misma ventana para conservar su longitud de 45 segmentos y el segundo algoritmo pronostica un nuevo segmento de señal.
4. El nuevo segmento pronosticado es clasificado por el primer algoritmo y se repite todo el procedimiento del tercer paso. Este proceso se repite hasta que la salida del primer modelo sobre cada segmento pronosticado sea una etiqueta “1” o el segundo algoritmo haya iterado 105 veces (420 segundos).
5. Una vez el primer algoritmo haya obtenido una etiqueta “1” sobre los datos pronosticados o el segundo algoritmo haya iterado 105 veces, los datos originalmente pronosticados se reintegran en la ventana HTW y se remueve el segmento inicial de esta ventana. El primer algoritmo clasifica una nueva señal de entrada y asigna un nuevo segmento al final de la ventana HTW para repetir el procedimiento desde el paso inicial. El tiempo de iteración entre ambos algoritmos debe tomar máximo cuatro segundos para que el clasificador pueda asignar una nueva etiqueta a un segmento de señal original (no pronosticado).

De esta manera, el tiempo en segundos del subsecuente evento somnoliento está determinado por $t = 4(\text{idx_iteration})$, donde cuatro es la longitud del segmento e idx_iteration es el índice de iteración donde el clasificador asignó una etiqueta “1” a un segmento pronosticado.

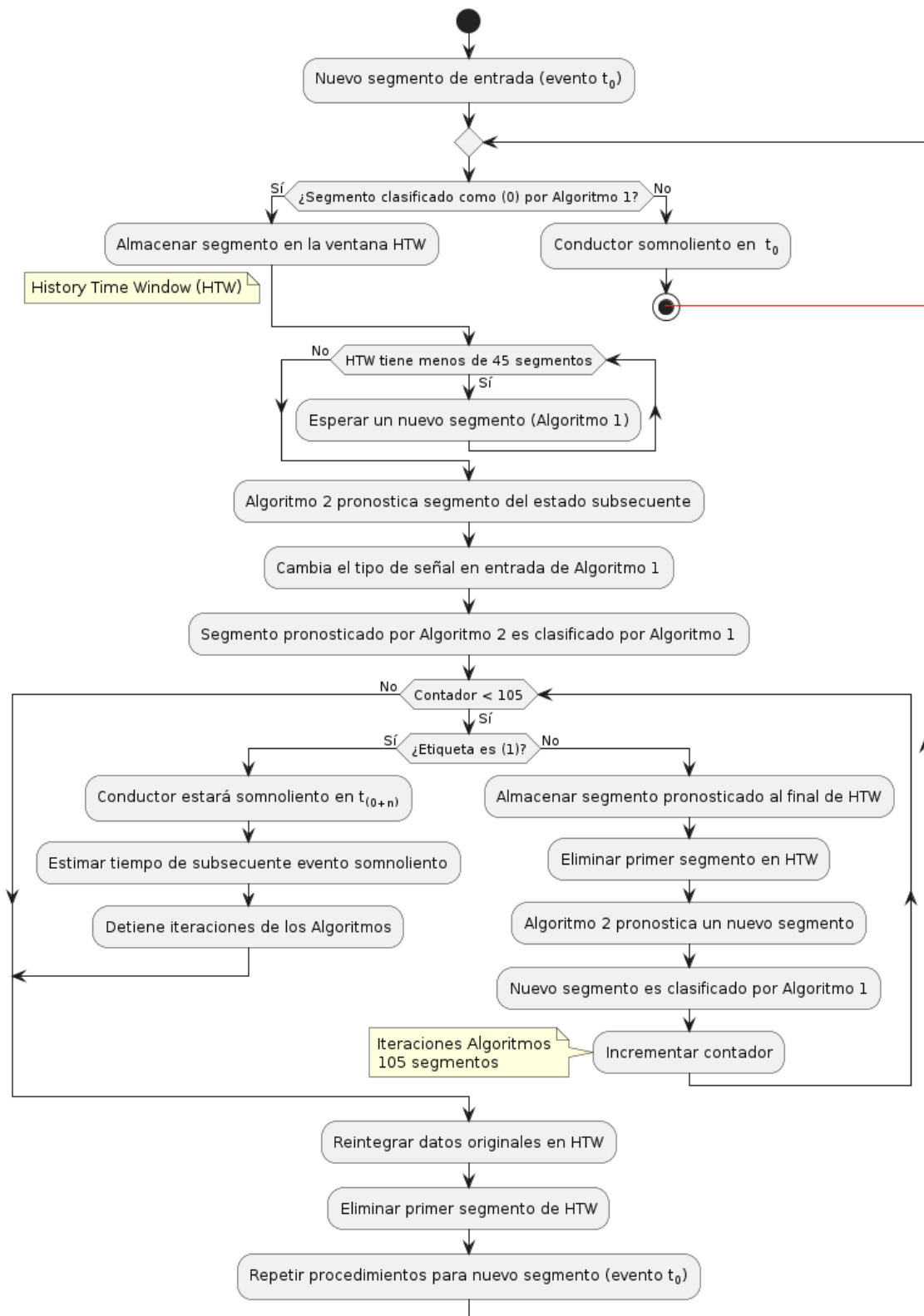


Figura 3.19: Diagrama de flujo en la operación de los algoritmos implementados en clasificación y pronóstico para el método uno. (i) Si el conductor está somnoliento en t_0 , el algoritmo uno terminará el proceso y regresa a su estado inicial (línea roja) para continuar clasificando nuevas muestras de la señal original.

El modelo de pronóstico está basado en una arquitectura CNN con una configuración de múltiples pasos temporales para predecir una secuencia de valores subsecuentes, Figuras 3.20 y 3.21. Esto permite hacer una predicción basada en datos históricos de ancho fijo, mejorando el rendimiento del modelo ya que puede estimar el cambio de información en el tiempo. Sin embargo, esta configuración requiere una mayor cantidad de datos históricos en HTW para mejorar la precisión de la salida. Un modelo basado en CNNs tiene la habilidad de capturar patrones locales en los datos para mejorar la precisión del algoritmo, pero esta ventaja disminuye drásticamente al capturar dependencias a largo plazo, lo cual se aprecia en los segmentos pronosticados al final de la ventana FTW. Por esta razón y para evitar utilizar una GPU más potente, en cada segmento de los datos fisiológicos, fue necesario reducir la frecuencia de muestreo a solo 200 Hz.

De esta manera, el modelo aprende a pronosticar 800 muestras tomando como entrada 36000 muestras fijas de la ventana HTW. Luego, en el segmento pronosticado se incrementa su frecuencia de muestreo a 500 Hz para conservar la misma dimensión y número de muestras del segmento de señal original. Aunque este proceso disminuye la cantidad de recursos computacionales utilizados por el modelo de pronóstico, concentra menos dependencias de información en una menor cantidad de muestras, reduciendo la precisión del modelo, pero posibilitando que los dos algoritmos tengan 105 iteraciones en un tiempo máximo de cuatro segundos.

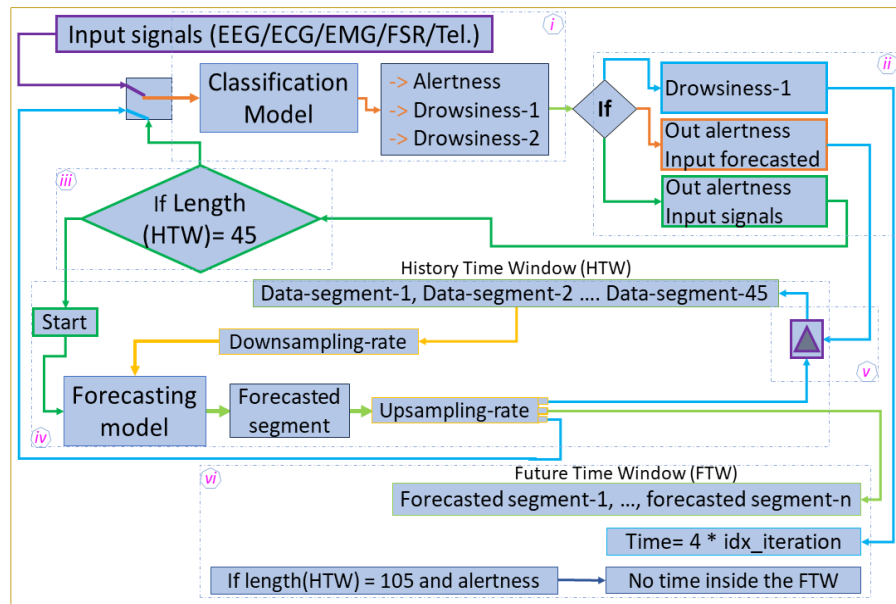


Figura 3.20: Esquema del sistema basado en el modelo de clasificación y pronóstico. (i) Cada cuatro segundos, el modelo de clasificación recibe una nueva señal de entrada del instante t_0 . (ii) En la etapa dos del sistema, se evalúa la salida del modelo de clasificación y su correspondiente señal de entrada para evaluar la etapa tres, la ubicación del segmento estimado en la ventana HTW o el cómputo del tiempo del subsecuente evento. (iii) Esta fase determina si la ventana HTW tiene los 45 segmentos requeridos y si la salida del modelo de clasificación es “alertness” para una entrada de nueva señal del instante t_0 ; esta fase se activa solo una vez cada cuatro segundos para iniciar la iteración del modelo de predicción. (iv) El modelo de pronóstico toma como entrada los 45 segmentos de la ventana HTW y estima un único segmento de cuatro segundos que es enviado al clasificador y a la ventana HTW. (v) Espera hasta que el segmento estimado sea clasificado y se obtenga una salida “alertness” para ubicar este segmento al final de la ventana HTW y eliminar el primer segmento de igual longitud de la misma ventana o desechar el segmento estimado. (vi) Se determina la cantidad de segmentos en la ventana FTW para definir si el conductor presentará un subsecuente evento somnoliento y el tiempo de este suceso o por el contrario, no se presentará un subsecuente evento en los siguientes siete minutos. Una vez termina la operación del algoritmo de pronóstico, de la ventana HTW se remueven todos los segmentos estimados para ubicar un nuevo segmento del instante t_0 y repetir todo el procedimiento.

Por otro lado, un mayor número de muestras en la ventana HTW requiere un tiempo de pronóstico o generación de los 105 segmentos mayor a cuatro segundos, Figura 3.21. En cuanto a los dos

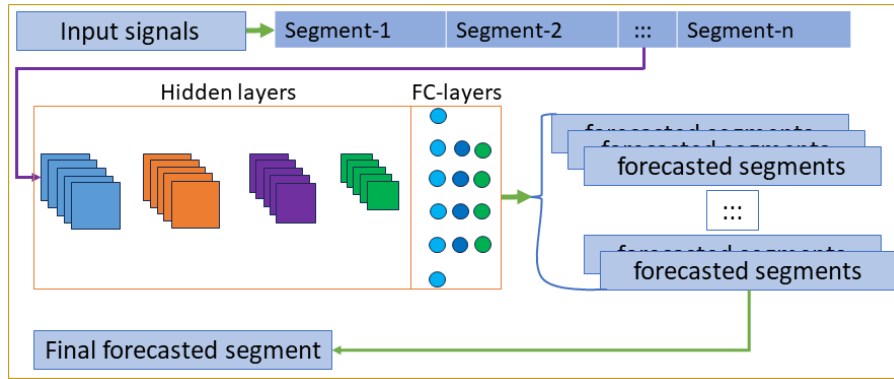


Figura 3.21: Las capas convolucionales aprenden automáticamente una gran cantidad de características locales en base a los filtros, lo que da como resultado mapas altamente específicos donde las capas completamente conectadas trabajan para pronosticar la señal.

modelos, estos se entrenaron individualmente. El clasificador se entrenó con todos los datos en las tres clases definidas, mientras que el segundo modelo se entrenó para pronosticar un segmento de señal del posible estado del conductor (clases “1”) en base a los datos de entrada. En esta etapa, se entrenó un modelo de pronóstico por cada submodelo de clasificación implementado. Los algoritmos de pronóstico utilizan la métrica del error cuadrático medio o RMSE para evaluar su desempeño, Ecuación 3.10. Este procedimiento representa la raíz cuadrada de las diferencias cuadráticas promedio entre los resultados previstos y los valores observados.

Para mejorar el proceso de aprendizaje, tanto en el modelo de clasificación como en el de pronóstico se utilizó la función EarlyStopping para detener el entrenamiento cuando la métrica de validación en la función de pérdida deja de mejorar. Además, se empleó un Learning Rate (LR) basado en una programación de la tasa de aprendizaje o (learning rate based-decay) y un algoritmo de descenso del gradiente mediante el optimizador Adam que mejora el entrenamiento del modelo.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3.10)$$

3.3.5. Método 2: algoritmos de regresión y pronóstico

Este método está configurado por un algoritmo de regresión para la etapa de monitorización en el tiempo t_0 , el cual entrega un valor continuo en el rango [0 a 2] y un segundo algoritmo de pronóstico para detectar los subsecuentes estados de somnolencia en el rango (t_0, t_{0+n}) ; con $n > 1$. El algoritmo de pronóstico utiliza directamente los niveles de somnolencia correspondientes a la salida del algoritmo de regresión.

3.3.5.1. Modelo basado en algoritmos de regresión

En este método, se implementó un modelo de regresión basado en CNNs mediante una arquitectura ResNet. Por conveniencia, el modelo implementado permite una salida continua dentro del rango (0 a 2) en pasos de 2/3. Esta salida permite estimar el nivel del estado somnoliento del conductor para el tiempo t_0 . Por lo tanto, todo valor menor que 2/3 corresponde a la etiqueta

“*alertness* $\equiv 0$ ”, de $2/3$ a $4/3$ representa la etiqueta “*drowsiness-1* $\equiv 1$ ” y de $4/3$ a 2 corresponde a “*drowsiness-2* $\equiv 2$ ”. Como la base de datos utilizada en esta sección es multimodal, cada conjunto de datos tiene representaciones intrínsecas y efectos estadísticos diferentes [21], implementar un único modelo que sea entrenado con todos los datos a la vez no presenta una ventaja en el rendimiento del modelo final. Esta característica agudiza el problema de tener un mapa de características estándar para todas las modalidades de estos datos [143]. Por lo tanto, para cada conjunto de datos se implementa un submodelo independiente llamado Independent Sub-Model (ISM) y están basados en una arquitectura ResNet. Finalmente, como se ilustra en la Figura 3.22, estos ISMs se ensamblan para configurar un modelo general (ensembled model).

Por otro lado, también se implementa un único modelo basado en un Global Feature Map (GFM) utilizando todos los data-set de la base de datos, Figura 3.23. En esencia, este modelo es similar al algoritmo de clasificación del método uno, pero aquí está configurado como un modelo de regresión, facilitando una salida continua. En este punto, se aprovecha las ventajas del método de ensamblaje de modelos (ensemble modeling), ya que es una técnica utilizada para predecir un resultado final basado en múltiples submodelos. Esto es posible utilizando conjuntos de datos de entrenamiento separados y algoritmos de modelado específicos [144]. No obstante, como se tiene varios conjuntos de datos y submodelos independientes (ISM), se espera que el error de clasificación final del modelo ensamblado sea menor al funcionar como un modelo único.

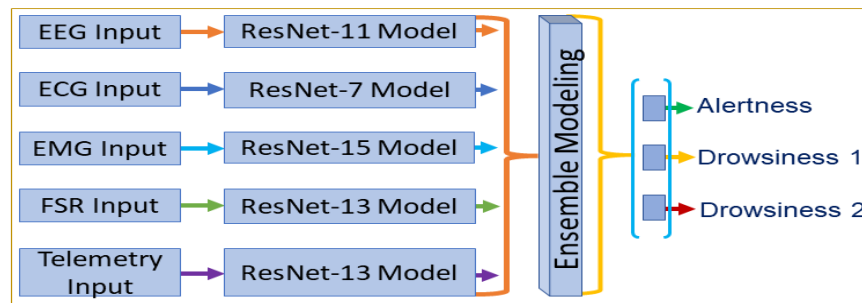


Figura 3.22: Este diagrama representa la implementación de cinco modelos independientes para ser ensamblados. (i) Cada modelo permite tener un valor de somnolencia continuo entre 0 y 2 en pasos de $2/3$. (ii) Si la salida del modelo es inferior a $2/3$, entonces el estado del conductor es considerado “alerta”; en caso contrario, se clasifica como somnolencia de grado “1” ó “2” según corresponda.

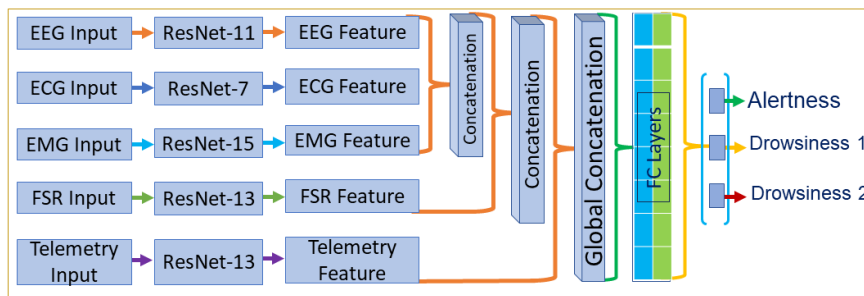


Figura 3.23: Diagrama de trabajo implementado para obtener un mapa de características único por concatenación de cada submodelo. (iii) La arquitectura de ResNet ayuda a reducir significativamente el problema del gradiente de fuga, lo cual incrementa el desempeño del submodelo, generando mapas de características con representaciones más significativas.

Cada ISM se diseñó buscando su mejor desempeño según la dimensión de cada conjunto de datos utilizado. Por ejemplo, el modelo *ECG-model* utiliza siete capas, donde los datos tienen la dimensión ($2000\text{-muestras} \times 1$) ya que disponen únicamente de la *II*-derivación del ECG estándar. En el diseño

de los ISMs, primero se implementó el submodelo de ECG y luego se modificó para adaptarlo al resto de ISMs según la dimensión de cada conjunto de datos. El bloque residual se implementó con capas convolucionales $2D$, una función de activación lineal rectificadora (ReLU) y un promedio $2D$ ($2D$ -average-pooling) sobre una ventana de tamaño (5×1) con un solo paso (1 -stride). Se utilizó el error cuadrático medio MSE como métrica de evaluación, Ecuación 3.11, la cual calcula la diferencia cuadrática promedio entre el valor objetivo y la salida de la red. Para optimizar el entrenamiento, se empleó el optimizador Adam (Adam optimizer) como algoritmo de descenso del gradiente. Además, se implementó una técnica de tasa de aprendizaje cíclico, que comenzó en 0,001, mejorando el aprendizaje del modelo y reduciendo el tiempo de entrenamiento.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.11)$$

3.3.5.2. Algoritmo de pronóstico basado en niveles de somnolencia

El algoritmo de pronóstico predice cuándo los valores DLs serán mayores que $2/3$ dentro del intervalo (t_0, t_{0+7}) o FTW. Este umbral corresponde a un valor intermedio entre las fases de alerta y somnolencia, donde el rendimiento del conductor comienza a disminuir durante la etapa “drowsiness-1”. El algoritmo de pronóstico estima los valores subsecuentes en base a los datos de la ventana HTW compuesta de DLs. En este método, las ventanas HTW y FTW están fijadas en siete minutos, equivalentes a 105 valores DLs cada una o 420 segundos: **a)** A los siete minutos en el bloque dos, al menos 25 % de los participantes mostró señales evidentes de somnolencia de grado uno, generando un deterioro en el desempeño al conducir. **b)** Durante el tercer bloque del experimento, al menos 50 % de los participantes experimentó somnolencia de grado uno en los primeros siete minutos de iniciado este bloque. **c)** Dentro de una ventana temporal de siete minutos es posible alertar al conductor si un subsecuente evento de somnolencia está próximo a ocurrir. Si esta ventana tiene una mayor longitud, el conductor podría tener una sensación de que su subsecuente estado somnoliento no es de emergencia y seguirá conduciendo por más tiempo del mínimo necesario para reaccionar adecuadamente.

En esta configuración, los algoritmos de regresión y pronóstico trabajan de manera alternada para monitorear el estado somnoliento en el tiempo t_0 y pronosticar el nivel de somnolencia para los siguientes siete minutos ($t_0 \rightarrow t_{0+7}$). La Figura 3.24 describe este flujo de trabajo.

1. El algoritmo uno estima el nivel de somnolencia (0 a 2) en los segmentos de entrada, donde los valores menores que $2/3$ se ubican en la ventana HTW. Este proceso se repite hasta que se completen 105 niveles en la ventana HTW; un nivel por cada segmento de 4 segundos.
2. El primer algoritmo deja de operar y el algoritmo dos emplea como entrada los 105 DLs de la ventana HTW para pronosticar 7-DLs, equivalentes a 28 segundos dentro de la ventana FTW.
3. Se calcula la media de estos siete valores y se verifica si este resultado excede el umbral fijado ($2/3$). Si este valor supera el umbral fijado, el índice de iteración del algoritmo determina el tiempo de previsión del subsecuente evento somnoliento de nivel “1”, llamado FTD-1. De lo contrario, si dentro del paso anterior no hay un valor superior a $2/3$, el algoritmo coloca los 7-DLs pronosticados al final de la ventana HTW, elimina los primeros 7-DLs de esta ventana para conservar la misma longitud de 105-DLs y genera 7-DLs nuevos.

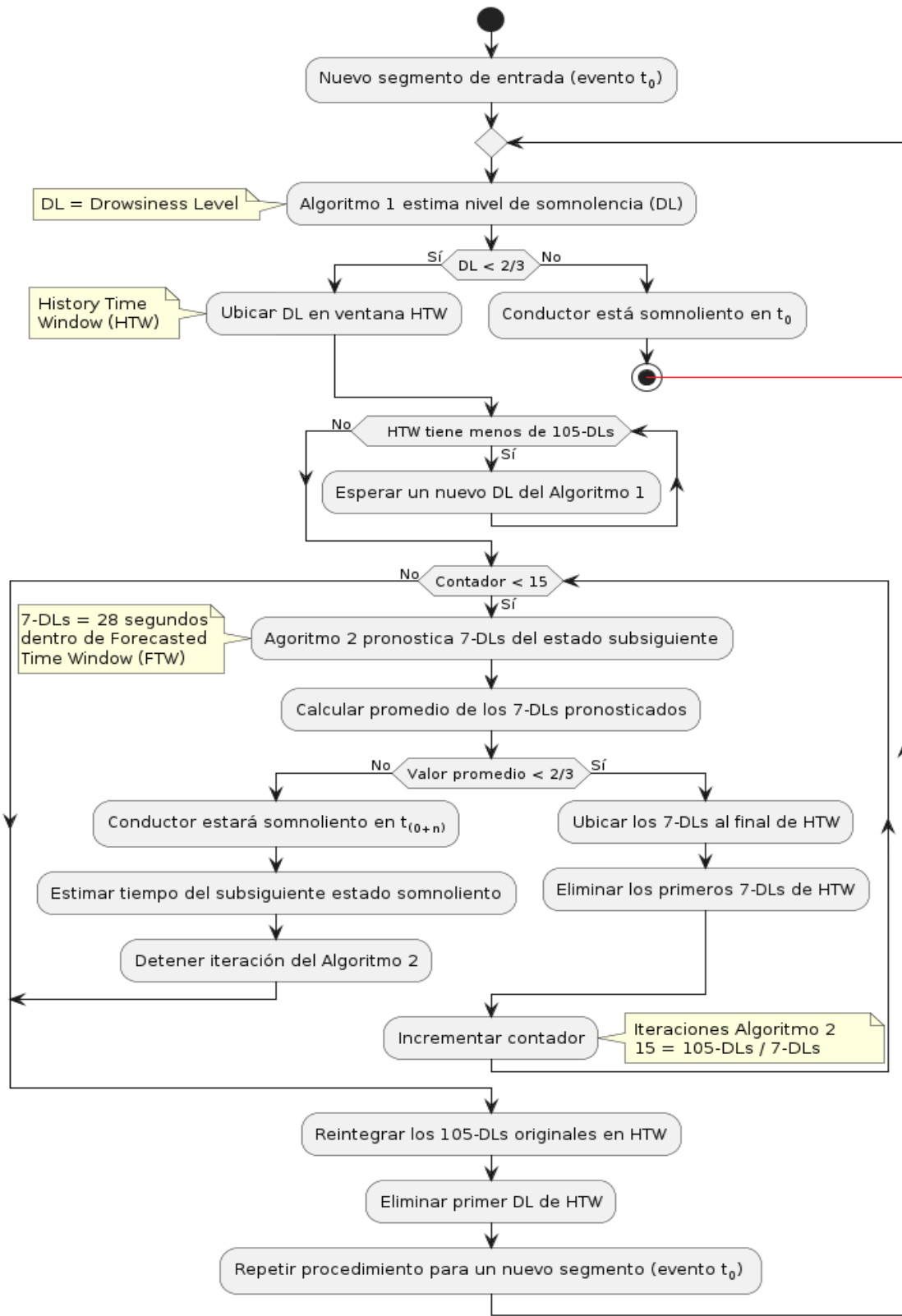


Figura 3.24: Diagrama de flujo en la operación de los algoritmos de regresión y pronóstico que utilizan las señales de entrada y los niveles de somnolencia en el método dos. (i) Cuando el conductor está somnoliento en el evento t_0 , el algoritmo de regresión regresa a su estado inicial (línea roja) para procesar nuevas muestras de entrada y reiniciar la ventana HTW inicial.

4. Se repite todo el procedimiento desde el primer paso hasta que haya un valor FTD-1 o se completen 15 iteraciones del algoritmo dos. Finalmente, se reintegran todos los DLs originales a la ventana HTW, se eliminan los primeros 7-DLs de esta ventana y se espera a que el algoritmo de regresión estime nuevos DLs para repetir todo el procedimiento, Figura 3.25.

El tiempo FTD-1 en segundos para un subsecuente evento de somnolencia está determinado por $t_{FTD-1} = (14 + 28(idx_iteration - 1))$; donde $idx_iteration$ es el índice de cada ventana de 7-DLs generada en la iteración del segundo algoritmo, 28 es la longitud de esta ventana en segundos y 14 es la mitad de la misma ventana evaluada dentro de FTW. Si no hay un valor FTD-1, entonces el conductor no tendrá un subsecuente evento de somnolencia dentro de los siguientes siete minutos.

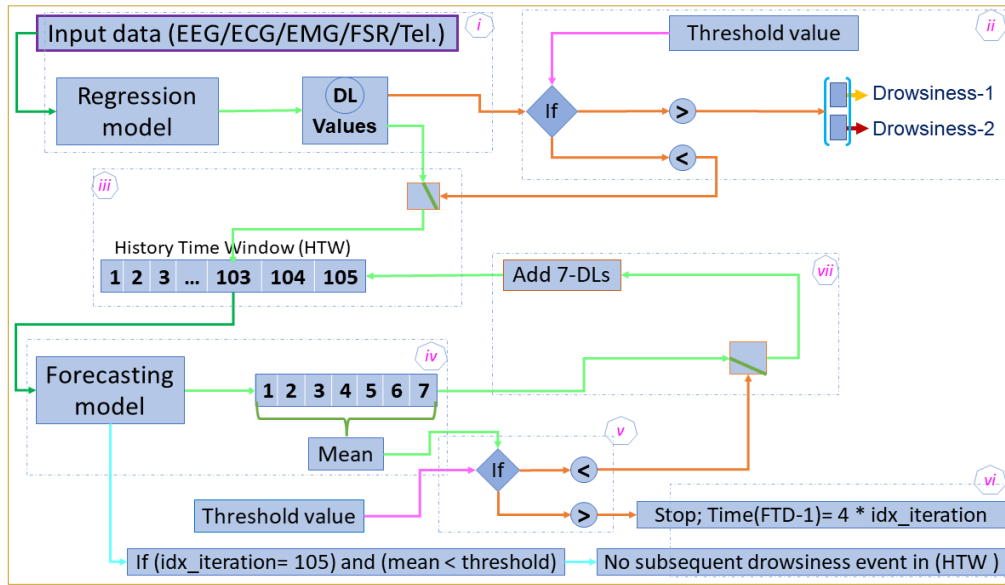


Figura 3.25: Esquema del sistema basado en el modelo de regresión y pronóstico que utiliza los valores DLs como entrada. (i) Representa la etapa de operación del algoritmo de regresión que entrega valores continuos en el rango (0, 2). (ii) En esta fase, los valores obtenidos de la etapa previa se comparan contra el valor de umbral (2/3). (iii) Si los valores DLs comparados en la etapa anterior son menores que 2/3, se ubican en la ventana HTW hasta completar 105 valores. (iv) El algoritmo de pronóstico utiliza como entrada todos los valores de la HTW y estima 7-DLs en los cuales se calcula su promedio. (v) El valor promedio obtenido en el paso anterior se compara contra el valor de umbral. (vi) Si el valor promedio es mayor que el umbral, se detienen la iteración del algoritmo de pronóstico y se calcula el tiempo de ocurrencia del subsiguiente evento somnoliento. (vii) Si en la etapa anterior el valor promedio es menor que el umbral, se ubican los 7-DLs al final de la ventana HTW, se remueven los primeros 7-DLs para conservar la misma longitud de esta ventana y si corresponde, se habilita una nueva iteración del algoritmo de pronóstico. El término $idx_iteration$ representa el número de iteraciones del algoritmo de pronóstico hasta que el promedio de los 7-DLs excede el umbral establecido.

El diseño del modelo de pronóstico está basado en capas bidireccionales de LSTM llamadas BiLSTM [42], Figura 3.26. Una LSTM es una variedad de redes neuronales recurrentes o Recurrent Neural Network (RNN), las cuales son una extensión de redes neuronales Feed-Forward convencionales, pero con una capacidad de gestionar entradas secuenciales con longitud variable. Esto les permite aprender dependencias a largo plazo entre pasos de tiempo en base a series de datos. Los modelos de RNNs configuran celdas para almacenar las entradas anteriores y aprovechar su información secuencial. La memoria temporal o estado recurrente oculto de una RNN les brinda la posibilidad de predecir cuál entrada será la siguiente en la secuencia de datos. Sin embargo, esta capacidad de memoria solo se limita a pocos pasos hacia atrás, creando dificultades para aprender dependencias a largo plazo, lo cual reduce la precisión del modelo.

De esta manera, si $\mathbf{x} = (x_1, x_2, x_3, \dots, x_L)$ representa una secuencia de longitud L , siempre

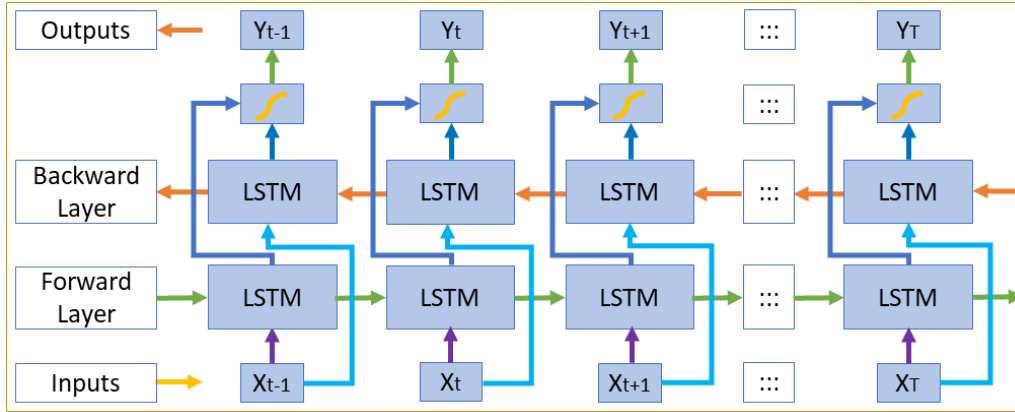


Figura 3.26: La arquitectura del modelo de pronóstico basado en BiLSTM hace que cada componente de la secuencia de entrada integre información tanto del pasado como del presente, lo cual permite que el modelo genere un valor más significativo tras la combinación de las capas LSTM.

que $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{W} \in \mathbb{R}^{n \times n}$ sea una matriz de pesos del modelo, entonces \mathbf{h}_l representa la memoria de una RNN para un paso l . Por lo tanto, el modelo de RNN actualiza su memoria de información en el paso l cuando $\mathbf{h}_l = \beta(\mathbf{W}_x \mathbf{x}_l + \mathbf{W}_h \mathbf{h}_{(l-1)} + \mathbf{b}_l)$. Donde β es una función no lineal como sigmoidea logística o ReLU, \mathbf{W}_x , \mathbf{W}_h son matrices de pesos y \mathbf{b}_l es el valor de sesgo constante [83]. Otro problema en las RNNs es su gradiente de fuga, donde estos modelos tienen dificultad al capturar las dependencias de información a largo plazo.

Por otro lado, un modelo LSTM extiende la memoria de los modelos simples de RNNs, permitiendo mantener y aprender dependencias a largo plazo en base a entradas temporales [83]. Esta configuración de memoria mantiene más información o características relevantes por mucho más tiempo y permite editar, leer y eliminar datos de su memoria. En una LSTM, esta memoria se conoce como celda cerrada y permite tomar la decisión de preservar o ignorar información almacenada. A diferencia de un modelo basado en CNNs, una LSTM utiliza las matrices de pesos adquiridas en la fase de entrenamiento para evaluar el tipo de información que se debe conservar o eliminar. Por esta razón, un modelo LSTM está configurado por:

- Al menos una puerta de remoción “forget gate” para preservar o remover información de la celda existente en base al valor de $\mathbf{h}_{(l-1)}$ y \mathbf{x}_l , Ecuación 3.12; donde \mathbf{f}_l es la salida de esta puerta en el rango $(0, 1)$ y \mathbf{b}_f es la constante de sesgo en el paso l .
- Una puerta de entrada “input gate” para determinar el límite hasta el cual la nueva información es agregada a la memoria. Esta puerta se compone de una capa de función sigmoidea para decidir qué información se actualizará, Ecuación 3.13 y una función tangente para crear una matriz de valores candidatos, Ecuación 3.14; donde \mathbf{i}_l representa la decisión de cambiar o no la información. La operación de estas dos capas genera los valores actualizados para la memoria, donde la puerta de remoción reemplaza los valores actuales por los valores candidatos $\mathbf{i}_l * \mathbf{c}_l$, Ecuación 3.15; donde \mathbf{f}_l es el resultado de esta puerta en el rango $(0, 1)$ para la combinación $(\mathbf{i}_l * \mathbf{c}_l)$.
- Una puerta de salida “output gate” que determina si el valor existente en la memoria favorece a la salida. Esta puerta utiliza la función sigmoidea para determinar que parte de la información en la celda LSTM contribuye con la salida, Ecuación 3.16. La función tangente $(-1, 1)$ mapea la información de la celda y la multiplica por el resultado de la función anterior, Ecuación

3.17; donde out_l representa el valor de salida en esta puerta y h_l en el rango $(-1, 1)$ es el resultado de la función tangente [83].

$$f_l = \beta(W_{f_h}[h_{(l-1)}], W_{f_x}[x_l], b_f) \quad (3.12)$$

$$i_l = \beta(W_{i_h}[h_{(l-1)}], W_{i_x}[x_l], b_i) \quad (3.13)$$

$$\tilde{c}_l^* = \tanh(W_{c_h}[h_{(l-1)}], W_{c_x}[x_l], b_c) \quad (3.14)$$

$$c_l = f_l * c_{(l-1)} + i_l + \tilde{c}_l^* \quad (3.15)$$

$$out_l = \beta(W_{out_h}[h_{(l-1)}], W_{out_x}[x_l], b_{out}) \quad (3.16)$$

$$h_l = out_l * \tanh(c_l) \quad (3.17)$$

Una extensión doble de una LSTM simple configura una BiLSTM [42], la cual dispone de dos LSTM en dirección opuesta. En una capa BiLSTM, la secuencia de datos fluye tanto hacia adelante como hacia atrás, aprovechando la información en ambos lados de la secuencia y modelando dependencias secuenciales en ambas direcciones. A diferencia de una capa LSTM estándar, una BiLSTM tiene una segunda capa que invierte la dirección del flujo de información y combina la salida de ambas capas LSTM, haciendo que cada componente de una secuencia de entrada tenga información del pasado como del presente [145]. Esta configuración mantiene conexiones de largo alcance y reconoce la relación entre los valores al principio y al final de una serie temporal, Figura 3.26. Además, esta arquitectura ayuda a resolver el problema de desaparición y explosión del gradiente de las RNNs simples.

Finalmente, el modelo de regresión se entrena con toda la base de datos en las tres clases definidas, pero el segundo modelo es entrenado por separado con el data-set de los niveles de somnolencia creado para este fin. En el entrenamiento del segundo modelo se utiliza el error cuadrático medio o MSE como métrica de evaluación. Esta función calcula la diferencia cuadrática promedio entre el valor objetivo y la salida DLs del modelo. Para optimizar el proceso de entrenamiento, se utiliza el algoritmo Adam como optimizador del descenso del gradiente. El entrenamiento del modelo se optimiza mediante una técnica programada para la tasa de aprendizaje basado en intervalos de tiempo (learning rate time-based), el cual se configura en 0,001 como valor de inicio.

3.3.5.3. Datos para los niveles de somnolencia

El data-set requerido por el algoritmo de pronóstico en el método dos, debe estar configurados como una serie temporal que permita entrenar y evaluar este algoritmo. En consecuencia, al no disponer de tal base de datos, es necesario implementar un procedimiento que permita utilizar los datos fisiológicos para transformarlos en valores DLs sin perder el orden en el que fueron adquiridos. En esencia, se puede considerar que los datos fisiológicos representan una serie ordenada

con mediciones tomadas en intervalos regulares (*segundos – bloque1 – bloque2*). Por lo tanto, para obtener los datos en los niveles de somnolencia, una vez entrenados todos los algoritmos de regresión, solo se utiliza una de las categorías de datos fisiológicos no segmentados, ya que representan una serie en configuración (*segundos – bloques*).

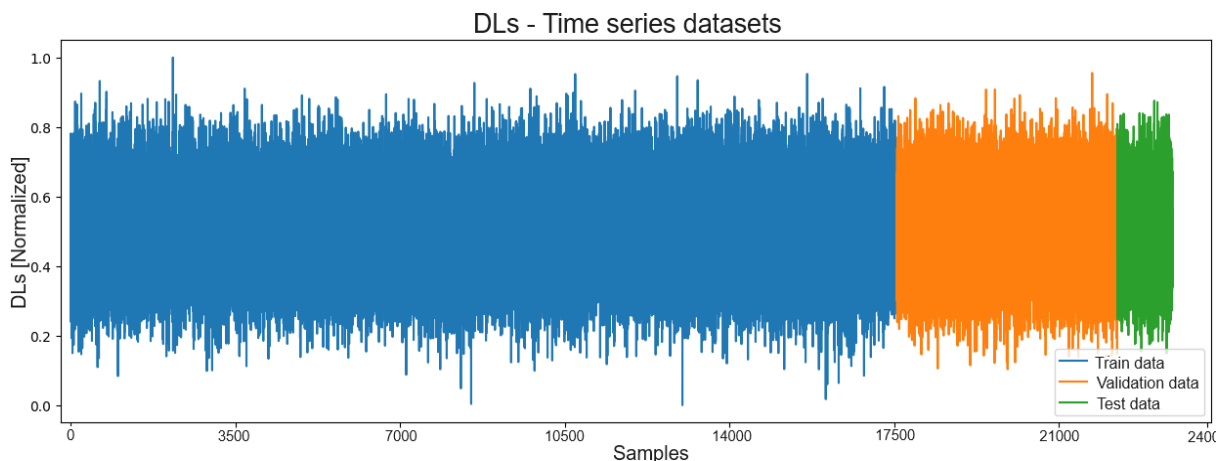


Figura 3.27: Orden de los datos en la división de la nueva serie de valores DLs organizados para entrenamiento, validación y prueba.

Los datos de los participantes se concatenan conservando la secuencia de registro en la cual se obtuvieron. Esto configura una *nueva serie* temporal basada en (*segundos – bloques – sujetos*) que conserva el orden de los datos y participantes. Por consiguiente, se emplea el algoritmo de regresión con mejor valor *test_accuracy* para obtener los valores DLs, el cual utiliza como entrada los segmentos ($seg_1, seg_2, seg_3, \dots, seg_n$) de la *nueva serie* y se obtiene la salida de valores DLs ($DL_1, DL_2, DL_3, \dots, DL_n$). Por consiguiente, cada resultado DLs se compara con los valores verdaderos para corroborar que corresponden con su respectiva clase. En este punto, todos los valores DLs se normalizaron a la escala (0 a $4/3$), esto genera que la nueva serie de valores DLs tenga igual número de muestras y con el mismo orden en que los datos fueron adquiridos. Finalmente, esta serie se divide en data-sets para entrenamiento (70%), validación (20%) y prueba (10%), Figuras 3.27 y Figura 3.28. Finalmente, estos nuevos datos son los que se utilizan para entrenar el algoritmo de pronóstico del método dos.

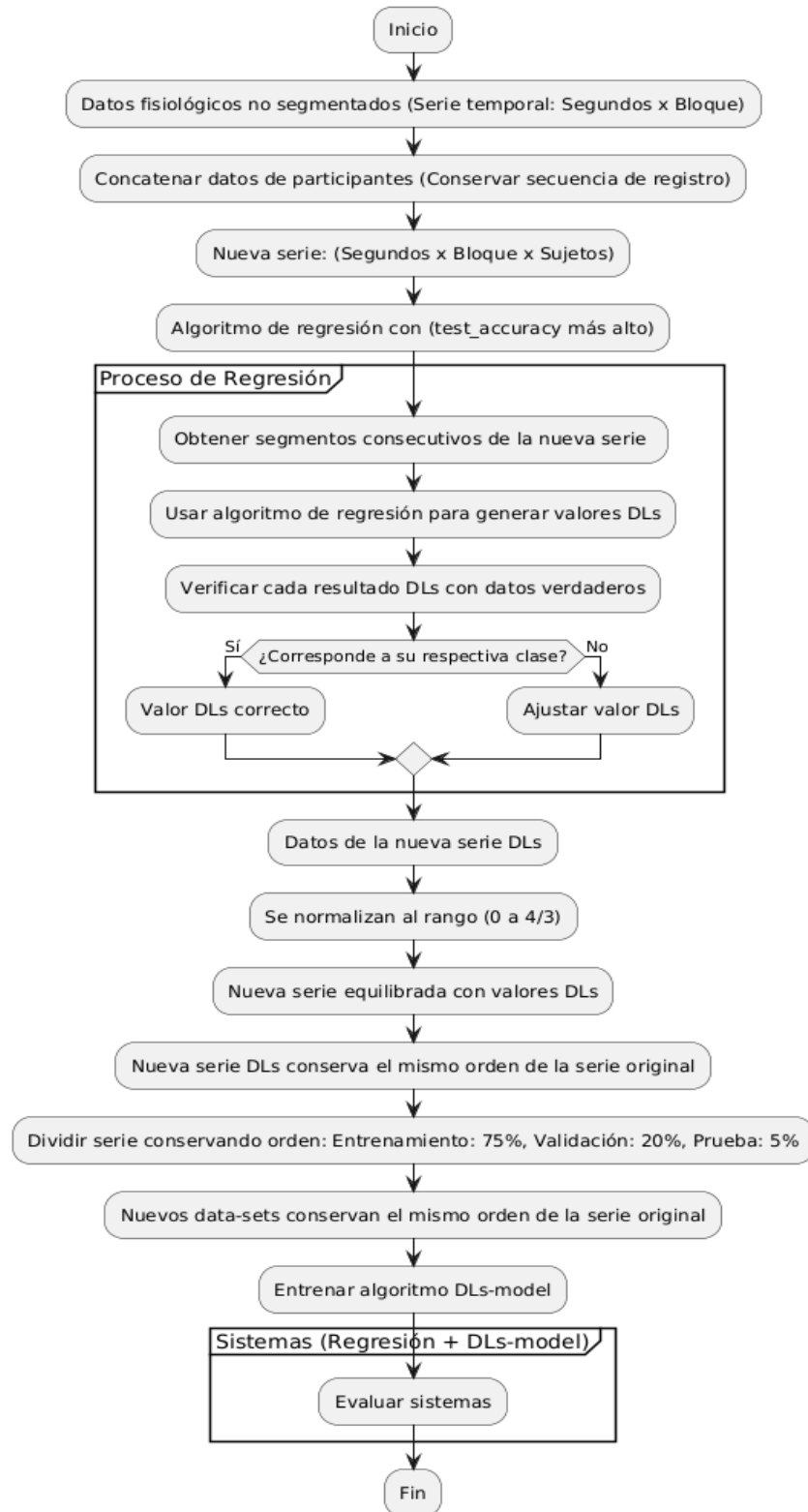


Figura 3.28: Flujo de trabajo realizado para obtener los datos de la serie temporal de valores DLs utilizando una categoría de datos fisiológicos y el algoritmo de regresión con mejor métrica de evaluación.

4.1. Introducción

En este capítulo se examinan los resultados obtenidos del análisis de los datos, los algoritmos implementados, las métricas de evaluación y los resultados logrados en cada una de las etapas presentadas. En relación al entrenamiento de los modelos, algunos se entrenaron en Google Colab [146] dado los requerimientos de GPU y memoria RAM requeridos. Sin embargo, todos los modelos se implementaron en Python [147] mediante la librería de código abierto TensorFlow-Keras [148, 149], mientras que el preprocesamiento de los datos se realizó a través de Matlab [108].

4.2. Detección de la intención de frenada de emergencia

Cuando un conductor está bajo efectos de la fatiga física, somnolencia o incluso con sobre carga cognitiva, incrementa su tiempo de reacción en eventos de frenada de emergencia. Por lo tanto, un sistema capaz de predecir o detectar con anticipación la necesidad de frenar ante una emergencia ayudaría a reducir el tiempo que se necesita para pisar el pedal de los frenos y detener la marcha del vehículo, reduciendo la posibilidad de una colisión. Un sistema con tales características bien podría funcionar en conjunto con un dispositivo orientado a pronosticar el tiempo de ocurrencia de subsiguientes eventos de somnolencia en el conductor. Por otro lado, esta sección expone los resultados obtenidos en la etapa para la detección de la intención de frenada de emergencia, lo cual constituye una fase preliminar para el análisis de los datos multi-sensor resultantes del experimento de conducción simulada.

4.2.1. Tiempo de reacción en frenada de emergencia

El análisis de los datos evidencia que el tiempo medio de reacción de frenado o BRT para la totalidad de los participantes fue de 691 ± 115 milisegundos. Los valores máximo y mínimo fueron 1195 y 455 milisegundos, respectivamente. La Tabla 4.1 resume los valores BRT para cada bloque y el promedio general en el conjunto de participantes. La diferencia porcentual en el valor de standard deviation (std) entre los bloques uno y dos es aproximadamente similar, pero entre los bloques uno y tres es del 15 %, lo que demuestra una mayor variación del desempeño de los sujetos conforme aumenta el tiempo en la tarea de conducción. La variación del valor std entre los bloques es una

consecuencia de la fatiga física y cognitiva acumuladas que se produce en los sujetos por el tiempo pasado en el simulador de conducción (tres bloques de 45 minutos cada uno). Por otro lado, la Figura 4.1 muestra que en el bloque tres se produjeron más colisiones que en el bloque uno, reforzando el argumento anterior.

Tabla 4.1: Resultados de la medición del tiempo de reacción de frenada de emergencia en cada uno de los tres bloques.

Block name	BRT: Average \pm std	BRT: Maximum	BRT: Minimum
Block 1	689 \pm 115	1195	455
Block 2	690 \pm 118	1195	460
Block 3	692 \pm 133	1195	455
Total Blocks (1,2,3)	690 \pm 122	1195	455

i) Entre el primero y el segundo bloque únicamente hay un valor std de (± 003). Sin embargo, el resultado std en los últimos 45 minutos del estudio es mayor (133) respecto a los primeros 45 minutos (115). Esto significa que algunos participantes frenaron más tarde que otros dado los posibles niveles de fatiga física y estrés por la carga de trabajo. *ii*) El valor máximo de BRT no es similar en todos los bloques porque este tiempo se ve afectado por el número de colisiones que sufrió el participante en cada bloque.

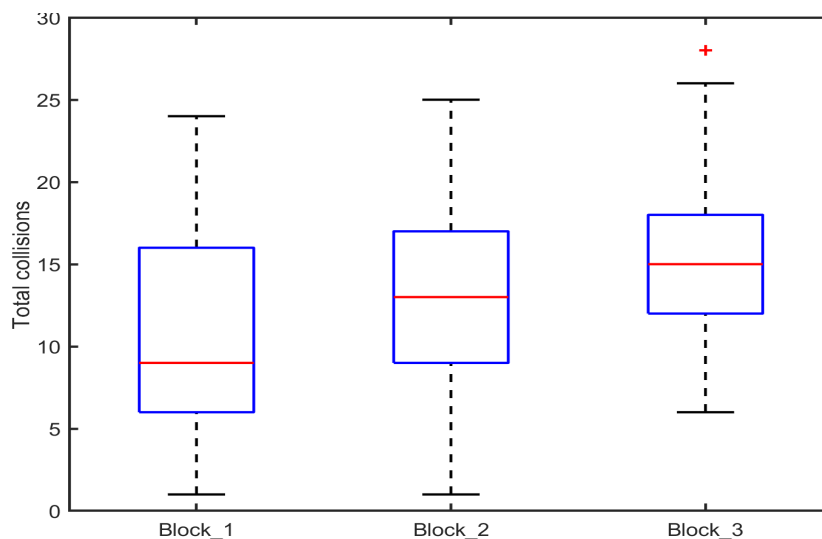


Figura 4.1: Representación del número total de colisiones por cada uno de los tres bloques del experimento. *(i)* El parámetro ($IQR = Q_3 - Q_1$) corresponde al rango intercuartílico y representa la variabilidad de los datos. *(ii)* En el bloque tres, el máximo valor atípico ($Q_3 + 1,5 * IQR$) es 27, pero el mínimo valor atípico ($Q_1 - 1,5 * IQR$) de este bloque es 6 y para el bloque dos es 29 y 3, respectivamente.

En relación con el número de colisiones que determina el desempeño de los participantes, en el segundo bloque de 45 minutos, la mediana (percentil 50) es 13 y tiene un valor de ocho como Inter-Quartile Range (IQR), Figura 4.1. Sin embargo, a pesar del descanso de diez minutos entre bloques, este parámetro varía en cada bloque del experimento. El valor IQR en los últimos 45 minutos de la tarea de conducción es seis y el del primer bloque es diez, aunque con una media más alta (en torno a 15 colisiones). Estos resultados pueden interpretarse como un efecto de la acumulación de fatiga física y cognitiva por la tarea realizada. Sin embargo, el desempeño del sujeto en la tarea está asociado a factores de experiencia en conducción, capacidades físicas y cognitivas de cada individuo para gestionar el estrés en el tiempo (diferencias intra-individuales).

Por otro lado, el desempeño del modelo implementado también fue evaluado al entrenarse con

los datos de imágenes en escala de grises correspondientes a características directas de los datos. El rendimiento del modelo basado en redes CNNs se validó utilizando las métricas de accuracy, precisión, área bajo la curva o Area Under the Curve (AUC) y la curva de características o Receiver Operating Characteristic (ROC). Estas métricas pueden evaluar qué tan bien el clasificador distingue entre las clases *objetivo* y *no-objetivo*; a mayor área bajo la curva ROC, mejor se desempeñará el clasificador, Tabla 4.2. La Figura 4.2 muestra los valores de la curva ROC en el grupo de 59 electrodos con datos del método matricial y la Figura 4.3 representa la curva AUC para el mismo modelo entrenado con imágenes en escala de grises. El valor AUC más bajo entre estos dos métodos se obtiene cuando el modelo se entrenó con imágenes en escala de grises.

Por el contrario, el clasificador puede distinguir muy bien entre segmento de frenada de emergencia y conducción normal, aun cuando este algoritmo se entrenó directamente con los segmentos obtenidos de las señales. Es de aclarar que, los mismos datos se utilizaron en el método matricial y para obtener las imágenes en escala de grises. Estas imágenes se utilizaron para entrenar un modelo idéntico, donde no cambia su configuración ni hiperparámetros, posibilitando una comparación de los resultados obtenidos.

Tabla 4.2: Resultados de las métricas accuracy y Area Under the Curve (AUC) para cada uno de los grupos de imágenes y electrodos.

Data		Accuracy (%)		ROC(AUC)	
Electrode groups	Set-image groups	Electrode	Image groups	Electrode	Image groups
4	Set 1	82.5	58.2	0.81	0.59
8	Set 2	79.1	57.7	0.82	0.60
13	Set 3	83.5	60.5	0.83	0.61
18	Set 4	86.2	64.1	0.81	0.66
33	Set 5	83.6	70.2	0.82	0.65
59	Set 6	87.5	68.1	0.86	0.68

(i) Resultado de la métrica de accuracy y ROC(AUC) en cada grupo de electrodos y conjunto de imágenes. (ii) Cada conjunto de imágenes representa a todos los electrodos de su respectivo grupo utilizados para calcular las imágenes en escala de grises. (iii) El valor accuracy en el grupo de cuatro electrodos es sólo un 5 % inferior que el resultado de accuracy en el grupo con 59 electrodos. (iv) El valor de ROC(AUC) en el grupo de 59 electrodos (0.865) no es significativamente mejor que el valor de la misma métrica para el grupo de cuatro electrodos (0.825), donde los demás grupos tienen resultados similares tanto en la métrica de accuracy como en los valores ROC(AUC). (v) El Set-5 tiene un valor accuracy solo 3.9 % inferior que el set-6, por lo tanto, el rendimiento del clasificador no sólo depende del número de imágenes, sino también de la calidad de la información contenida en cada imagen de espectrograma. (vi) El valor ROC(AUC) en el Set-1 tiene el valor más bajo (0.59) debido al menor número de imágenes obtenidas con este conjunto de electrodos.

En cada modelo para los grupos de electrodos e imágenes se buscó la mejor tasa de aprendizaje para incrementar el valor AUC. Sin embargo, los resultados en los sets de imágenes son menores y el tiempo de entrenamiento es mayor comparado con los resultados obtenidos en los grupos de electrodos, Figura 4.3 y Tabla 4.2. Adicionalmente, en la Figura 4.4 se puede apreciar los resultados de las métricas de precisión, AUC y accuracy para todos los grupos. El grupo de cuatro electrodos tiene un valor accuracy de 0.825 y es aproximadamente 5 % menor que el resultado de esta métrica en el grupo de 59 electrodos (0.875). A pesar del menor valor de la métrica de precisión en el grupo de ocho electrodos (0.780), su valor AUC es 0.825 y muestra que el clasificador todavía puede diferenciar entre conducción normal y frenado de emergencia. Estos resultados resaltan un comportamiento consistente relacionado con la mínima diferencia de información presente en las

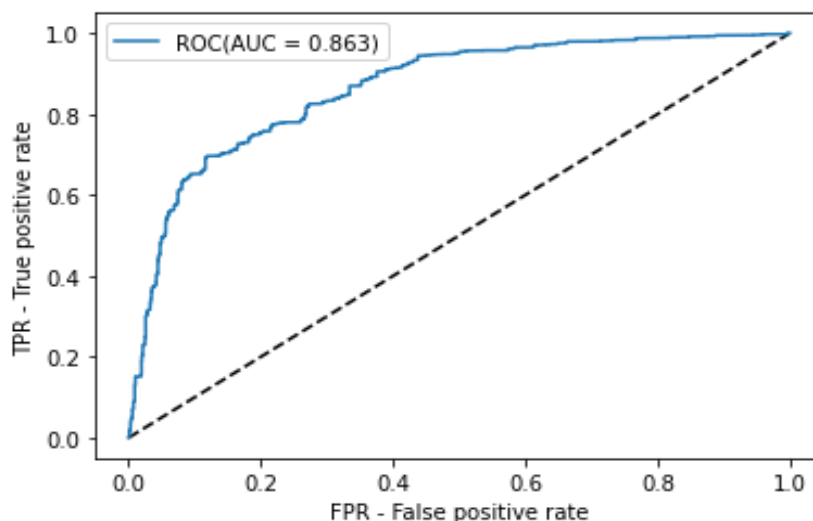


Figura 4.2: Representación de la curva ROC con un valor AUC de 0.863 en el grupo de 59 electrodos con datos matriciales. (i) La métrica de ROC mide cómo están cambiando la tasa de verdaderos positivos (recall) y la tasa de falsos positivos, haciendo que esta métrica se aplique con mayor frecuencia en clasificadores binarios. (ii) AUC es una evaluación del clasificador porque el umbral varía en todos los valores posibles. (iii) Se considera que no es necesario comparar más gráficos de ROC(AUC) entre los diferentes grupos de electrodos e imágenes porque tendrán conclusiones similares.

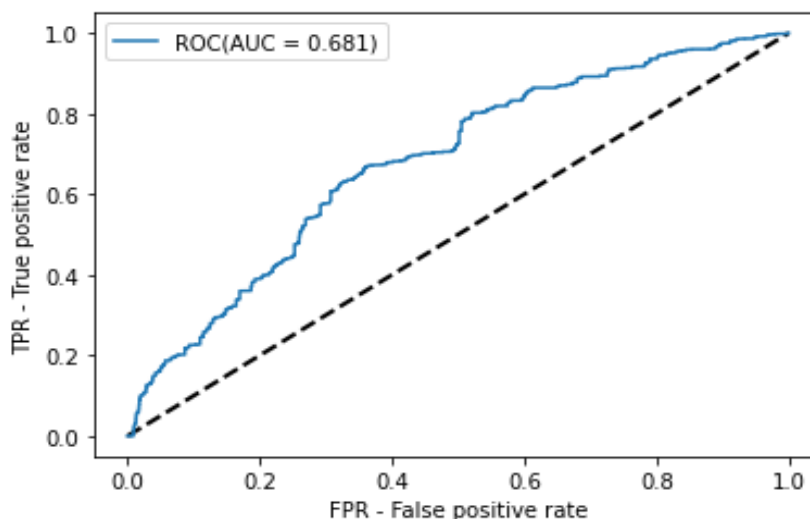


Figura 4.3: Curva ROC correspondiente a los resultados del modelo entrenado con imágenes en el set-6 de 59 electrodos, donde hay un valor AUC de 0.681. (i) este modelo consume mayores recursos computacionales y tiempo de entrenamiento en comparación con el modelo entrenado directamente con señales.

dos clases (segmentos objetivo y no-objetivo). Además, el grupo de 59 electrodos tiene un valor accuracy (0.875) y AUC de 0.865 mayor a todos los demás grupos de electrodos.

4.2.2. Potenciales relacionados con eventos

En esta sección se estimó la latencia promedio de la onda P300 en las regiones de la corteza visual primaria PVC y la corteza motora primaria (PMC) en cada bloque del experimento. Al tener una respuesta alineada con un estímulo visual en el experimento, la onda P300 es la más

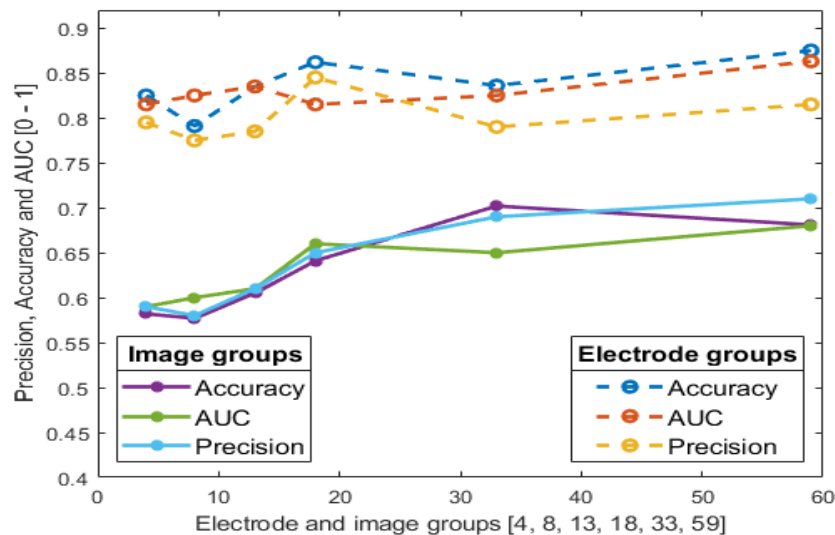


Figura 4.4: Resultados de las métricas AUC y accuracy para todos los conjuntos de electrodos e imágenes. (i) A medida que aumenta el número de electrodos en los datos de matrices, las métricas de evaluación aumentan mínimamente. No obstante, el tiempo de entrenamiento y los recursos computacionales aumentan a medida que crece el número de electrodos. (ii) Los resultados al utilizar los grupos de imágenes fue afectado significativamente por la cantidad de estas obtenidas a partir del número de electrodos en cada grupo.

representativa en estas regiones cerebrales. Esta onda se presenta como una larga deflexión positiva en un intervalo de 250 a 550 milisegundos después de presentarse un estímulo a intervalos aleatorios frecuentes [150]. La onda P300 es un componente de potenciales relacionados con eventos (Event Related Potential (ERP)) que puede utilizarse para medir el proceso de toma de decisiones y la velocidad de clasificación o discriminación del estímulo presentado. Latencias más cortas se asocian con una mayor capacidad y mejor rendimiento cognitivo en el procesamiento de información [151]. Además, la amplitud (microvoltios) de la onda P300 está relacionada con factores relevantes de la tarea y la atención del sujeto. Así, la amplitud y latencias de esta onda se pueden ver afectados por la fatiga o somnolencia del participante durante el experimento.

Adicionalmente, es necesario establecer si hay dependencia entre las clases (*objetivo* y *no-objetivo*) que permitan afirmar que la onda P300 sólo se presenta en los eventos para frenada de emergencia como respuesta al estímulo visual establecido por las luces de frenada del vehículo líder. Si no hay diferencia entre las dos clases, los resultados de la onda P300 no serían concluyentes respecto del estado cognitivo de los participantes durante el experimento. Para ello, en los datos con matrices $2D$, se estimó el coeficiente de correlación que mide qué tan fuerte es la dependencia entre las dos clases. El valor de relación está normalizado entre (-1 y 1), siendo 1 la mayor dependencia positiva, -1 una fuerte dependencia inversa y un valor cero indica que no hay fuerza de asociación. La Figura 4.5 resume los valores obtenidos, donde el valor IQR es 0.33 con un valor medio (Q_2) en 0,1. Esto demuestra que la mayoría de los segmentos para frenada de emergencia pueden diferenciarse muy bien de las muestras de conducción normal y la onda P300-ERP es una respuesta única del estímulo visual para frenada de emergencia.

Por lo tanto, los valores de latencia se resumen en la Figura 4.6 para la región PVC y Figura 4.7 para la región PMC. El promedio general de los segmentos de señal objetivo de las regiones PVC y PMC evidencia ondas ERP espacio-temporales significativamente positivas, Figura 4.8. Después del inicio de estímulo, los electrodos $C3$ y $O2$ resaltan picos positivos relevantes de $9,1 \mu V$ a 355 milisegundos y $8,1 \mu V$ a 350 milisegundos, respectivamente. Además, el electrodo Cz muestra un

pico de $9,4 \mu V$ a 340 milisegundos en la región PMC.

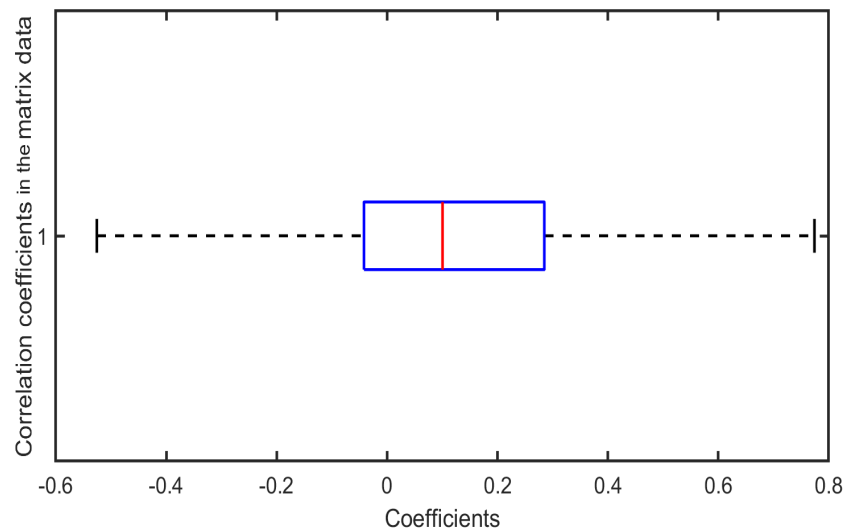


Figura 4.5: (i) En los coeficientes de correlación para el conjunto de datos matriciales, el percentil 25 está en el punto $-0,05$ y el máximo valor atípico ($Q_3 + 1,5 * IQR$) es $0,33$. (ii) El IQR está entre los valores $(-0,05$ a $0,25)$ con un Q_2 cercano a cero, demostrando que entre el conjunto de datos matriciales para las clases *objetivo* y *no-objetivo* no existe una fuerza de asociación relevante.

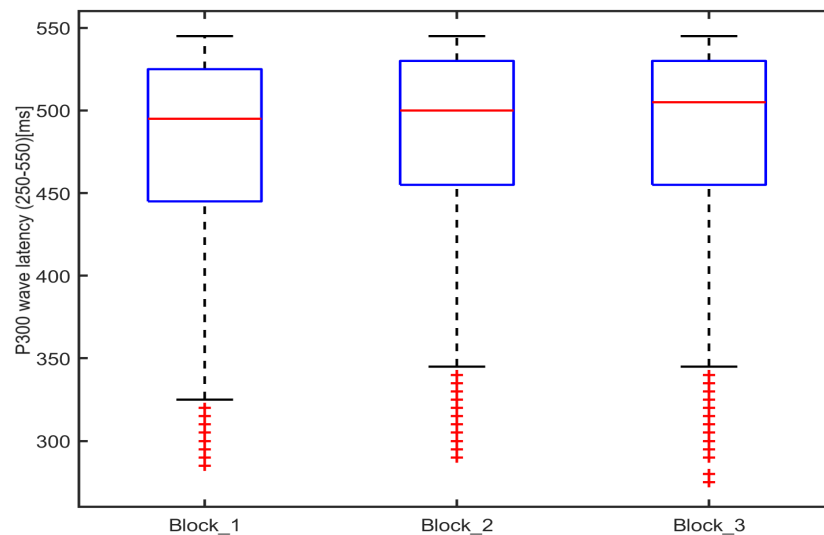


Figura 4.6: Latencias de la onda P300 en la región de la Corteza Visual Primaria. (i) El bloque tres presenta el mayor número de valores mínimos atípicos ($Q_1 - 1,5 * IQR$) con 330 milisegundos. (ii) Por el contrario, este parámetro es de 225 para el bloque uno, con un percentil 25 de 495 milisegundos.

4.3. Detección y pronóstico de eventos somnolientos

En esta sección se presentan los resultados concernientes a la detección del estado somnoliento del conductor en el tiempo t_0 y el tiempo de aparición de subsecuentes eventos de somnolencia dentro de una ventana temporal de siete minutos. Para ello, en el método uno se asigna una etiqueta de $[0, 1$ o $2]$ a cada segmento de cuatro segundos mediante un algoritmo de clasificación y un algoritmo de

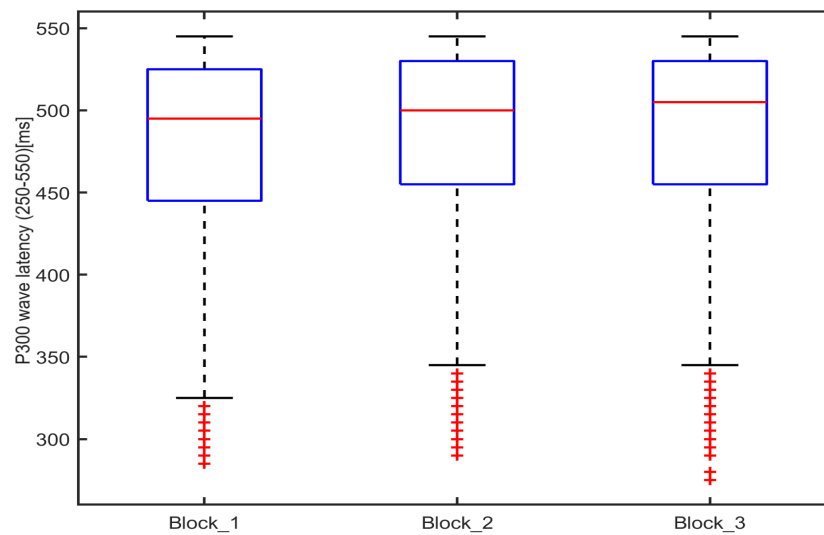


Figura 4.7: Latencia de la onda P300 en la región de la Corteza Motora Primaria. (i) Se aprecia que la diferencia del número de valores mínimos atípicos ($Q_1 - 1,5 * IQR$) entre el bloque uno y tres es de aproximadamente 15 milisegundos, siendo el último bloque el que presenta un mayor tiempo de este parámetro (335) milisegundos.

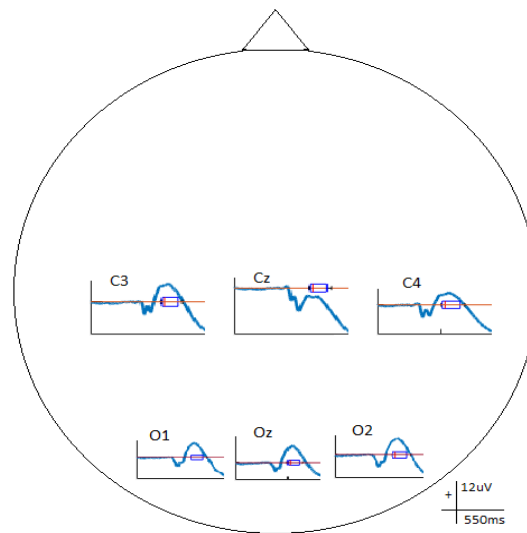


Figura 4.8: Promedio general de la onda P300-ERP en la corteza motora primaria (arriba) y en la corteza visual primaria (abajo). (i) Se evidencia que en la región central se ha generado un fuerte impacto en los ERPs, observable en los electrodos Cz y Oz .

pronóstico para generar una señal del posible estado de somnolencia dentro del rango ($t_0 \rightarrow t_{0+7}$). Por otro lado, en el método dos se analiza el nivel de somnolencia en el tiempo t_0 mediante un algoritmo de regresión con una salida continua en el rango (0 a 2) y se implementa un algoritmo de pronóstico para estimar el nivel de somnolencia dentro del intervalo (t_0, t_{0+7}). Cada método configura un sistema donde se implementaron diversas combinaciones de los datos y algoritmos para obtener el mejor modelo.

4.3.1. Método 1: algoritmos de clasificación y pronóstico

En base a que las señales se dividieron en segmentos de cuatro segundos, el monitoreo del estado somnoliento del conductor tiene una resolución equivalente a esta longitud. Adicionalmente, este primer método permite tener una resolución de cuatro segundos en el pronóstico de subsecuentes eventos de somnolencia. Por otro lado, al disponer de una base de datos multi-sensor compuesta por datos de EEG, ECG, EMG, FSR y Telemetría del vehículo, es necesario implementar un sistema multi-modelo, permitiendo encontrar diversas características de los datos.

4.3.1.1. Modelos de clasificación

Cada modelo de clasificación se entrenó utilizando su correspondiente categoría de datos, donde además se implementaron diferentes combinaciones entre los datos y cantidad de electrodos en el data-set de EEG. Estos modelos se evaluaron utilizando la métrica de accuracy, donde 1.0 es el máximo valor que se puede obtener. La Tabla 4.3 presenta los resultados de los modelos entrenados con la base de datos balanceada mediante el método SMOTE. El modelo *GFM-(EEG-ECG)* tiene el valor *validation_accuracy* más alto comparado con los demás resultados de esta métrica, pero es únicamente 0.4 % mejor que el del modelo EEG-ResNet-5 (0,950). Sin embargo, el mejor modelo utiliza el conjunto completo de electrodos EEG y ECG en el método basado en concatenación de los mapas de características globales. Por el contrario, el modelo ECG-ResNet-7 dispone de un electrodo de ECG en la *II*-derivación y obtiene un resultado en validación de 0,918, que es 3.2 % menor que el del mejor modelo. Adicionalmente, el modelo EMG-ResNet-9 contempla una mayor cantidad de bloques convolucionales que el modelo de EEG, pero obtiene un valor *validation_accuracy* menor (0,862) en comparación con los resultados de los modelos que emplean datos de EEG.

Por otro lado, el modelo EEG-[10]-ResNet-5 tiene un valor *validation_accuracy* 0.1 % mejor que el del modelo EEG-ResNet-5. Todos los demás modelos en las reducciones de electrodos EEG presentan valores *validation_accuracy* superiores a los resultados de los modelos que emplean datos de EMG, los cuales disponen de dos electrodos. Además, el modelo *GFM-model*, configurado mediante la concatenación de los mapas de características, tiene un *validation_accuracy* de 0.922, siendo 2.8 % menor comparado con este parámetro en el modelo EEG-ResNet-5. El valor accuracy de validación en el modelo Telemetry-ResNet-6 es aproximadamente 1.7 % menor en comparación con el resultado del modelo ECG-ResNet-7 y es 1.3 % mayor que el resultado del modelo FSR-ResNet-6. Adicionalmente, el modelo *GFM-(EEG-EMG)* utiliza un combinación de 18 electrodos y presenta un valor accuracy de validación que es 6.1 % menor que el resultado de esta métrica en el modelo EEG-ResNet-5 que dispone de 16 electrodos. Por otro lado, cada data-set se utilizó con la tasa de muestreo original porque los modelos de clasificación no requieren recursos informáticos excesivos en su entrenamiento en comparación con los modelos de pronóstico en este mismo método.

Adicionalmente, en la Tabla 4.4 se resumen los resultados obtenidos por los modelos de clasificación entrenados mediante la base de datos balanceada con el método SMOTETomek (oversampling y undersampling). Estos modelos presentan resultados con al menos un 16.6 % menores que los resultados de los mismos modelos entrenados con la base de datos que tiene el doble de muestras. De igual manera, el modelo EMG-ResNet-9 tiene el valor accuracy de validación más bajo (0.687) de todos los modelos de la Tabla 4.4. Por el contrario, el modelo EEG-ResNet-5 tiene el mejor resultado de validación (0.798), mostrando un comportamiento similar a los resultados de la Tabla 4.3. Como estos modelos tienen la misma configuración que los modelos de comparación en la Tabla 4.3, se puede evidenciar que el número de muestras en la base de datos impacta en al

Tabla 4.3: Resultados en el entrenamiento del primer algoritmo para la detección del estado del conductor utilizando la base de datos balanceada mediante el método SMOTE.

Models specification	Train Accuracy	Validation Accuracy	Input shape
EEG-ResNet-5	0.959	0.950	(2000, 16)
ECG-ResNet-7	0.923	0.918	(2000, 1)
EMG-ResNet-9	0.874	0.862	(2000, 2)
FSR-ResNet-6	0.903	0.890	(20, 18)
Telemetry-ResNet-6	0.921	0.903	(4, 10)
GFM-model**	0.934	0.922	(*, *)
EEG-[10]-ResNet-5	0.958	0.951	(*, *)
EEG-[8]-ResNet-5	0.953	0.947	(*, *)
EEG-[4]-ResNet-4	0.934	0.926	(*, *)
EEG-[2]-ResNet-4	0.918	0.907	(*, *)
GFM-(EEG-ECG)	0.959	0.954	(*, *)
GFM-(EEG-EMG)	0.906	0.891	(*, *)
GFM-(ECG-EMG)	0.875	0.864	(*, *)

La métrica de Accuracy fue utilizada para todos los algoritmo entrenados. (*,*) Representa la misma dimensión de entrada para cada conjunto de datos. (**) Significa que este modelo no ha utilizado los datos de FSR y Telemetría debido a su frecuencia de muestreo. (i) En el modelo EEG-ResNet-5 se utilizó todos los electrodos del conjunto de datos EEG, pero el data-set ECG solo tiene un electrodo.

Tabla 4.4: Resultados en el entrenamiento del primer algoritmo en la detección del estado del conductor al utilizar la base de datos balanceada con el método SMOTETomek.

Models specification	Train Accuracy	Validation Accuracy	Input shape
EEG-ResNet-5	0.813	0.798	(2000, 16)
ECG-ResNet-7	0.771	0.756	(2000, 1)
EMG-ResNet-9	0.702	0.687	(2000, 2)
FSR-ResNet-6	0.741	0.725	(20, 18)
Telemetry-ResNet-6	0.736	0.706	(4, 10)
GFM-model**	0.768	0.741	(*, *)
EEG-[10]-ResNet-5	0.805	0.781	(*, *)
EEG-[8]-ResNet-5	0.783	0.774	(*, *)
EEG-[4]-ResNet-4	0.784	0.770	(*, *)
EEG-[2]-ResNet-4	0.768	0.752	(*, *)
GFM-(EEG-ECG)	0.825	0.815	(*, *)
GFM-(EEG-EMG)	0.728	0.715	(*, *)
GFM-(ECG-EMG)	0.707	0.697	(*, *)

En cada uno de estos modelos se utilizó la métrica de Accuracy para evaluar su desempeño en el entrenamiento. (*,*) En este modelo se utilizó la misma dimensión de entrada en cada conjunto de datos fisiológicos. (**) Este modelo no utilizó los datos de FSR y Telemetría debido a la frecuencia de muestreo. (i) Aquí se utilizó la base de datos balanceada mediante los métodos de oversampling y undersampling, resultando en un 55 % menor en el número de muestras.

menos un 14.6 % en el desempeño de los modelos implementados.

4.3.1.2. Algoritmo de pronóstico basado en señales

Los modelos de pronóstico se evaluaron mediante la métrica RMSE, ya que es la más precisa y ampliamente utilizada en el procedimiento de previsión (Forecasting) de datos. En este caso, la métrica está en las mismas unidades que los datos de entrada (niveles de somnolencia-DLs). En la Tabla 4.5, el modelo *ECG-model* tiene el valor RMSE de validación más bajo (0,217) en comparación con los otros modelos. Por el contrario, el modelo *EMG-model* tiene el resultado *validation_RMSE* más alto (0.382) de todos los modelos con una sola categoría de datos, aproximadamente un 40 % peor que en el modelo *EEG-model*. Además, todos los modelos que involucran datos de EMG presentan resultados RMSE de validación mayores que 0.413, que es el resultado del modelo *EEG-[2]-model*.

Por el contrario, el modelo *FSR-model* utilizó datos con una frecuencia de muestreo menor en comparación a los datos fisiológicos, pero presentó un valor *validation_RMSE* de 0,328, siendo relativamente similar al resultado del modelo *EEG-model*. Además, los datos FSR son el resultado de un conjunto de sensores no intrusivos, los cuales detectan manifestaciones externas de somnolencia en el conductor, contrario a lo que ocurre con los electrodos de EEG. El modelo *GFM-model*, basado en tres tipos de datos fisiológicos, tiene un valor RMSE de validación en 0.497, siendo un modelo menos preciso. Además, su entrenamiento consume más recursos computacionales dado que utiliza una combinación de 19 electrodos. Todos los modelos que utilizan los datos de EMG presentan resultados de RMSE más altos en comparación a los modelos con una sola clase de sensores.

Tabla 4.5: Resultados del entrenamiento de los modelos para la estimación del estado subsecuente del conductor utilizando la base de datos balanceada con el método SMOTE.

Models specification	Train RMSE	Validation RMSE	Input shape
Forecasting EEG-model	0.277	0.294	(200, 16)
Forecasting ECG-model	0.198	0.217	(200, 1)
Forecasting EMG-model	0.318	0.382	(200, 2)
Forecasting FSR-model	0.290	0.328	(20, 18)
Forecasting Telemetry-model	0.303	0.344	(4, 10)
Forecasting GFM-model**	0.451	0.497	(* , *)
Forecasting EEG-[10]-model	0.306	0.335	(* , *)
Forecasting EEG-[8]-model	0.324	0.358	(* , *)
Forecasting EEG-[4]-model	0.345	0.361	(* , *)
Forecasting EEG-[2]-model	0.386	0.413	(* , *)
Forecasting GFM-(EEG-ECG)-model	0.252	0.306	(* , *)
Forecasting GFM-(EEG-EMG)-model	0.473	0.504	(* , *)
Forecasting GFM-(ECG-EMG)-model	0.522	0.567	(* , *)

Métrica Root Mean Square Error (RMSE). (*,*) Representa la misma dimensión en datos de entrada para cada conjunto. (**) Significa que este modelo no ha utilizado los datos de FSR y Telemetría. (i) El modelo Forecasting EEG-ResNet-5 empleó el conjunto de datos compuesto por los 16 electrodos, mientras que el conjunto de datos ECG solo tiene un electrodo, siendo un 27 % mejor.

Adicionalmente, la Figura 4.9 describe un segmento de señal ECG pronosticado (color azul) y los datos verdaderos (rojo) utilizados para evaluar el desempeño del modelo. Los datos estimados

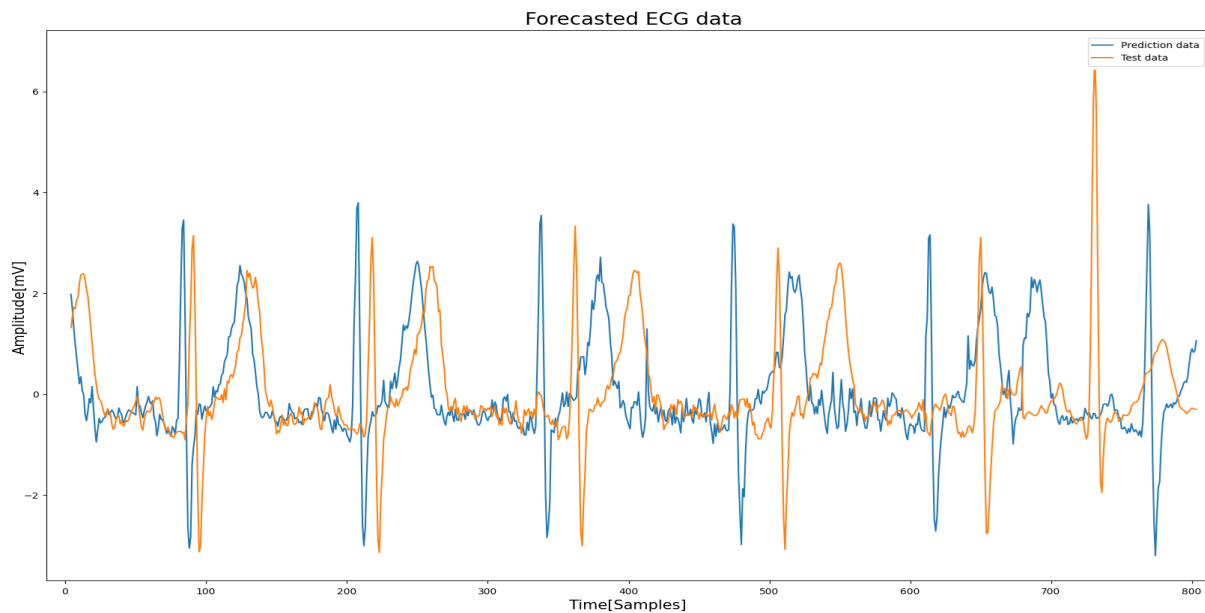


Figura 4.9: Segmentos de señal pronosticados y datos verdaderos (objetivo) a la salida del algoritmo de pronóstico, el cual utiliza múltiples muestras históricas para estimar un segmento de señal ECG a 200 Hz en frecuencia de muestreo.

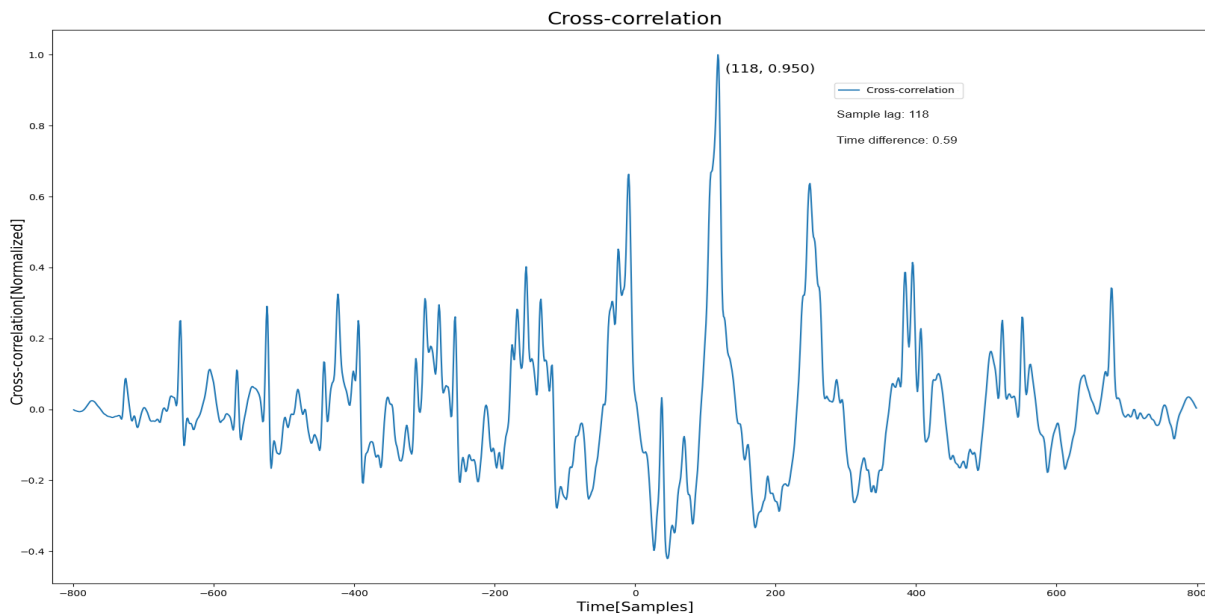


Figura 4.10: Resultados de correlación cruzada entre el segmento de valores verdaderos y los datos estimados de un segmento de señal ECG. Se aprecia que existe un grado de similitud de 0.950 con un desfase de 118 muestras entre estos segmentos.

muestran que se preserva bien la forma de la señal ECG original, indicando que el modelo puede generar un segmento de señal basado en los datos de entrada con 180 segundos (36000 muestras). La Figura 4.10 muestra el grado de precisión en la señal pronosticada y la similitud entre estas dos señales, donde se obtiene un valor de correlación cruzada de 0.950 con un desfase de 118 muestras. La Figura 4.11 representa un segmento pronosticado y los datos verdaderos de la señal EMG-1 y la Figura 4.12 corresponde a la estimación de los valores de correlación cruzada entre los datos verdaderos y estimados, donde se obtiene un grado de similitud de solo 0.750 con 70 muestras de

desfase, siendo este el modelo menos preciso en comparación con los demás. Asimismo, la Figura 4.13 ilustra un segmento estimado de señal EEG en el electrodo FP1, donde se conserva la similitud con los datos verdaderos. En la Figura 4.14 se aprecia que el grado de similitud es 0.875 cuando hay cuatro muestras de desfase entre la señal original y los puntos de predicción (color azul). Esto sugiere una baja relación entre las dos señales y una mayor dispersión de muestras en el lado positivo alrededor del punto central de correlación en comparación con el modelo *ECG-model*, que tiene menos puntos de dispersión, pero un mayor desfase entre las señales verdadera y estimada.

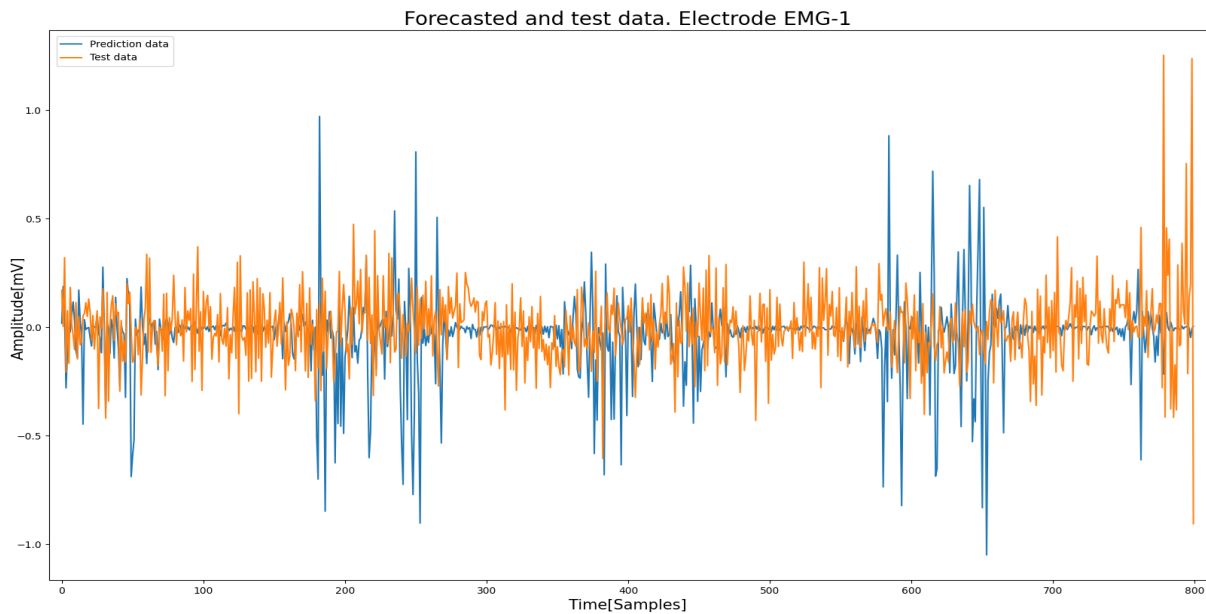


Figura 4.11: Este modelo de pronóstico es el menos preciso comparado con los demás, aún cuando utiliza 36000 muestras históricas para estimar un segmento de señal EMG de 800 muestras. Esto se relaciona con el nivel de ruido presente en los datos EMG debido al movimiento del electrodo al registrar la actividad muscular, que es de muy baja amplitud.

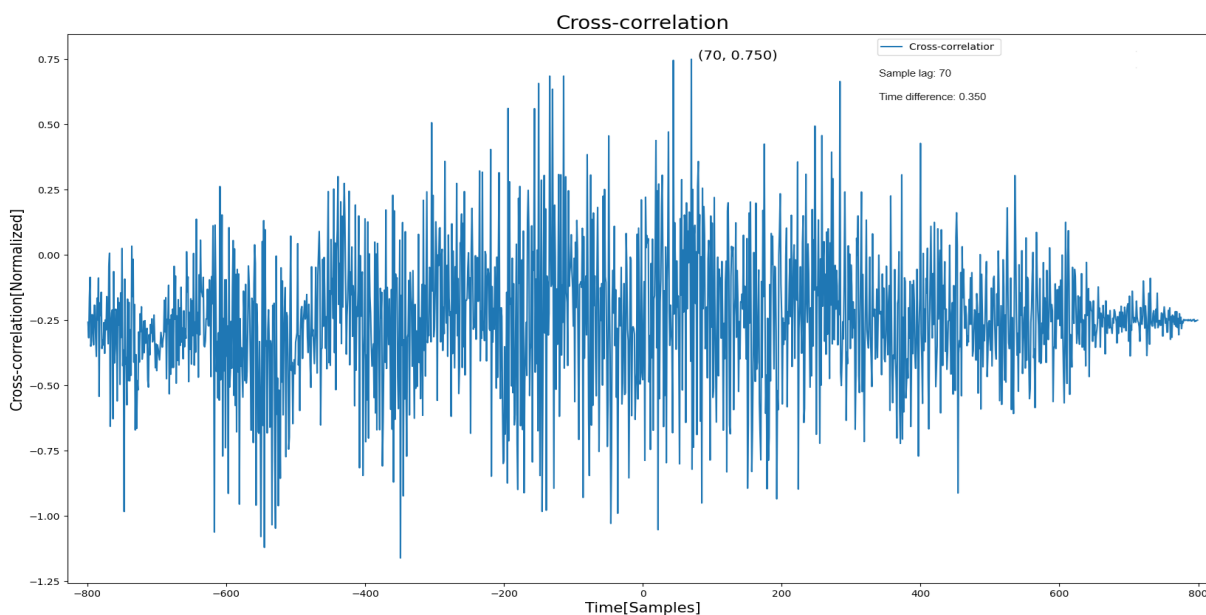


Figura 4.12: Representación de la correlación cruzada entre la señal estimada de EMG-1 y los datos verdaderos, donde se aprecia un grado de similitud de 0.750 con 70 muestras de desfase entre los segmentos y una amplia dispersión de los datos.

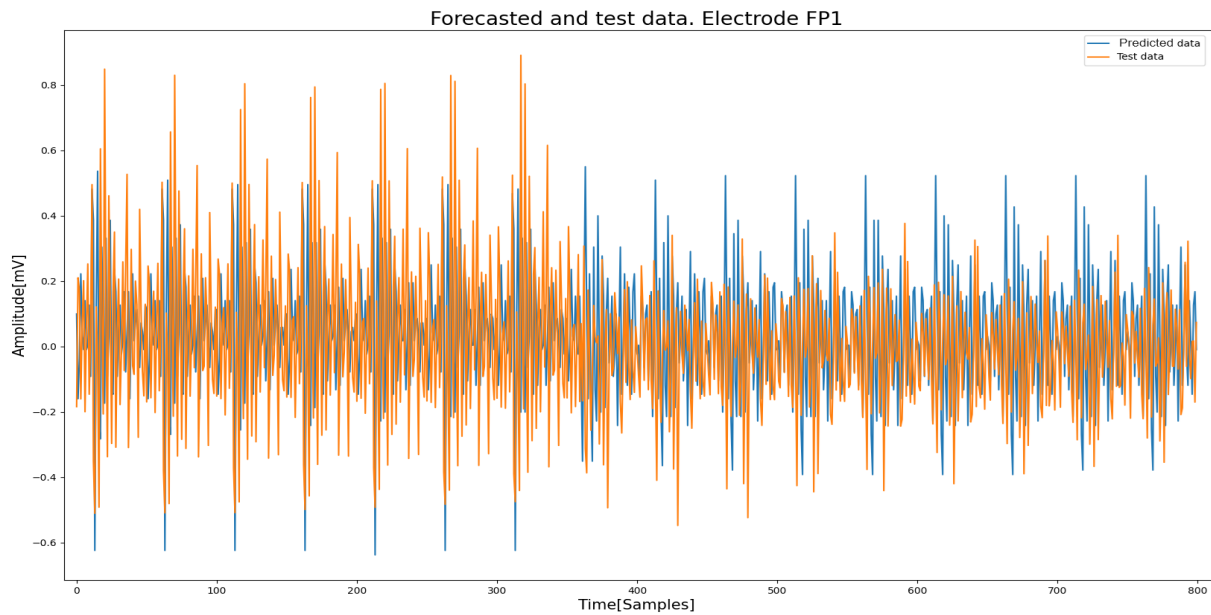


Figura 4.13: Estimación de un segmento del electrodo EEG-FP1, donde los datos pronosticados (color azul) conservan bien la forma de la señal objetivo.

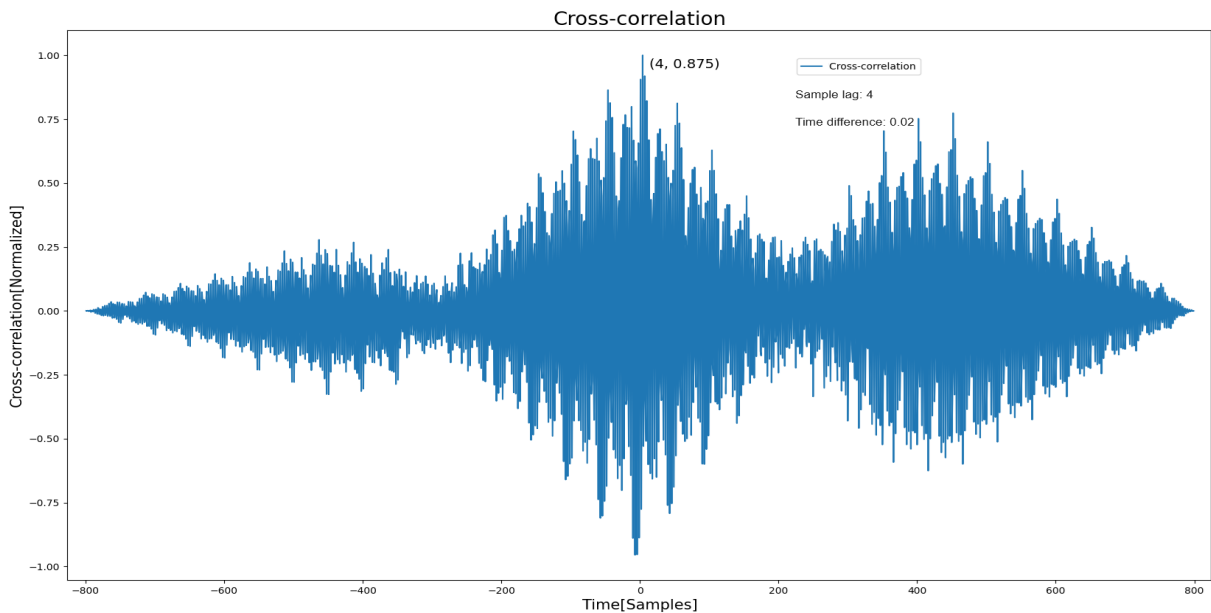


Figura 4.14: Correlación cruzada entre el segmento objetivo y los datos estimados en la señal EEG-FP1. Se aprecia que existe un grado de similitud de 0.875 con 8 milisegundos de retraso.

Por otro lado, la Tabla 4.6 muestra los resultados de pronóstico de los mismos modelos entrenados con data-sets balanceados mediante el método SMOTETomek. Se evidencia que esta base de datos incrementa un 0.187 (nivel de somnolencia) los resultados del modelo *ECG-model*, el cual alcanza un valor mínimo de 0.404, siendo mayor que la mitad (0.34) de un nivel en la escala DLs. Los resultados de la métrica RMSE de validación del modelo *FSR-model* también son al menos 1.4 veces más altos que los resultados del mismo modelo obtenidos al utilizar la base de datos más grande, 0.453 y 0.328, respectivamente. El modelo *GFM-(ECG-EMG)-model* tiene el peor resultado RMSE de validación

con un 0.733, que es superior a un nivel de la escala DLs. Además, todos los modelos que utilizan los datos de EMG tiene los peores resultados comparados con los demás modelos, donde los valores RMSE de validación son mayores que la mitad de un nivel (2/3) en la escala DLs. Los submodelos entrenados en la Tabla 4.6 tienen las mismas configuraciones que las utilizadas en los submodelos de la Tabla 4.5, posibilitando una comparación de sus resultados.

Tabla 4.6: Estimación del estado somnoliento del conductor mediante la base de datos balanceada con el método SMOTETomek.

Models specification	Train RMSE	Validation RMSE	Input shape
Forecasting EEG-model	0.426	0.451	(200, 16)
Forecasting ECG-model	0.381	0.404	(200, 1)
Forecasting EMG-model	0.482	0.526	(200, 2)
Forecasting FSR-model	0.409	0.453	(20, 18)
Forecasting Telemetry-model	0.453	0.488	(4, 10)
Forecasting GFM-model**	0.537	0.565	(* , *)
Forecasting EEG-[10]-model	0.487	0.513	(* , *)
Forecasting EEG-[8]-model	0.493	0.525	(* , *)
Forecasting EEG-[4]-model	0.525	0.549	(* , *)
Forecasting EEG-[2]-model	0.573	0.595	(* , *)
Forecasting GFM-(EEG-ECG)-model	0.463	0.488	(* , *)
Forecasting GFM-(EEG-EMG)-model	0.673	0.707	(* , *)
Forecasting GFM-(ECG-EMG)-model	0.694	0.733	(* , *)

(* , *) Significa que en este modelo se utilizó la misma dimensión en los datos de entrada de cada conjunto, mientras que el modelo fue evaluado con la métrica RMSE. (**) Este modelo no ha utilizado los datos de FSR y Telemetría. (i) Aquí se utilizó todos los electrodos en cada conjunto de datos, pero la base de datos se balanceó mediante los métodos undersampling más oversampling, lo cual reduce en un 45 % el tamaño de cada data-set.

En relación con el tiempo de subsecuentes eventos somnolientos, la Tabla 4.7 detalla un rango de este tiempo con los valores de la métrica *test_accuracy* implementada para evaluar el sistema configurado por el par de modelos *clasificación-pronóstico*. Además, la Figura 4.15 muestra la evolución de este parámetro a lo largo del tiempo de la ventana FTW. A medida que el evento somnoliento subsecuente está más cerca de siete minutos, la exactitud de todos los sistemas disminuye. Los sistemas *Forecasting_ECG/ECG-ResNet-7* y *Forecasting_EEG[10]/EEG[10]-ResNet-5*, basados en los datos de ECG y EEG, alcanzaron los mejores valores *test_accuracy* a los siete minutos, llegando a valores finales de 0.734 y 0.715, respectivamente. El sistema *Forecasting_ECG/ECG-ResNet-7* alcanzó el mejor valor *test_accuracy* (0.785) para el rango [3 a 5] minutos. Este sistema es el único que mantiene el valor *test_accuracy* durante el rango [0 a 3] minutos con un valor de 0.784.

Por otra parte, todos los sistemas que involucran datos de EMG tienen los valores *test_accuracy* más bajos en comparación a los demás sistemas. Esto es más evidente en el sistema *Forecasting_GFM(ECG-EMG)/GFM(ECG-EMG)*, el cual tiene un valor *test_accuracy* de 0.680 en el rango [0 a 1] minuto y cae hasta 0.604 al final de la ventana FTW. Finalmente, el sistema *Forecasting_FSR/FSR-ResNet-6* alcanzó un valor *test_accuracy* de 0.705 en la ventana de [1 a 3] minutos. Este valor está 4.2 % por encima del resultado *test_accuracy* (0.663) del modelo con datos de Telemetría, pero el resultado en el modelo con datos FSR desciende hasta 0.634 a los siete minutos.

Por otro lado, los modelos que involucran los datos de EEG con más de cuatro electrodos,

Tabla 4.7: Resultados de precisión en los diferentes sistemas para el tiempo de pronóstico de subsecuentes eventos somnolientos.

Names of the systems (Forecasting/Classification)	Test Accuracy per time range			
	[0 to 1 minute]	[1 to 3 minutes]	[3 to 5 minutes]	[5 to 7 minutes]
(Forecasting_EEG)/(EEG-ResNet-5)	0.773	0.766	0.758	0.715
(Forecasting_ECG)/(ECG-ResNet-7)	0.784	0.784	0.785	0.734
(Forecasting_EMG)/(EMG-ResNet-9)	0.708	0.678	0.642	0.603
(Forecasting_FSR)/(FSR-ResNet-6)	0.741	0.739	0.705	0.662
(Forecasting_Tel.)/(Tel.-ResNet-6)	0.712	0.702	0.663	0.634
(Forecasting_GFM)/(GFM-model)**	0.719	0.698	0.702	0.663
(Forecasting_EEG-[10])/(EEG-[10]-ResNet-5)	0.771	0.767	0.751	0.714
(Forecasting_EEG-[8])/(EEG-[8]-ResNet-5)	0.762	0.752	0.723	0.683
(Forecasting_EEG-[4])/(EEG-[4]-ResNet-4)	0.743	0.740	0.725	0.675
(Forecasting_EEG-[2])/(EEG-[2]-ResNet-4)	0.759	0.738	0.715	0.677
(Forecasting_GFM(EEG-ECG))/(GFM(EEG-ECG))	0.738	0.720	0.709	0.686
(Forecasting_GFM(EEG-EMG))/(GFM(EEG-EMG))	0.698	0.696	0.665	0.641
(Forecasting_GFM(ECG-EMG))/(GFM(ECG-EMG))	0.680	0.654	0.635	0.604

(**) Significa que este modelo no ha utilizado los datos de FSR o Telemetría. (i) La primera columna representa la combinación de los modelos de pronóstico y clasificación. (ii) Se Utilizó la métrica *test_accuracy* para evaluar el rendimiento de cada sistema, utilizando los datos de prueba reservados en cada data-set balanceado mediante el método SMOTE.

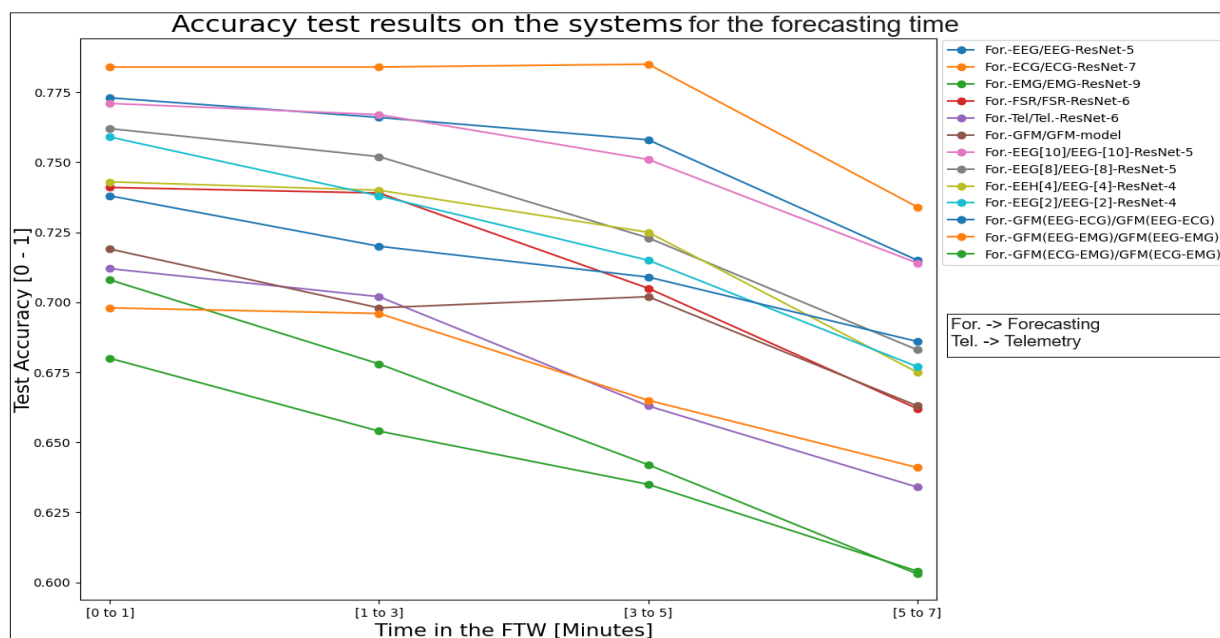


Figura 4.15: (i) El sistema *Forecasting-ECG/ECG-ResNet-7* es el único que mantiene el resultado de la métrica *test_accuracy* durante los primeros cinco minutos de la ventana FTW. (ii) todos los sistemas que utilizan datos de EMG presentan los resultados, lo cual es debido a la precisión de su respectivo modelo de pronóstico.

presentan los mejores valores *test_accuracy* en el rango [3 a 5] minutos. Por ejemplo, el modelo *Forecasting_EEG-[10]-EEG-[10]-ResNet-5* tiene un valor de 0.765 en el primer minuto y llega a 0.758 en el rango [0 a 5] minutos. Es de señalar que, el desempeño de todos los sistemas empeora conforme aumenta el rango de tiempo en la ventana FTW, siendo el intervalo de siete minutos el que perores valores *test_accuracy* tiene. Estos rangos de evaluación fueron tomados por conveniencia dado que la ventana temporal es de solo siete minutos. Sin embargo, el número de muestras en los datos de prueba están distribuidas de manera uniforme en todo el intervalo de los siete minutos. Este conjunto de datos no se utilizó en ninguna otra prueba, incluyendo los subconjuntos de entrenamiento y validación. Además, el valor *test_accuracy* se calcula como la fracción de los

valores pronosticados correctos entre el número total de elementos estimados con cada data-set.

4.3.2. Método 2: algoritmos de regresión y pronóstico

En esta sección se presentan los resultados del método configurado con un algoritmo de regresión para detectar el estado somnoliento en el tiempo t_0 y un algoritmo de pronóstico para monitorear al conductor en el rango temporal ($t_0 \rightarrow t_{0+7}$). No obstante, esta es una tarea altamente desafiante puesto que implica un proceso de pronóstico del estado somnoliento futuro del conductor para los subsecuentes siete minutos. Como entrada del segundo algoritmo se emplean los resultados del modelo de regresión, correspondientes a tres niveles de somnolencia en una escala de (0 a 2) con un nivel de alerta y dos de somnolencia.

4.3.2.1. Modelos basados en algoritmos de regresión

En este método, para cada categoría de datos disponible se entrenó un modelo de regresión identificado como ISM, resultando en cinco submodelos; dos modelos por combinación de submodelos (Ensemble modeling) y dos modelos por concatenación de mapas de características (GFM-model), Tabla 4.8. En su fase de entrenamiento, estos modelos fueron evaluados mediante los métodos MSE y *test_accuracy*, utilizados para evaluar qué tan bien funcionan los modelos con los datos de prueba después de entrenar cada ISM. El MSE calcula la diferencia cuadrática promedio entre los datos objetivo y el resultado del modelo. Se considera que el modelo funciona mejor cuando el valor MSE es lo más cercano posible a cero. Después de entrenar cada modelo, se aplicó la prueba *test_accuracy*, utilizando su respectivo conjunto de datos reservados de prueba. La Tabla 4.8 presenta el rendimiento obtenido por los modelos cuando se entrenaron utilizando los conjuntos de datos balanceados mediante el método SMOTE (oversampling).

Para todos los ISMs, el valor MSE de validación es inferior a la mitad de un nivel DLs (2/3). El modelo *EEG-ISM* tiene el mejor resultado MSE de validación (0.11), el mejor valor *test_accuracy* (0.964) y es 3.8 % mejor que el modelo *ECG-ISM* que tiene un resultado de 0.926. Sin embargo, el modelo basado en datos de EEG utiliza los 16 electrodos con segmentos de cuatro segundos. Esta dimensionalidad en el conjunto de datos tiene el potencial para proporcionar el mejor rendimiento del modelo al disponer de una mayor variabilidad en la información. Por otro lado, el modelo *EMG-ISM* tiene un alto valor MSE de validación (0.31) y un bajo valor *test_accuracy* (0.867) al igual que el modelo de Telemetría que tiene un valor *test_accuracy* de 0.869 y un resultado *validation_MSE* de 0.30. El conjunto de datos EMG tiene sólo dos electrodos que se ubicaron en la parte inferior de la pierna derecha. Por tanto, no detectan directamente patrones de somnolencia, sino señales de fatiga física que pueden incrementar los estados somnolientos del participante. En cuanto al resultado del modelo *Telemetry-ISM*, este se atribuye a su conjunto de datos que tiene una frecuencia de muestreo en solo 1 Hz, generando una menor concentración de patrones complejos de información.

Adicionalmente, el modelo *Ensemble Modeling* tiene un valor *test_accuracy* de 0.902 y es 3.4 % inferior que el resultado de la misma métrica en el *GFM-model* (0.936), entrenado con varias categorías diferentes de datos, mientras que, cada ISM se entrenó de forma independiente con su respectivo data-set. Por otro lado, en el modelo *GFM-model* que incluye datos de EEG y ECG el valor *test_accuracy* es 0.936, lo cual es una mejora del 0.5 % respecto del modelo *GFM-model* (0.941). Aún así, este primer modelo es un 2.3 % inferior que el valor *test_accuracy* logrado en el modelo *EEG-ISM* (0.964). El valor MSE de validación en el *GFM-model*(EEG, ECG) es

Tabla 4.8: Resultados de la detección del estado del conductor mediante la base de datos balanceada con el método SMOTE.

Model Name	Train MSE	Validation MSE	Test Accuracy	Input Shape	Model Size [MB]
EEG-ISM	0.10	0.11	0.964	(2000, 16)	(31.42)
ECG-ISM	0.14	0.13	0.926	(2000, 1)	(1.69)
EMG-ISM	0.28	0.31	0.867	(2000, 2)	(3.38)
FSR-ISM	0.24	0.26	0.905	(20, 16)	(1.35)
Telemetry-ISM	0.27	0.30	0.869	(4, 10)	(0.91)
GFM-Model	0.26	0.28	0.936	(-, 304640)	(55.89)
Ensemble Modeling	-	-	0.902	(*, *)	(0.45**)
GFM-Model (EEG, ECG)	0.19	0.20	0.941	(-, 268800)	(39.78)
Ensemble Model (EEG, ECG)	-	-	0.937	(*, *)	(0.25**)

Métrica Mean Square Error (MSE). (**) Corresponde únicamente a los modelos ensamblados. (*) Se relaciona con la misma dimensión de entrada de cada conjunto de datos. (i) El modelo EEG empleó un conjunto de datos que consta de 16 electrodos, mientras que el conjunto de datos de ECG solo tiene la *II*-derivación. (ii) El conjunto de datos de prueba se utilizó después de entrenar cada ISMs para calcular la métrica *test_accuracy* y evaluar qué tan bien el modelo clasificó las tres clases con datos no utilizados. Por otro lado, se calculó la métrica de MSE durante la fase de entrenamiento de los modelos.

0.38 y está por debajo de la mitad de un nivel en la escala DLs (0.67). Además, el modelo (*Ensemble-model*(EEG,ECG) tiene un valor *test_accuracy* de (0.937), siendo un 3.5 % mejor que el *Ensemble Modeling* (0,902). Sin embargo, estos modelos se entrenaron utilizando los conjuntos de datos más representativos (EEG, ECG). Finalmente, el valor *test_accuracy* del modelo *Ensemble Model*(EEG, ECG) es 0.937 y está un 2.7 % por debajo que el resultado de esta métrica en el modelo *EEG-ISM* que utiliza 16 electrodos.

Tabla 4.9: Resultados en la detección del estado del conductor mediante la base de datos balanceada con el método SMOTETomek.

Model Name	Train MSE	Validation MSE	Test Accuracy	Input Shape	Model Size [MB]
EEG-ISM	0.28	0.32	0.850	(2000, 16)	(31.42)
ECG-ISM	0.35	0.39	0.761	(2000, 1)	(1.69)
EMG-ISM	0.38	0.42	0.751	(2000, 2)	(3.38)
FSR-ISM	0.39	0.42	0.772	(200, 16)	(1.35)
Telemetry-ISM	0.40	0.44	0.725	(4, 10)	(0.91)
GFM-Model	0.37	0.41	0.815	(-, 304640)	(55.89)
Ensemble Modeling	-	-	0.791	(*, *)	(0.45**)
GFM-Model (EEG, ECG)	0.33	0.38	0.824	(-, 268800)	(39.78)
Ensemble Model (EEG, ECG)	-	-	0.806	(*, *)	(0.25**)

Métrica Mean Square Error (MSE). (**) Aplica para los modelos ensamblados. (i) El MSE de validación es mayor que el MSE de entrenamiento, lo que es consistente con un sobreajuste como resultado de la menor cantidad de datos disponibles.

De acuerdo con los resultados de los ISMs que utilizaron la base de datos balanceada mediante el método SMOTETomek (undersampling y oversampling), se observó que todos estos modelos tienen un efecto de sobreajuste, Tabla 4.9. Además, el modelo *EEG-ISM* tuvo un rendimiento de aproximadamente 11 % peor que el mismo modelo entrenado con el conjunto de datos más grande SMOTE (oversampling). Cada ISM revela una reducción de por lo menos 9 % en comparación con los mismos modelos de la Tabla 4.8. El desempeño de estos modelos demuestra que el valor MSE de validación es al menos la mitad de un nivel de la escala DLs (0.67). Es de señalar que la dimensión de los datos de entrada y los parámetros globales de los modelos permanecieron sin cambios para facilitar una comparación de los resultados obtenidos con estas dos bases de datos.

4.3.2.2. Algoritmo de pronóstico basado en niveles de somnolencia

El modelo de pronóstico implementado (*DLs-model*) utiliza como entrada los 105 valores históricos de la ventana HTW y estima siete valores según la escala Drowsiness Levels (DLs). Luego, calcula el promedio de estos 7-DLs, evalúa si este resultado excede el primer umbral de somnolencia (*drowsiness-1*) de $2/3$ y mide el tiempo FTD-1 del subsiguiente evento somnoliento. Al estimar tal parámetro, esta técnica es completamente diferente a la implementada en el método uno. Como el *DLs-model* estima 7-DLs, establece una resolución temporal FTD-1 de 28 segundos en la ventana FTW, pero con una resolución temporal de entrada fijada en cuatro segundos. Se utilizó la métrica MSE para evaluar el desempeño de este modelo; se mide la diferencia entre la salida del modelo y el estado verdadero (objetivo). Cuanto más cerca esté el MSE de cero, mejor será el rendimiento del modelo, lo que se traduce en estimaciones más precisas.

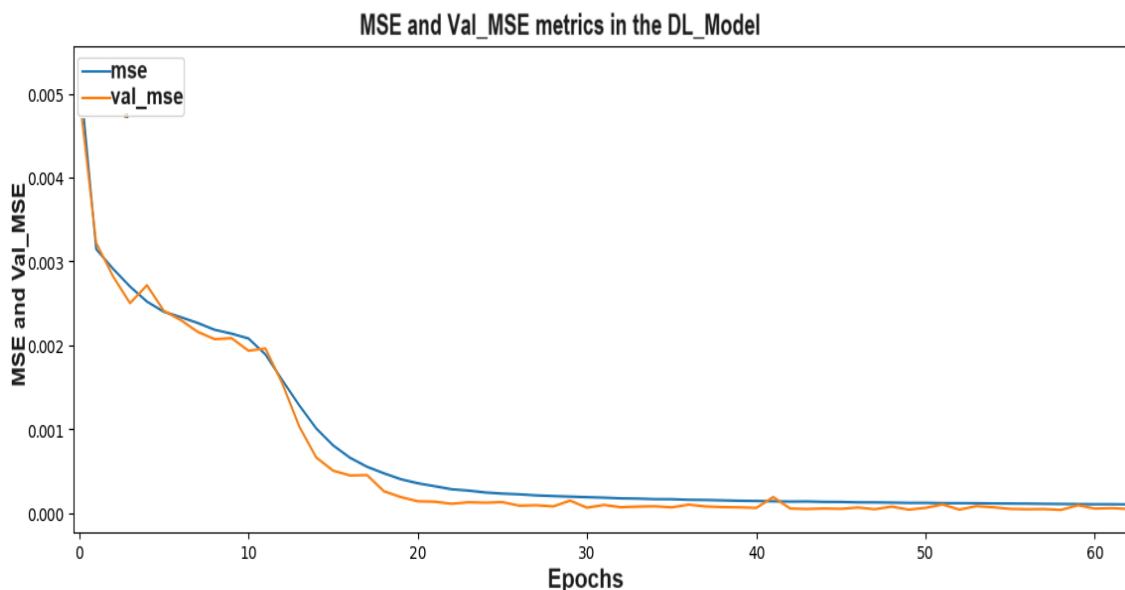


Figura 4.16: El modelo de pronóstico funciona notablemente bien usando una ventana HTW con al menos 105 valores, donde el MSE tiene las mismas unidades que de los valores de entrada al modelo (DLs).

En la Figura 4.16 se aprecian los resultados de la métrica MSE de entrenamiento y validación del modelo. El valor más bajo en la etapa de validación es aproximadamente $5 \cdot 10^{-4}$, lo cual está muy por debajo de la mitad de un nivel de somnolencia ($2/3$). Además, el gráfico muestra que el modelo converge muy bien y no tiene sobreajuste significativo. La Figura 4.17 corresponde a los valores de la función de pérdida del modelo; converger a cero a medida que aumentan los ciclos de entrenamiento. Por otro lado, con base en los resultados obtenidos, el algoritmo estima cualquier valor del tiempo FTD-1 dentro de la ventana FTW en pasos de 28 segundos.

Por otro lado, la Figura 4.18 muestra el valor de tiempo t_{FTD-1} de un subsecuente evento somnoliento localizado a los 322 segundos (valor intermedio de la ventana 12). En esta Figura se trazó todos los valores pronosticados para evaluar la variación del nivel de somnolencia después de calcular el tiempo t_{FTD-1} . La Figura 4.19 compara los valores DLs históricos con muestras pronosticadas normalizadas en las dos ventanas temporales, donde se aprecia que no hay un valor de tiempo t_{FTD-1} , estableciendo que el conductor no presentaría somnolencia en los subsiguientes siete minutos. Adicionalmente, en la Figura 4.20 se evidencia la variación de los valores DLs históricos y pronosticados normalizados y graficados en la misma ventana FTW con un valor de tiempo t_{FTD-1}

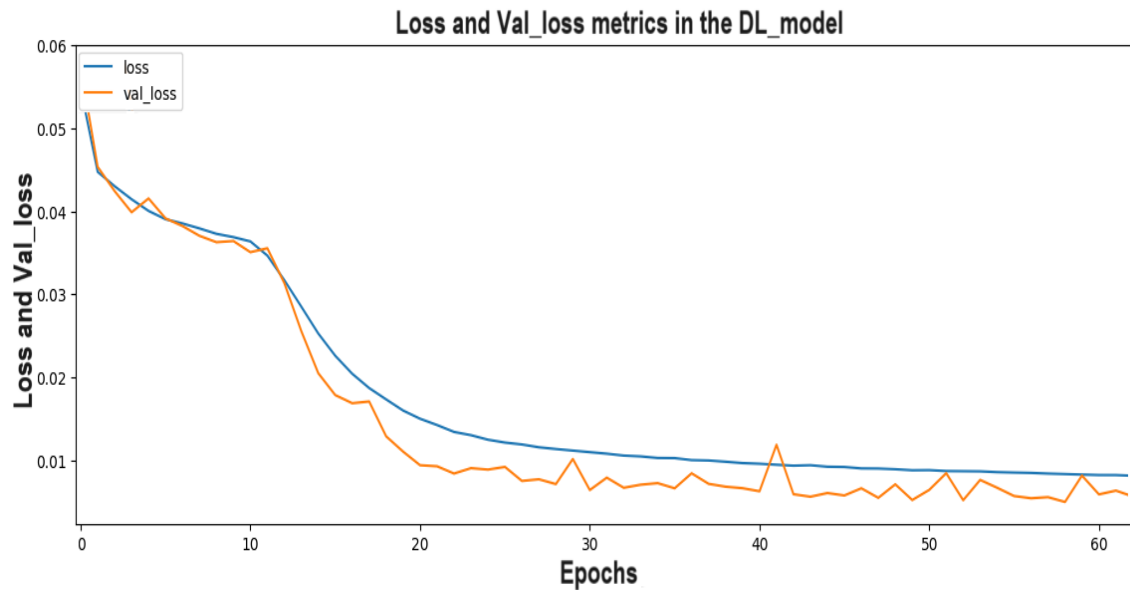


Figura 4.17: Representa la validación y el comportamiento del modelo en la etapa de entrenamiento. (i) La función de pérdida tiende a cero a medida que se incrementa el número de ciclos de entrenamiento, llegando a un valor de pérdida mínimo de 0.005.

a los 42 segundos de esta ventana, siendo este el promedio de la segunda ventana de 28 segundos. La Figura 4.21 no muestra valores de tiempo t_{FTD-1} en la ventana FTW. En este caso, el conductor no experimentaría ninguna escala de somnolencia dentro de los subsiguientes siete minutos, a menos que se presente una variación significativa en los DLs para el tiempo t_0 que modifique el subsiguiente estado del conductor.

4.20

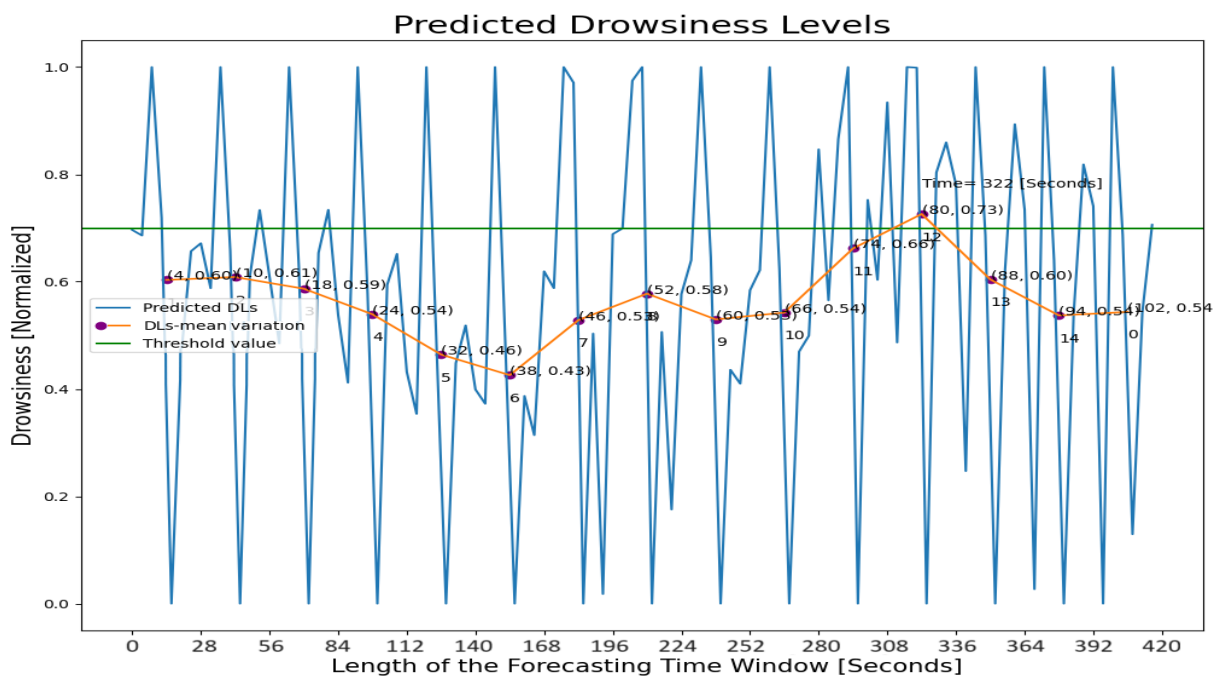


Figura 4.18: La línea de color rojo representa la variación de los valores de somnolencia pronosticados en intervalos de 28 segundos en la ventana FTW. (i) Cada punto en color púrpura representa el promedio de una ventana de 28 segundos.

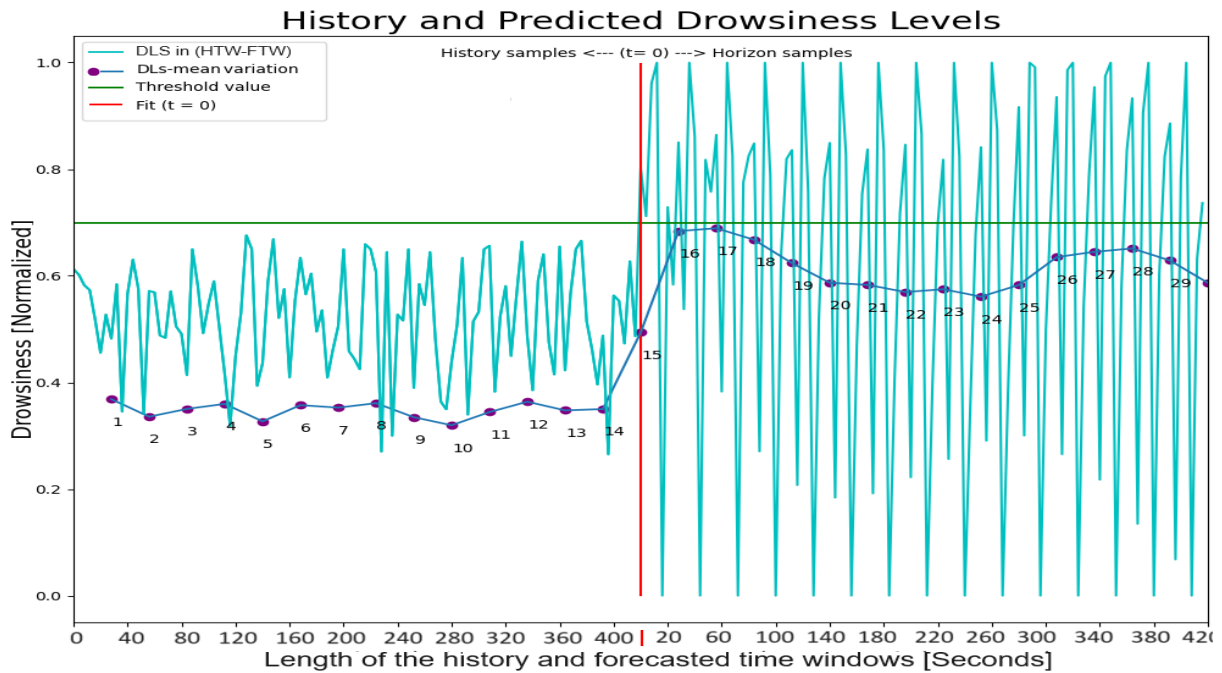


Figura 4.19: La línea vertical en color rojo indica el punto final de la ventana de muestras históricas y el comienzo de la ventana de muestras pronosticadas. (i) La línea en color azul representa la variación de los valores DLs sobre las dos ventanas temporales.

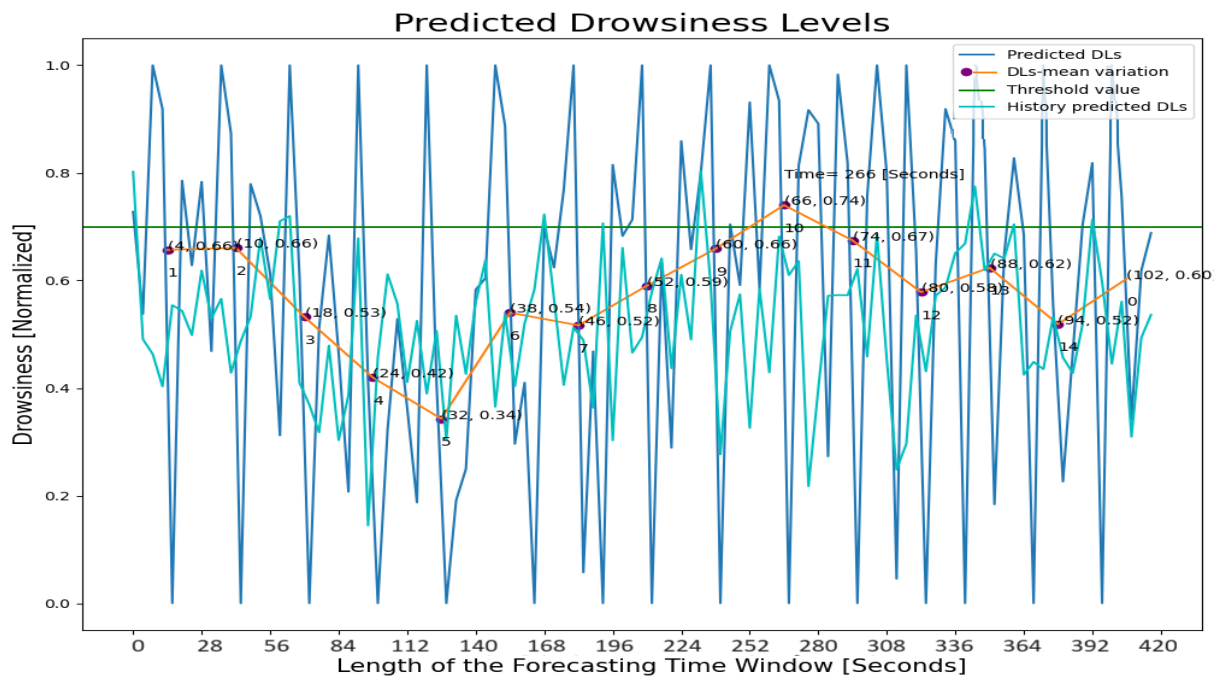


Figura 4.20: Para que sea posible una comparación, se ha trazado los valores DLs históricos y pronosticados en la misma ventana FTW. (i) El promedio de los valores 7-DLs pronosticados del conductor superó el umbral de somnolencia 2/3 a los 266 segundos (décima ventana de 28 segundos).

El modelo de pronóstico es capaz de predecir el tiempo en el cual el estado del conductor sobrepasará el umbral de somnolencia (2/3) dentro de la ventana FTW. Por lo tanto, se utilizó la métrica *test_accuracy* con los datos reservados de prueba para evaluar el desempeño de los

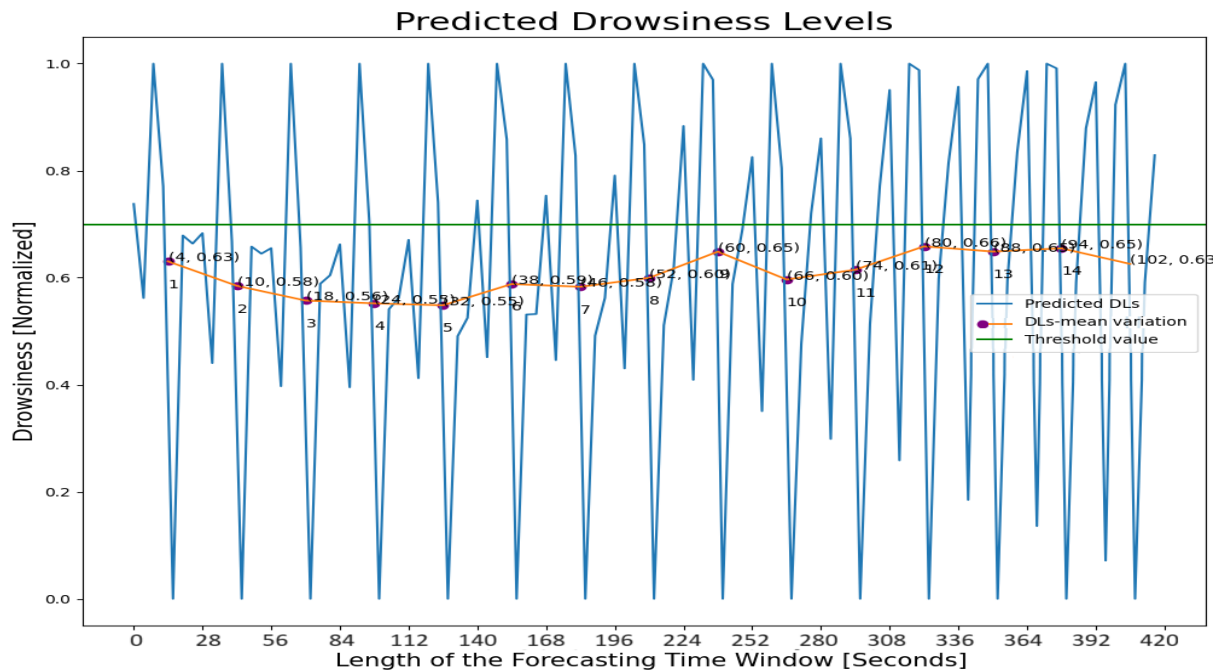


Figura 4.21: Valores DLs en la ventana temporal FTW, donde no se evidencia niveles de somnolencia mayores a $2/3$, descartando tener eventos somnolientos subsecuentes, a menos que el cambio en los valores DLs del tiempo t_0 puedan modificar los valores pronosticados.

sistemas (*regresión + pronóstico*) basados en combinaciones de los ISM, estimando así la exactitud del tiempo t_{FTD-1} en los rangos $[0 \text{ a } 5]$ y $[5 \text{ a } 7]$ minutos. El parámetro *test_accuracy* se implementó en función de los resultados del modelo de pronóstico y los valores verdaderos en los datos de prueba correspondientes a segmentos donde el conductor experimentó el primer nivel de somnolencia (drowsiness-1). Este data-set no se utilizó en las pruebas de entrenamiento ni validación. Por consiguiente, al utilizar la combinación (*EEG-ISM + DLs-model*) con 16 electrodos de EEG, se alcanza el mejor valor *test_accuracy* (0,887) para un t_{FTD-1} menor a cinco minutos. De lo contrario, el valor de esta métrica disminuye gradualmente hasta alcanzar 0,798 al final de la ventana HTW. Además, el sistema (*ECG-ISM + DLs-model*) obtiene un valor *test_accuracy* de 0.871 en un rango $[0 \text{ a } 5]$ minutos, siendo un 1.6 % menor que el valor de la combinación anterior. La configuración del sistema (*FSR-ISM + DLs-model*) alcanzó un valor *test_accuracy* de 0.844 para el tiempo FTD-1 entre $[0 \text{ a } 5]$ y un valor de 0.746 entre los $[5 \text{ a } 7]$ minutos.

Por otra parte, el sistema que utiliza los datos de Telemetría obtuvo un valor *test_accuracy* de 0.707 y 0.655 para los intervalos $[0 \text{ a } 5]$ y $[5 \text{ a } 7]$ minutos, respectivamente. Para el sistema basado en los modelos (*Ensemble Modeling + DLs-model*) se obtuvo un valor *test_accuracy* de 0.809 para $[0 \text{ a } 5]$ minutos, pero este valor cae a 0.767 entre los $[5 \text{ a } 7]$ minutos. El sistema configurado por *Telemetry-ISM + DLs-model* alcanzó un valor *test_accuracy* de 0.788 para el tiempo FTD-1 en el rango $[0 \text{ a } 5]$ minutos y un resultado de 0.751 para intervalo temporal final. En contraste a estos resultados, al utilizar la base de datos balanceada mediante el método SMOTETomek, la cual tiene únicamente el 49 % de las muestras de la base de datos balanceada mediante el método SMOTE, la mejor combinación (*EEG-ISM + DLs-model*) obtuvo un valor *test_accuracy* de 0,645 en el rango $[0 \text{ a } 5]$ minutos. Con esta misma base de datos, el sistema (*ECG-ISM + DLs-model*) obtuvo un valor *test_accuracy* de 0.615 en un rango $[0 \text{ a } 5]$ minutos y 0.567 en el intervalo $[5 \text{ a } 7]$ minutos. Los sistemas restantes tiene resultados *test_accuracy* inferiores a 0.600 entre los $[0 \text{ a } 5]$ minutos y

0.550 en el rango de [5 a 7] minutos.

Finalmente, en el método dos, la combinación de un modelo de regresión y el algoritmo DLs-model, se obtuvieron mejores resultados en la estimación del tiempo t_{FTD-1} en comparación con los resultados obtenidos en el método uno para este mismo parámetro. Además, los modelos de pronóstico en el método uno requieren mayores recursos computacionales, por ello se ejecutaron en Colab versión pro mediante el uso de una GPU de mayores capacidades (NVIDIA A100). No obstante, el algoritmo DLs-model requiere una mayor cantidad de datos y con mejor variabilidad para incrementar los valores de precisión obtenidos.

5.1. Detección de la intención de frenada de emergencia

En el mejor de los casos, al emplear imágenes en escala de grises para entrenar el modelo y detectar la intención de frenada de emergencia, se obtuvo un AUC de 0.66, donde se utilizó el grupo de 59 electrodos. Además, al emplear únicamente cuatro electrodos, este modelo solo alcanzó un valor AUC de 0.60. Es claro que, al no haber un método universal que permita obtener la mejor configuración para generar las imágenes a partir de las señales de EEG, se requiere aplicar diversos procedimientos que permitan diferenciar las imágenes *objetivo* de las clases *no-objetivo*. Por lo tanto, entrenar un modelo CNN con imágenes en escala de grises consume más tiempo de procesamiento, recursos computacionales y tiempo de entrenamiento comprado con entrenar el mismo modelo directamente con los segmentos de señal. A pesar de que los resultados (AUC= 0,82, precisión= 80 %) obtenidos al entrenar los algoritmos directamente con los segmentos de señal están por encima de los obtenidos en los trabajos [106, 112], estos valores aún son inferiores a los reportados en el análisis del estado del arte para esta clase de trabajos.

A pesar de los resultados reportados en la etapa de predicción de la intención de frenada de emergencia, la ventaja de esta técnica es la simplicidad del modelo implementado y su método tensorial de entrenamiento con solo cuatro electrodos. Aún así, mejorar el rendimiento del modelo implementado requiere incrementar el número de muestras en los datos al igual que su número de capas. Por otro lado, esta técnica de análisis puede utilizarse para detectar fatiga y somnolencia en diversos contextos donde los operadores podrían estar en riesgo. Por ejemplo, conductores de vehículos, operadores de tráfico aéreo, operadores remotos de drones, de grúas y equipos de riesgo eléctrico. Este personal se enfrenta constantemente a una amplia variación de los niveles de la carga de trabajo, impactando negativamente en el nivel de fatiga y somnolencia del operador.

5.2. Detección y pronóstico de eventos somnolientos

En esta sección se discuten los resultados obtenidos de los dos métodos orientados a monitorear la condición somnolienta en el tiempo t_0 y estimar el estado subsecuente del conductor en los siguientes siete minutos ($t_0 \rightarrow t_{0+7}$). Además, se implementaron varios submodelos con diferentes combinaciones de datos y cantidades reducidas de electrodos. Esto permitió identificar la mejor categoría de datos y electrodos necesarios para una precisión superior al 80 % en el método de

pronóstico y superior al 90 % para el modelo de clasificación.

5.2.1. Tamaño de los datos y desempeño de los modelos

Con el fin de reducir la dimensión del data-set de EEG, se implementó diferentes combinaciones con un número reducido de estos sensores, ya que en el experimento se utilizaron 16 electrodos. Esto permitió evaluar el desempeño de los sistemas que utilizaron estos datos. En el método uno, el sistema que utiliza datos de EEG con dos electrodos tiene un resultado *validation_RMSE* de 0.413, pero al combinarse en el sistema *Forecasting_EEG-[2]/EEG-[2]-ResNet-4*, se obtiene un valor *test_accuracy* de 0.715 en el rango [3 a 5] minutos. Por el contrario, al utilizar los 16 electrodos en el sistema *Forecasting_EEG/EEG-ResNet-5* se alcanzó un valor *test_accuracy* de 0.758 para el mismo rango de [3 a 5] minutos, dando un valor promedio de 0.753 para todo el intervalo de la ventana FTW. Del mismo modo, al considerar el valor *validation_accuracy* del modelo de clasificación *EEG-ResNet-5* (0.950), este resultado se ve reducido hasta en un 23.5 % en el sistema *Forecasting_EEG/EEG-ResNet-5*, llegando únicamente a un valor *test_accuracy* de 0.715 para el rango de siete minutos.

De igual manera, se evidencia una reducción del 26.8 % en el valor *validation_accuracy* del modelo de clasificación *GFM-(EEG-ECG)* que utiliza los 16 electrodos, pasando de 0.954 a un resultado de 0.686 en el sistema *Forecasting_GFM(EEG-ECG)/(GFM(EEG-ECG))* para el rango completo de la ventana FTW. Asimismo, en el modelo de clasificación *GFM-model* se obtuvo un resultado *validation_accuracy* de 0.922 y su modelo de pronóstico alcanzó un valor RMSE en validación de 0.497. No obstante, se presenta un descenso del 20.3 % en la métrica *test_accuracy* obtenida por el sistema configurado por estos dos modelos (*Forecasting_GFM)/(GFM-model)*. Esto demuestra que la combinación de diferentes categorías de datos no mejora la precisión del modelo de pronóstico, por el contrario, reduce en más del 20 % la precisión de los sistemas configurados por los modelos de (*clasificación + pronóstico*). Las señales pronosticadas de los subsecuentes estados somnolientos tienen menor fuerza de relación con los datos de la clase “alertness” o “drowsiness-1”, presentando más muestras dispersas que afectan negativamente la precisión del modelo de clasificación, ya que este algoritmo debe asignar una nueva etiqueta a los segmentos pronosticados.

Por otro lado, al evaluar de manera independiente los modelos de pronóstico, para el 73 % de los datos de prueba, las muestras estimadas tienen un resultado de correlación cruzada de al menos 0.50 como valor medio del grado de similitud entre el segmento verdadero y las muestras pronosticadas. La Figura 5.1 representa una sección de los datos en la ventana HTW, las muestras pronosticadas y los datos para prueba de este segmento. Al considerar que su modelo de pronóstico obtuvo un resultado *validation_RMSE* de 0.217 estimando los datos del segmento ECG en el tiempo t_{0+4} segundos, su modelo de clasificación un valor *validation_accuracy* de 0.918 y su sistema combinado un resultado *test_accuracy* de 0.784 en el rango de [0 a 1] minuto, su valor de correlación (0.950) demuestra un grado de similitud entre las muestras pronosticadas y el segmento verdadero, Figura 5.2. De este modo, el primer segmento pronosticado es el más precisos de todos y tiene menos muestras atípicas.

Por el contrario, entre mayor número de segmentos pronosticados se integran en la ventana HTW, mayor número de muestras atípicas hay en los subsecuentes segmentos pronosticados, llevando a que el segmento 105 sea el menos preciso de todos. Adicionalmente, en la Figura 5.3 se destaca los valores de un segmento pronosticado y los datos verdaderos en una señal de EEG para el electrodo PO7. A partir de la muestra 500, los datos estimados difieren de los valores de prueba, su resultado

de correlación muestra que no existe fuerza de relación entre estos dos segmentos, Figura 5.4. Esto es congruente al analizar los datos de correlación cruzada, evidenciando un grado de similitud máxima de 0.875 con 5 muestras de desfase, Figura 5.5. Además, si una distribución presenta una Curtosis mayor que cero, tiene colas más pesadas y un centro más marcado en comparación con la distribución normal, haciendo que los valores de los datos estén más concentrados en torno a la media y haya menos valores atípicos. El caso opuesto indica que los datos están menos concentrados alrededor de la media y hay más valores atípicos [152].

De este modo, en la Figura 5.6 se muestra que los datos pronosticados tienen un valor de Curtosis de 1.274 con una std de 0.133. Complementariamente, la Figura 5.7 destaca una Curtosis de 1.664 y un resultado std de 0.073. En consecuencia, la comparación y análisis de estos resultados indica que al incrementar el número de muestras pronosticadas en la ventana HTW, el algoritmo de pronóstico genera un mayor número de muestras atípicas lejos de la media en comparación con los datos de prueba (verdaderos).

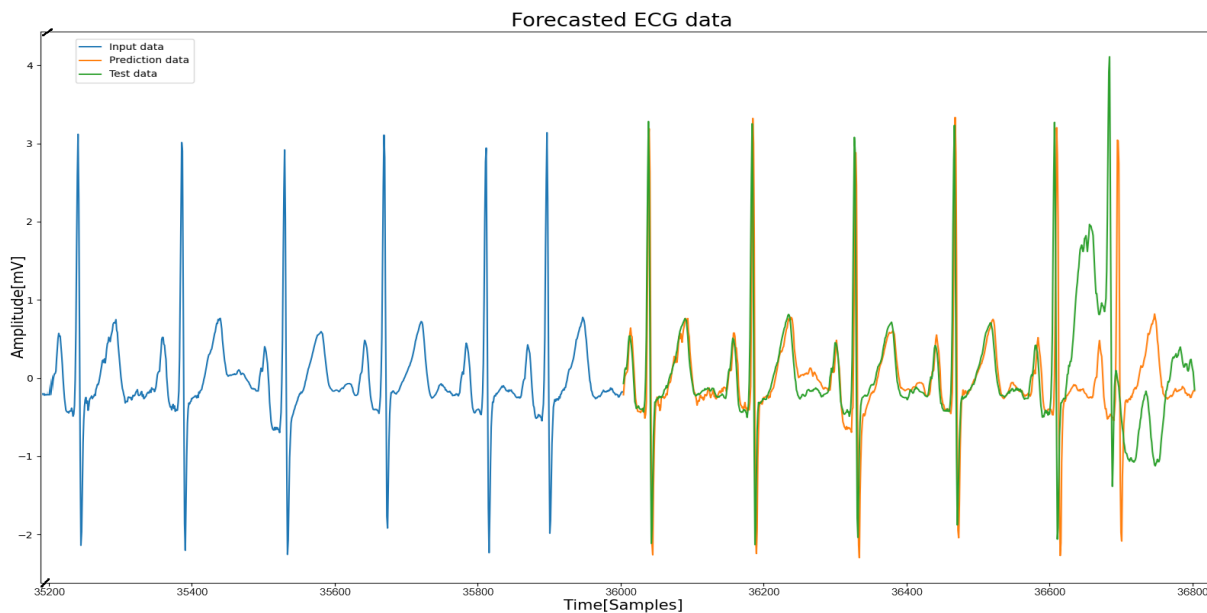


Figura 5.1: Representación de una sección de los datos de la ventana HTW de la señal ECG junto a un segmento pronosticado y los datos verdaderos. (i) Se aprecia que las últimas muestras estimadas difieren de los datos de prueba (color verde), lo que no sucede con las primeras muestras estimadas de este segmento.

En cuanto al modelo de clasificación *EEG*-[2]-*ResNet*-4, este satisface el objetivo de tener un valor *validation_accuracy* de al menos 90% con solo dos sensores. El par de electrodos se ubicaron en la región Primary Visual Cortex (PVC), la cual tiene una menor actividad dinámica cuando el conductor está somnoliento, pero es mayor en estado de alerta, Tabla 5.1. Además, esta variación es un buen indicador para explicar el estado funcional de atención [153–155]. El modelo *EEG*-[10]-*ResNet*-5 tiene al menos dos electrodos en las regiones Occipital, Parietal y Frontal, alcanzando un valor *accuracy* de validación más alto en el clasificador (0.951) y un valor *validation_RMSE* de 0.335 en el algoritmo de pronóstico. Sin embargo, la combinación de estos dos modelos tiene un valor *test_accuracy* de 71.4% para el rango de siete minutos con 10 electrodos. Por el contrario, el sistema *Forecasting_ECG/ECG-ResNet*-7 es la única combinación de modelos que tiene la menor cantidad de electrodos y un valor *test_acurracy* de 73.4% en el rango [5 a 7] minutos ya que solo se dispone de la *II*-derivación del ECG estándar.

Esta actividad neuronal focalizada indica que algunos sensores registran una mayor actividad

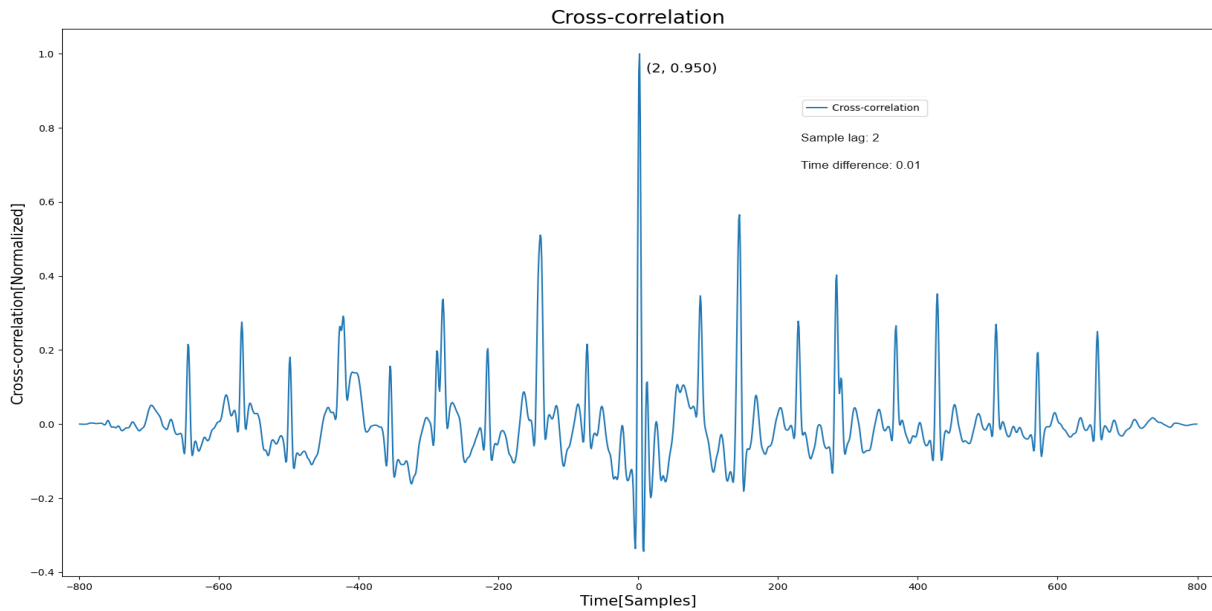


Figura 5.2: Resultados de la correlación cruzada entre el segmento de datos de prueba en la señal ECG y las muestras pronosticadas por el algoritmo. (i) El resultado de correlación no considera los valores de entrada en la ventana HTW.

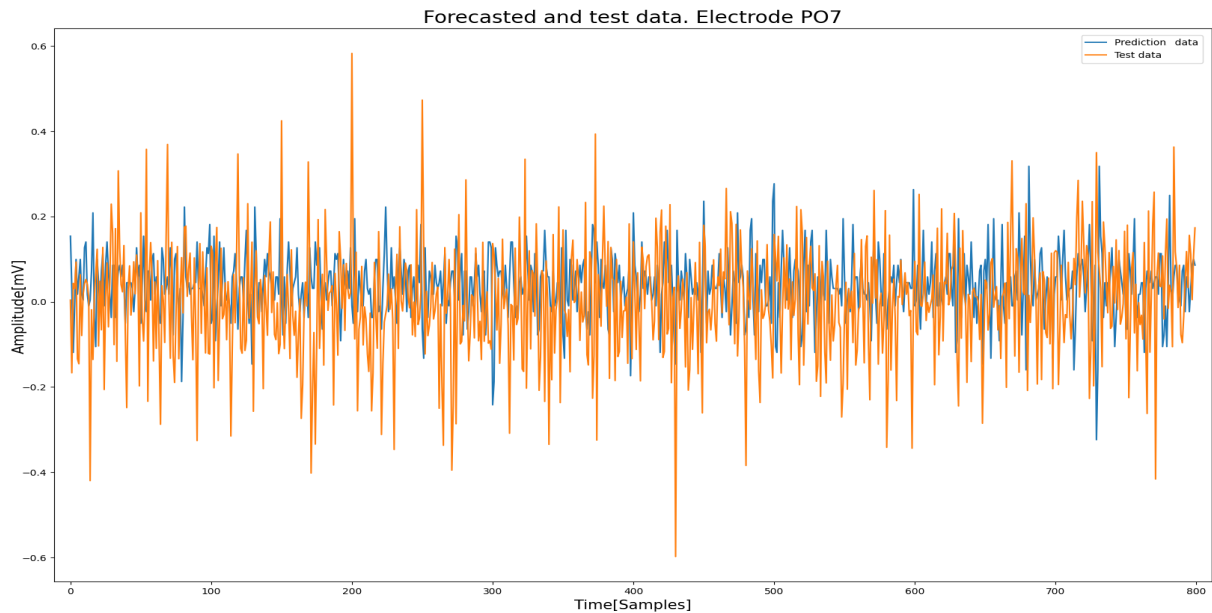


Figura 5.3: Segmento estimado de la señal EEG en el electrodo PO7 y los datos de prueba para el mismo sensor a una frecuencia de muestreo de 200Hz, correspondiente a la salida del algoritmo de pronóstico.

cerebral durante los eventos somnolientos del conductor, descartando la necesidad de utilizar el set completo de 16 electrodos de EEG. Por otro lado, el sistema *Forecasting_FSR/FSR-ResNet-6* obtiene un valor *test_accuracy* de 0.705 para el rango [3 a 5] minutos, mostrando la posibilidad de utilizar sensores no intrusivos en la detección y pronóstico de eventos subsecuentes de somnolencia del conductor.

En relación con el método dos, al comparar los resultados *validation_accuracy* de los modelos de clasificación con los valores *test_accuracy* de los modelos de regresión, estos algoritmos presentan

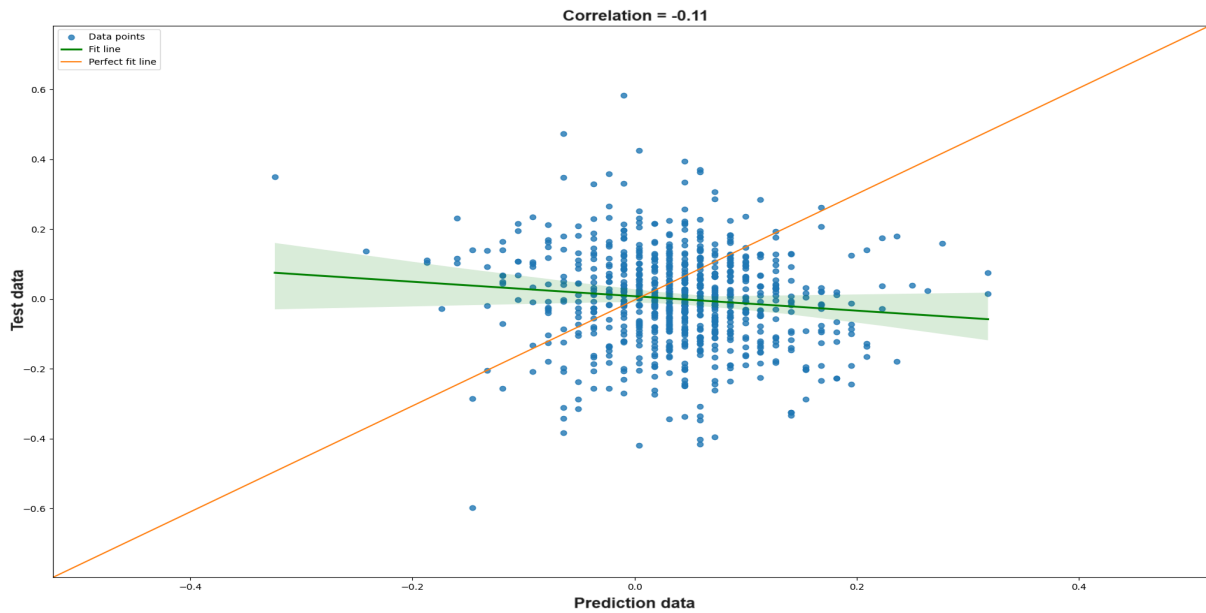


Figura 5.4: Resultados de la correlación entre el segmento de prueba y los datos pronosticados para la señal EEG en el electrodo PO7, donde el valor de la fuerza de relación entre estos dos segmentos es -0.11. (i) Estos valores no estiman el grado de similitud entre los dos segmentos utilizados, pero sí estiman su grado de relación.

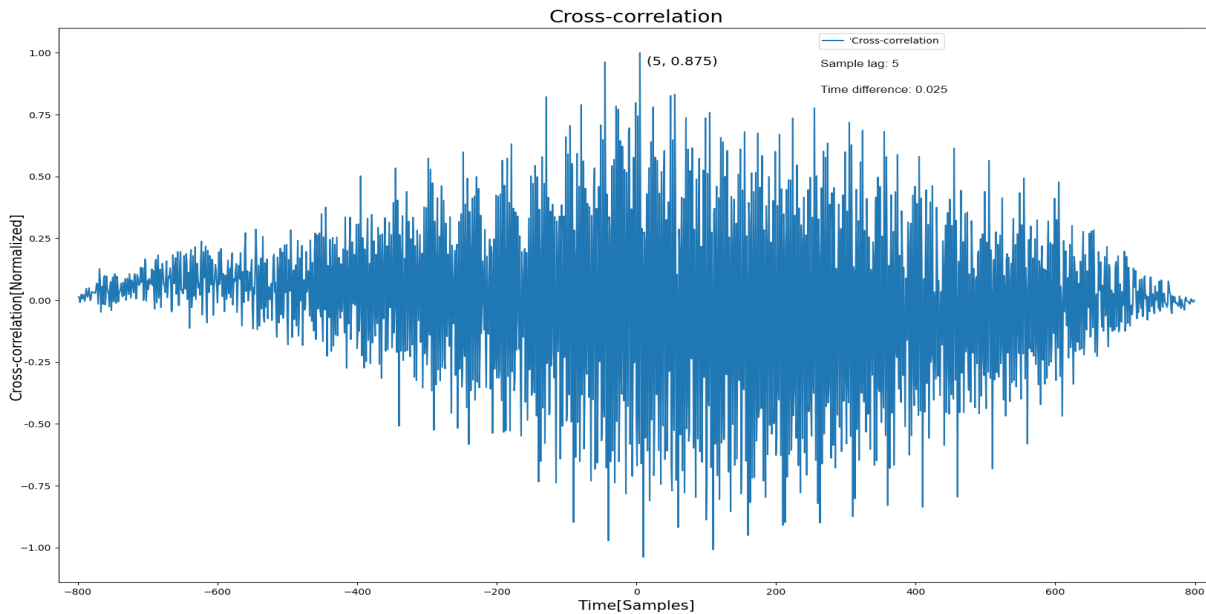


Figura 5.5: Los valores de correlación cruzada muestran que hay un desfase de 5 muestras entre la señal pronosticada y los datos de prueba, determinando un grado de similitud máxima de 0.875 entre estos dos segmentos.

resultados similares entre ellos. Así, el algoritmo de clasificación *EEG-ResNet-5* alcanzó un valor accuracy en validación de 0.950 y el modelo de regresión *EEG-ISM* obtuvo un valor de 0.964 en la métrica *test_accuracy*, generando una variación de 1.4% en el resultado de esta medición. Además, este comportamiento se repite en los modelos que utilización una combinación de datos EEG y ECG tanto en los algoritmos de clasificación como de regresión, pero esto no se cumple en los modelos que emplean datos de EMG, los cuales presentan los resultados más bajos de *validation_accuracy*. Así, los algoritmos de regresión *EMG-ISM* y *Telemetry-ISM* no cumplen el objetivo de alcanzar

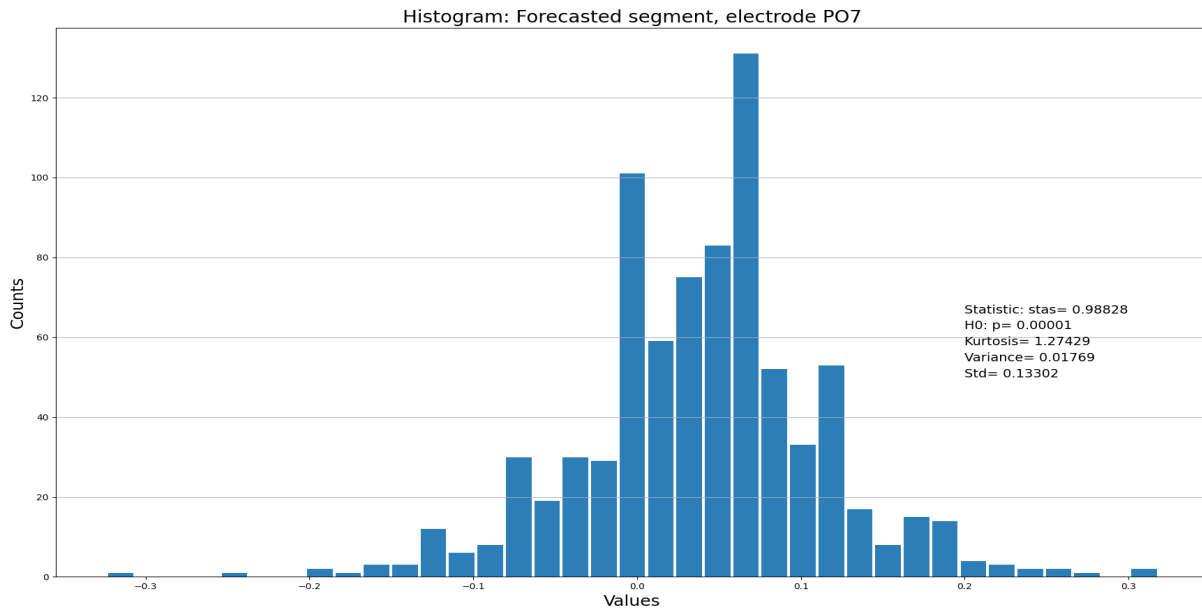


Figura 5.6: Como el segmento de pronóstico no tiene distribución normal, su valor de Curtosis (1.274) es menor comparado con este parámetro en los datos verdaderos. Esto genera más muestras dispersas lejos de la media de los datos.

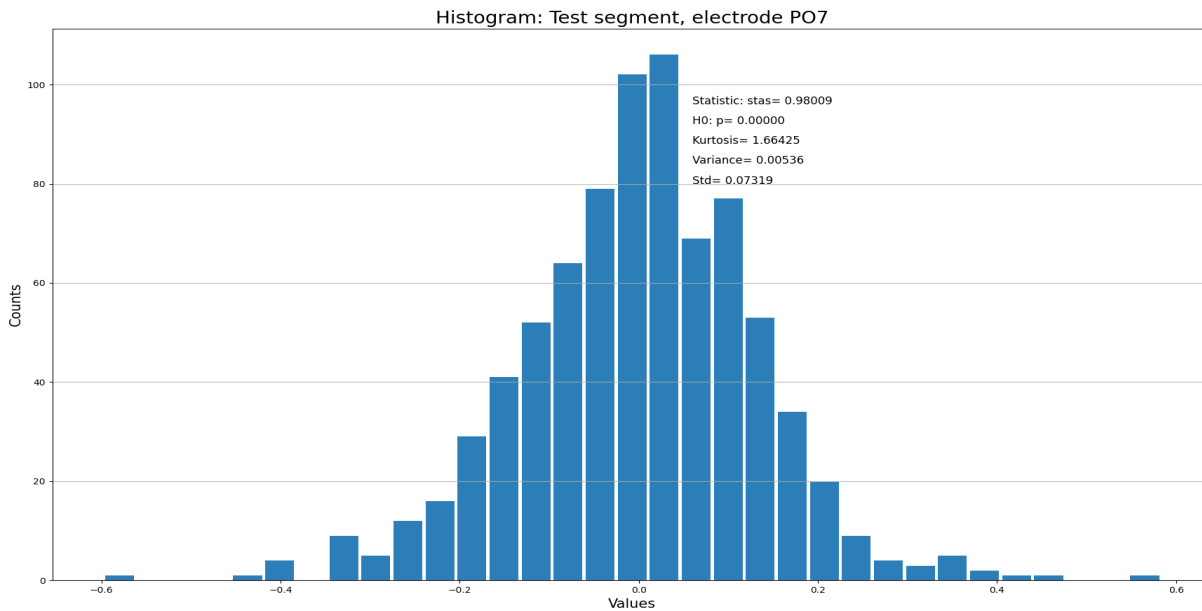


Figura 5.7: El valor de Curtosis del segmento es cercano a 1.664 con un valor std de 0.073. (i) Estos valores muestran que el segmento pronosticado tienen menos muestras atípicas en los extremos.

un valor *test_accuracy* de al menos 90 %, puesto que solo obtuvieron un valor de 0.867 y 0.869, respectivamente. Por último, aunque la métrica *test_accuracy* del modelo *FSR – ISM* es 0.905 y la del modelo *Ensemble Modeling* es 0.902, se cumple con el objetivo fijado de obtener un resultado mínimo de 90 % en la métrica de evaluación para los algoritmos de regresión y clasificación.

Tabla 5.1: Número de electrodos ubicados por región cerebral en la detección y predicción de eventos somnolientos.

Electrode groups	Brain regions	Electrode names
2	PVC	PO7, PO8
4	PMC, Pfc	C3, C4, F3, F4
8	PVC, PAA, PMC, Pfc	PO7, PO8, P3, P4, C3, C4, F3, F4
10	PVC, PAA, PMC Pfc, FC	PO7, PO8, P3, P4, C3 C4, F3, F4, FP1, FP2
16	Occipital, Parietal Frontal, Temporal	PO7, PO8, Oz, T8, P3, Pz, P4, T7 C3, Cz, C4, F3, Fz, F4, FP1, FP2

(i) Primary Visual Cortex (PVC), Visual Association Cortex (VAC), Posterior Association Area (PAA), Primary Motor Cortex (PMC), Motor Association Cortex (MAC), Premotor Cortex (PmC), Prefrontal Cortex (Pfc), FC. (ii) En el grupo con cuatro electrodos se consideró las regiones PMC y Pfc porque son las más activas al recibir un estímulo y en el movimiento de las extremidades.

5.2.1.1. Datos en los niveles de somnolencia

Con el fin de obtener el data-ser de valores DLs necesarios para entrenar el algoritmo de pronóstico en el método dos, se utilizó los datos fisiológicos de EEG no segmentados, los cuales representan una serie temporal en una configuración (*segundos – bloques*). También se utilizó el modelo de regresión *EEG – ISM*, ya que presenta el resultado *test_accuracy* más alto de todos estos modelos (0.964). De este modo, se configuró el data-set con valores DLs sin muestras sintéticas, facilitando su comparación contra los datos obtenidos en el análisis del vídeo. Desafortunadamente, este método no posibilita generar datos sintéticos adicionales para incrementar el tamaño del data-set DLs de entrenamiento del algoritmo *DLs-model*, siendo esta su principal desventaja. Este inconveniente subraya la importancia de tener en cuenta un mayor tiempo en el experimento, aumentando así el número de muestras finales que se pueden obtener.

5.2.2. Ventanas: history and forecast time windows

De acuerdo con los resultados obtenidos, este trabajo demuestra que es posible estimar el tiempo de ocurrencia de subsecuentes eventos somnolientos utilizando una HTW con muestras históricas. El mejor sistema en el método uno (*Forecasting_ECG/(ECG-ResNet-7)*) logra estimar este tiempo con una precisión del 73.4% al final de la ventana FTW. No obstante, con el método dos, esta métrica se incrementa hasta el 79.80% para un t_{FTD-1} entre cero y siete minutos mediante el sistema *EEG-ISM + DLs-model*. Además, los dos métodos implementados en este trabajo reducen la posibilidad de tener una relación lineal entre la longitud de los datos de entrada al modelo de pronóstico y el tiempo estimado por este algoritmo, representando un enfoque opuesto al trabajo de [11]. Este factor es importante al considerar que el tiempo de conducción monótona o experimental aumenta la probabilidad de sentir somnolencia [156, 157], lo cual se puede ver reflejado en los datos obtenidos al final del experimento.

Por otro lado, los dos métodos implementados tienen una resolución temporal de cuatro segundos para t_0 . Sin embargo, el modelo de pronóstico del método uno genera máximo 105 predicciones con frecuencia de muestreo a 200 Hz que deben ser clasificadas por el primer algoritmo en los mismos cuatro segundos. Esto no es eficiente en términos computacionales, ya que los modelos de pronóstico que utilizan más de ocho electrodos EEG requieren de una GPU NVIDIA A100 Tensor Core con

una VRAM de al menos 20GB. Esto se debe a la dimensión de entrada en los datos ($36000 \times 8 \times 1$) muestras, generando una salida con dimensión ($179925 \times 179925 \times 32$) en la primera capa. Al utilizar una GPU de menor capacidad y una HTW más grande, el tiempo de pronóstico es mayor a cuatro segundos, lo cual hace imposible implementar este modelo para funcionar con datos en tiempo real. Además, como la longitud de la ventana HTW es 180 segundos, esto reduce la precisión final del algoritmo de pronóstico.

Por otro lado, en el método dos, las ventanas HTW y FTW tienen la misma longitud de siete minutos. En este caso, el algoritmo de pronóstico utiliza 105 valores históricos, reduciendo los recursos computacionales requeridos para pronosticar 7-DLs dentro de la ventana FTW. Este proceso solo requiere dos segundos para estimar 105 DLs, facilitando poder escalar el modelo para funcionar con datos en tiempo real. Además, al estimar únicamente 7-DLs, los datos generados tienen menos puntos dispersos, por ello mejora la métrica *test_accuracy* en el sistema *regresión + pronóstico*. En cuanto a la resolución temporal en t_0 , esta sigue siendo de cuatro segundos, pero la resolución de salida en la ventana FTW para $t_0 \rightarrow t_{0+7}$ cambia a 28 segundos. Este procedimiento mejora la generalización de los modelos para función con datos en tiempo real, ya que el análisis de los vídeos mostró que los eventos de somnolencia duran al menos 55 segundos consecutivos. Los trabajos de [158,159] encontraron que existe una fuerte interacción entre la tasa de accidentes vehiculares por eventos somnolientos y la experiencia de conducción, pero además, muestran la presencia de eventos somnolientos en los datos de los participantes.

5.2.3. Análisis comparativo de los resultados

Los resultados obtenidos por los algoritmos de regresión y clasificación para la detección del estado somnoliento en t_0 están dentro del estado del arte. Si bien, en el trabajo [97] reportan una precisión del 100 % mediante un algoritmo de clasificación, es de considerar que los datos utilizados en ese trabajo fueron obtenidos en un experimento de conducción simulada. Por lo tanto, la generalización y conveniencia del modelo implementado va a cambiar drásticamente. Además, únicamente utilizaron datos de EEG como medida fisiológica, igual que en el trabajo de [98]. Como se demuestra en los resultados de esta tesis, este tipo de señales son el método más efectivo para detectar somnolencia, pero son altamente susceptibles al ruido, especialmente en condiciones reales. Por otro lado, el trabajo de [33] emplea un enfoque de datos multimodal. El algoritmo implementado está basado en una combinación de CNN con LSTM, obteniendo una precisión del 96 %. En comparación, en esta tesis se reporta el mejor modelo (*EEG_model*) con valor de precisión del 96.4 % en una sola categoría de datos fisiológicos. Además, se reportan los resultados obtenidos con datos FSR, siendo estos los sensores menos intrusivos para el conductor.

Por otro lado, en la predicción de eventos somnolientos subsecuentes, el trabajo de [160] reporta un tiempo de hasta 6.7 minutos. Sin embargo, a pesar de que este trabajo utiliza datos fisiológicos y del vehículo, el estado real de somnolencia (valor absoluto) se basa en el método subjetivo de la escala KSS con una puntuación de somnolencia *ORD* en medidas de diez minutos. Este método induce a errores de clasificación en las señales fisiológicas, ya que su segmentación se hace en base al valor verdadero (KSS) de somnolencia. Representa mediciones subjetivas de los sujetos según su perspectiva de somnolencia. Adicionalmente, los trabajos de [11,105] utilizan un método multimodal que permite detectar y predecir los eventos somnolientos. Sus modelos implementados se entrenan con la primera parte de los datos, se adapta con el data-set del último participante y se prueba con los datos restantes de los demás participantes. Este proceso de adaptación no genera una mejora significativa en los modelos, pero si requiere de mayor tiempo de procesamiento y entrenamiento de

los modelos.

Adicionalmente, utilizar una entrada de datos con longitud variable mayor a 15 minutos implica que en ese lapso no es posible predecir el estado somnoliento. La variación de la longitud en los datos de entrada al modelo está en relación lineal con el factor de exposición a eventos somnolientos (a mayor tiempo conduciendo, hay una mayor probabilidad de somnolencia en el sujeto). Por lo tanto, aunque el error es de 2 minutos para el 95 % de los datos, el modelo tendrá un error de ± 1.33 minutos para el mejor de los casos. En contraste, el modelo de pronóstico del método dos (*DLs_model*) de esta tesis tienen una mayor precisión en eventos somnolientos que están presentes a menos de cinco minutos de la ventana FTW. Además, este factor puede mejorar al tener una mayor cantidad de datos (*LDs*) e información contextual para entrenar este modelo. No obstante, al igual que en los trabajos orientados a detectar o clasificar el estado somnoliento [97–100] y en los trabajos para predecir eventos somnolientos subsecuentes [11, 104, 105, 160], es necesario evaluar la generalización de los modelos con datos de participantes en un mayor rango de edad.

5.2.4. Evaluación del desempeño durante el experimento

La evaluación del desempeño del conductor es un método utilizado para evaluar el impacto de la somnolencia en la ejecución de la tarea, lo cual da una idea del grado de somnolencia en el sujeto. Si el número de infracciones y errores de conducción aumentan, significa que el nivel de somnolencia también se ha incrementado en el chofer. La Figura 5.8 representa el desempeño de un participante en el bloque dos (arriba) y tres (abajo), donde se aprecia un incremento en el número de infracciones al conducir. Aun cuando los escenarios fueron configurados con carriles vehiculares en ambos sentidos, el cambio de carril sin indicar con intermitentes es común durante los dos bloques. Del mismo modo, la Figura 5.9 representa el desempeño global de todos los participantes durante el bloque dos (arriba) y tres (abajo). Se evidencia que, para todos los participantes, la infracción fuera de pista es predominante durante el bloque dos (40 eventos) y bloque tres con 80 sucesos.

Asimismo, para todos los sujetos del bloque dos, se evidencia un total de 173 colisiones frente a 14 colisiones del bloque tres, donde el número de infracciones, eventos y errores durante la conducción es mucho mayor. Esto demuestra una reducción del desempeño global de los participantes durante el experimento. Además, algunos de estos eventos también son consecuencia de sucesos como falta de atención, aburrimiento, estrés y fatiga. En este punto, el tiempo pasado en el simulador y el entorno del experimento son factores que impactan en el estado anímico de los participantes. Por ejemplo, en los dos escenarios en los cuales los voluntarios debían conducir a una velocidad máxima de 30km/h, el error de conducción (fuera de pista) es más representativo en el bloque tres que en el bloque dos.

Por otro lado, la Figura 5.10 representa la suma de las muestras para las tres clases por cada minuto con todos los sujetos del experimento en el bloque dos. A medida que aumenta el tiempo de conducción en este bloque, el número de muestras en la clase “alertness” disminuye y se incrementa el total de muestras en las clases “drowsiness-1” y “drowsiness-2”. De este modo, en el minuto 29 del segundo bloque experimental hay 19 sujetos que presentan somnolencia de primer grado y 11 voluntarios con somnolencia de nivel dos. Si bien, el número de muestras en la clase “drowsiness-2” entre todos los sujetos aumenta levemente desde el minuto nueve, al final de este bloque hay 19 sujetos con somnolencia de grado dos.

Adicionalmente, la Figura 5.11 presenta el resultado de la variación del número de muestras en las tres clases para todos los sujetos en cada minuto del tercer bloque. En este caso, hay un 50 %

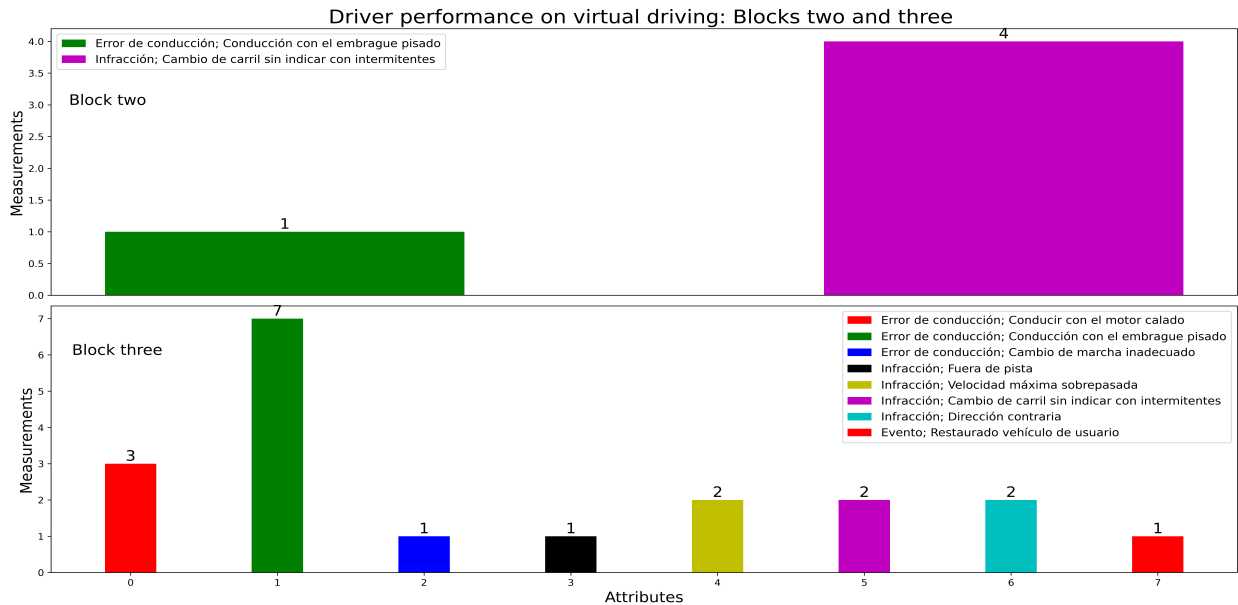


Figura 5.8: Este sujeto muestra una mayor cantidad de infracciones al pasar del bloque dos (arriba) a la conducción durante el bloque tres (abajo). (i) Cambiar de carril sin indicar con intermitentes es una infracción común en estos dos bloques del experimento.

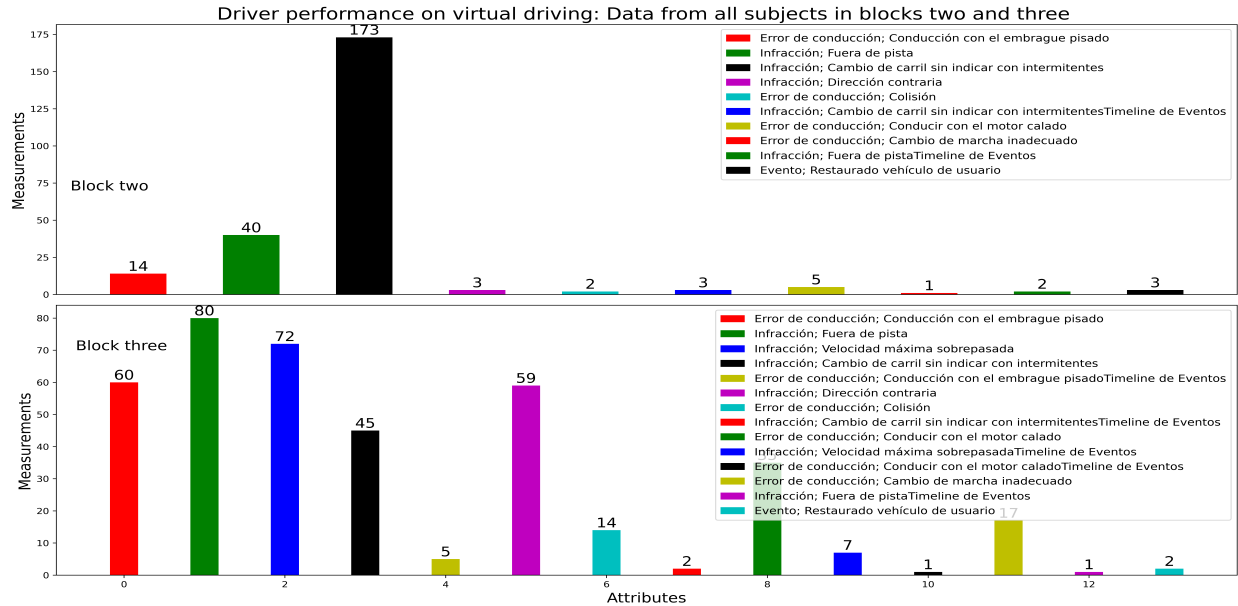


Figura 5.9: Se evidencia un incremento en el número de infracciones y errores durante la conducción en el bloque dos (arriba) y tres (abajo). (i) La infracción fuera de pista es la que presenta una mayor cantidad de eventos durante el bloque tres, demostrando que la somnolencia impactó negativamente en todos los participantes durante el experimento.

de sujetos que presentaron somnolencia en el minuto siete y alrededor del 16 % de los participantes demostró señales de somnolencia en el minuto tres. De este modelo, en el minuto diez del bloque tres, por lo menos el 50 % de los participantes presentaron somnolencia, llegando a más del 90 % de los sujetos con señales claras de somnolencia en el minuto 27 y al 100 % en el último minuto. Finalmente, al analizar los resultados de la Figura 5.11 y los datos de la Figura 5.9, se evidencia que la cantidad de infracciones y errores de conducción están relacionados con el impacto del nivel

de somnolencia en los participantes a lo largo de estos dos bloques experimentales. Conforme se incrementa el número de sujetos con muestras de las clases “drowsiness-1” y “drowsiness-2”, también aumenta el número de colisiones e infracciones al conducir, siendo mucho más evidente en los últimos 16 minutos del bloque tres.

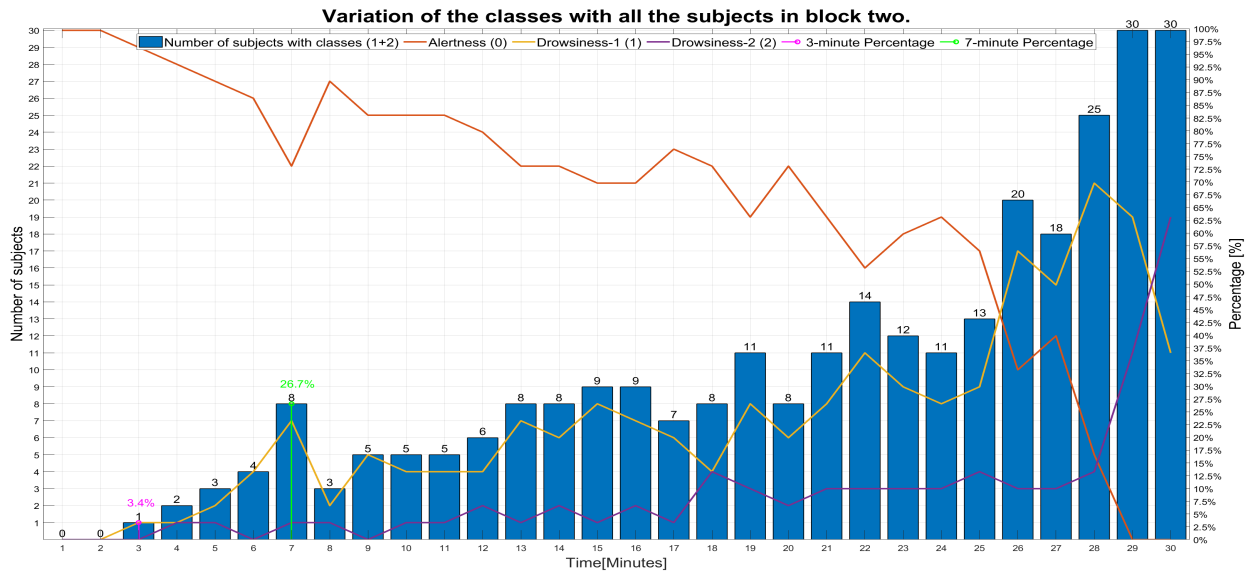


Figura 5.10: (i) La línea de color verde representa el 26.7% de participantes que experimentaron somnolencia. (ii) La línea de color magenta corresponde al 3.4% de participantes que presentaron somnolencia. (iii) En los minutos 29 y 30, el 100% de los sujetos presentan somnolencia en este bloque.

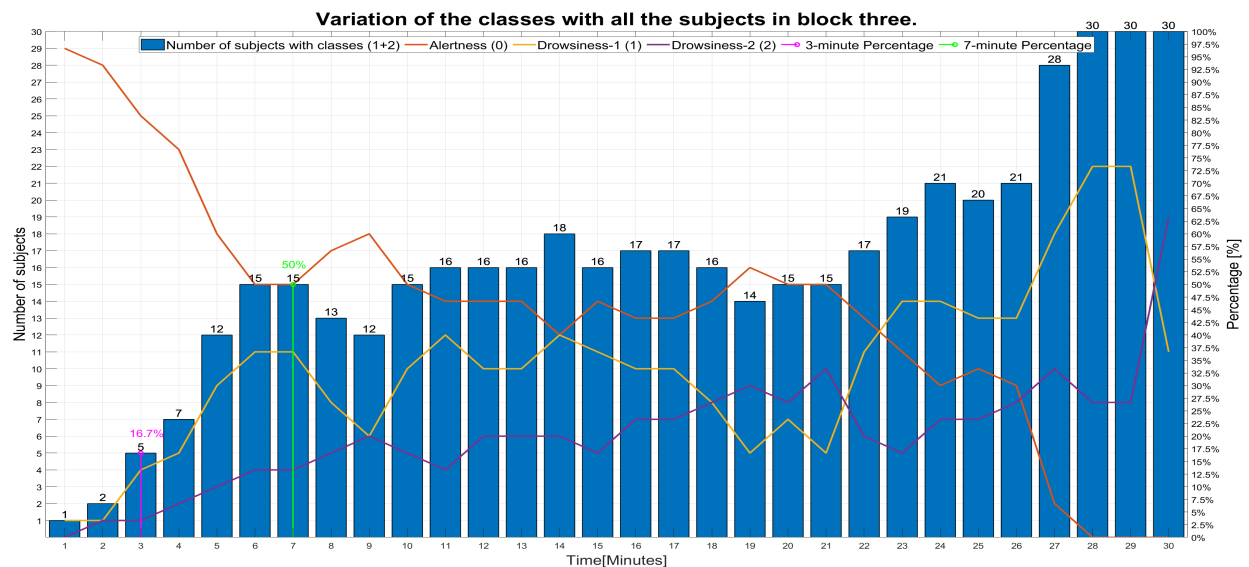


Figura 5.11: (i) El 36.6% de los participantes experimentó somnolencia leve y el 13.3% exhibió señales de somnolencia fuerte. (ii) La línea de color magenta en el minuto tres representa el 16.7% de participantes que experimentaron somnolencia.

CONCLUSIONES

En la primera etapa de esta investigación, se implementó un algoritmo para predecir la intención de frenada de emergencia y estimar el tiempo de reacción, logrando un AUC de 0,820 con solo cuatro electrodos y entrenando el algoritmo con segmentos en disposición tensorial. La estimación del tiempo de reacción evidenció que es posible reducir la distancia de frenado hasta en 15.35 metros para un vehículo a 80km/h. Adicionalmente, la parte central de esta investigación se orientó en detectar el estado somnoliento del conductor en t_0 para pronosticar el subsiguiente evento de somnolencia dentro de una ventana temporal ($t_0 \rightarrow t_{0+7}$). Se utilizaron sensores de EEG, ECG, EMG, FSR y Telemetría del vehículo para conformar una base de datos multimodal. El método uno consta de un algoritmo de clasificación y un modelo de pronóstico para generar una señal de salida del estado subsecuente. El método dos utiliza un algoritmo de regresión y un modelo de pronóstico basado en los niveles de somnolencia del algoritmo de regresión.

El mejor modelo (*Forecasting ECG-model*) del método uno obtuvo un resultado *validation_RMSE* de 0.217. No obstante, la combinación de los modelos de clasificación y pronóstico necesitan una sincronía temporal que se ve afectada cuando los recursos computacionales no son suficientes, especialmente en el pronóstico mediante señales de EEG con más de cuatro electrodos. En el método dos, el algoritmo de regresión genera un valor continuo en el rango (0 a 2) correspondiente al nivel de somnolencia en el tiempo t_0 . El modelo de regresión (*EEG-ISM*) logró el mejor valor *test_accuracy* (0.964) cuando se utilizaron los datos de EEG con los 16 electrodos. No obstante, es de notar que el modelo *FSR-ISM* obtuvo un valor *test_accuracy* de 0,905 utilizando datos de sensores no intrusivos que están implementados directamente en los pedales y el volante del vehículo.

6.1. Conclusiones

En relación con la detección y predicción de eventos subsecuentes de somnolencia en el conductor, la primera conclusión que valida la hipótesis planteada está relacionada con la combinación de los modelos de *regresión* y el algoritmo de *pronóstico*, permitiendo predecir el estado subsecuente del conductor dentro del intervalo $t_0 \rightarrow t_{0+7}$ minutos; se utiliza una entrada multimodal y un arreglo secuencial con valores de los niveles de somnolencia obtenidos del algoritmo de regresión. Asimismo, con la mejor combinación de estos modelos (*EEG-ISM + DLs-model*) se obtiene un resultado *test_accuracy* de 88.7% y 79.8% para todo t_{FTD-1} en el rango de cinco y siete minutos, respectivamente. Esto representa una mejora del 10.2% en comparación con los resultados obtenidos

con el método uno para todo t_{FTD-1} en el rango de la venta FTW. Además, la técnica implementada en el modelo *DLS-model* requiere máximo una GPU con 1GB y no más de 4GB en memoria RAM, lo que permite escalar el modelo para funcionar en un dispositivo con datos en tiempo real.

Otra importante conclusión que reafirma la hipótesis de esta investigación se basa en que la combinación de los modelos de *clasificación + pronóstico* del método uno predicen cualquier evento somnoliento subsecuente dentro del intervalo ($t_0 \rightarrow t_{0+7}$) utilizando como entrada la secuencia de 36000 muestras históricas. El modelo de clasificación con la menor cantidad de sensores (*ECG-ResNet-7*) obtuvo un resultado *validation_accuracy* de 91.8% para t_0 y utilizó solo la II-derivación de ECG. Para la mejor combinación *Forecasting_ECG/ECG-ResNet-7* se obtuvo un valor *text_accuracy* de 78.5% en el rango de 5 minutos. Sin embargo, al utilizar más de cuatro electrodos de EEG se requiere una GPU con al menos 20GB (NVIDIA A100 Tensor Core) y una memoria RAM de 16GB. Estos resultados y requerimientos imposibilitan poder escalar el modelo para funcionar en un dispositivo con datos en tiempo real.

En relación con el conjunto de sensores más adecuados para detectar somnolencia en conductores de vehículos, se puede concluir que las señales de EEG con cuatro electrodos y el algoritmo de clasificación permiten detectar el estado somnoliento en t_0 con una precisión del 92.6%, lo cual está dentro del estado del arte para esta clase de trabajos. No obstante, los sensores fisiológicos implementados sobre el cuerpo de los participantes resultaron ser invasivos e incómodos. Por lo tanto, la implementación de sensores FSR resulta ser el método más confortable para el conductor, pero con una precisión menor en 6% comparado con el algoritmo de clasificación y en 5.9% respecto del modelo de regresión. Además, los sensores FSR implementados sobre los asientos, pedales y volante del vehículo resultan de gran utilidad para monitorear el confort de los pasajeros y del conductor. En el lado opuesto, se puede concluir que los datos de EMG son los sensores menos efectivos y precisos para detectar y predecir eventos somnolientos subsecuentes en el conductor, ya que no detectan señales directas de somnolencia.

Otro aspecto importante de esta investigación es el experimento realizado para obtener la base de datos multimodal. En este punto, se concluye que las condiciones preliminares establecidas en la selección de los participantes son indispensables a fin de medir la respuesta directa de la somnolencia por desactivación. De otro modo, productos psicoestimulantes como la cafeína alteran drásticamente la respuesta del sistema nervioso del participante en el experimento. Además, la etapa de adaptación al simulador de conducción en el participante permitió reducir la carga cognitiva extrínseca (sesgo de adaptación) que se presenta durante los primeros minutos del experimento. También es importante destacar el número de participantes y el tiempo experimental. Si bien, no se ha establecido un número mínimo o máximo de voluntarios necesarios para esta clase de trabajos, los requerimientos de los modelos muestran que una mayor cantidad de datos genera una contundente mejora en el entrenamiento de los algoritmos implementados.

Por otro lado, en base al algoritmo implementado en la primera etapa de esta investigación, se concluye que es posible predecir la intención de frenada de emergencia con una precisión del 82.2% utilizando cuatro electrodos de EEG. Esto puede llevar a reducir el tiempo de reacción de frenado BRT en 689 milisegundos, activando los frenos del vehículo antes que el conductor. Esta importante reducción ayudaría a evitar un accidente automovilístico en el momento que el conductor esté somnoliento o con fatiga. Además, el método de entrenamiento del modelo CNN implementado demostró que no siempre se requiere transformar las señales a imágenes en escala de grises para entrenar esta clase de algoritmos. No obstante, para mejorar la precisión y generalización de este algoritmo de predicción se requiere de una mayor cantidad de datos obtenidos en diferentes horas y condiciones experimentales.

6.2. Trabajo futuro

A pesar de que el modelo de pronóstico del método uno no es el más eficiente dado los recursos computacionales que necesita, este algoritmo puede ser fácilmente adaptado en la generación de datos sintéticos para incrementar el tamaño de data-sets o completar series temporales que tienen datos ausentes. También pueden adaptarse para predecir el estado subsecuente del ritmo cardíaco o el tiempo de ocurrencia de algunos tipos de fibrilación mediante muestras históricas de datos ECG. En este mismo orden de ideas, es posible aplicar esta técnica para predecir el nivel de fatiga y efecto cognitivo en aplicaciones Brain-Computer Interface (BCI) que suele utilizar no más de cuatro sensores. Además, los algoritmos de pronóstico del método uno y dos podrían lograr mejoras futuras utilizando técnicas adicionales de aprendizaje como Inductive Transfer Learning, ya que permite entrenar los modelos con datos relacionados en una tarea diferente y luego transferir el “conocimiento” inicial para entrenar las capas finales de los mismos modelos con los datos y tarea deseada.

En relación con la base de datos multimodal configurada con las señales obtenidas en el experimento, estos datos representan un aporte adicional de esta investigación. La base de datos puede utilizarse para evaluar parámetros como el estilo de manejo mediante los sensores FSR en contraste con el análisis de vídeo. También puede emplearse las señales de EMG para determinar la evolución de la fatiga física y el deterioro del desempeño en el conductor durante el tiempo del experimento. Además, como en esta tesis solo se reporta el análisis de los dos últimos bloques del experimento, queda el primer bloque para evaluar el tiempo de reacción, la carga cognitiva o fatiga de los participantes. El primer bloque del experimento se configuró como un escenario con variación ascendente de la dificultad de la tarea que consta de varios obstáculos, la dificultad del circuito y los diferentes requerimientos de velocidad por cada número de vueltas.

Otro aspecto crítico de esta investigación está relacionado con las señales obtenidas del experimento. Para mejorar el algoritmo de pronóstico se recomienda obtener una mayor cantidad de datos en un experimento con más de 95 minutos. Sería prudente considerar una etapa experimental en diferentes horas del día para medir la evolución del nivel de somnolencia en el participante. Por lo tanto, la etapa de adaptación al simulador y las condiciones preliminares que limitan el consumo de sustancias psicoestimulantes deben ser estrictamente controladas a fin de evitar sesgos en la respuesta del participante a las variables experimentales. En esta misma línea, se hace necesario incorporar un mayor tamaño de la muestra en un rango de edad más amplio, ya que en este experimento solo participaron conductores menores de 35 años. También resultaría oportuno que la cantidad de voluntarios esté configurada por igual número de hombres y mujeres. Además, es crucial incorporar información como la variabilidad inter-sujetos e intra-sujetos para verificar si hay una mejora de la conveniencia y generalización de los modelos.

Finalmente, dado que las señales de EEG, ECG y EMG pueden ser intrusivas, la atención podría centrarse en mejorar un modelo basado en sensores FSR y datos de Telemetría del vehículo. Así, un dispositivo para la detección de somnolencia con esta clase de sensores no depende de la voluntad del conductor para que se utilice constantemente durante la ejecución de su tarea. Tampoco se vería afectado por condiciones ambientales del entorno o problemas de iluminación dentro del vehículo. Además, dado que las señales FSR no interfieren con la actividad del conductor y no presentan riesgo alguno para su seguridad o la de los demás ocupantes del vehículo, puede implementarse un dispositivo para obtener esta clase de datos en condiciones reales, sin que esto induzca somnolencia en el conductor.

ANEXO A: PUBLICACIONES Y CONGRESOS

Los siguientes son algunos de los resultados obtenidos en este trabajo de investigación, los cuales se publicaron en revistas y presentaciones en congresos.

- Mora, H. J., and Pino, E. J. “Simplified Prediction Method for Detecting the Emergency Braking Intention Using EEG and a CNN Trained with a 2D Matrices Tensor Arrangement.” *International Journal of Human–Computer Interaction* 39.3 (2023): 587-600.
- Mora, H. J., Echaveguren, T. B., and Pino, E. J. “Employing a multimodal database to detect the driver somnolence and forecast the horizon time of drowsiness events using a ResNet architecture.” *Biocybernetics and Biomedical Engineering* (2024) (En revisión).
- Mora, H. J., Echaveguren, T. B., and Pino, E. J., (2022, November). “A System Detection of Atrial Fibrillation Using One ECG Derivation and Inductive Transfer Learning”. In *International Conference on Biomedical and Health Informatics* (pp. 69-80). Cham: Springer Nature Switzerland.
- Pino, E. J., Mora, H. J., Sepúlveda, M. A., Chavez, J. A., Lecannelier, E. A., De Carvalho, P., ... and Magjarevic, R. (2022, November). “Ballistocardiogram database from unobtrusive sensors in sitting volunteers for the evaluation of beat detection algorithms”. In *International Conference on Biomedical and Health Informatics* (pp. 281-288). Cham: Springer Nature Switzerland.
- Mora, H. J., Echaveguren, T. B., and Pino, E. J. “A system detection of Atrial Fibrillation using one ECG derivation and Inductive Transfer Learning.” *ICBHI 2022 - International Conference on Biomedical and Health Informatics*.
- Mora, H. J., Echaveguren, T. B., and Pino, E. J. “Non-intrusive sensors to predict subsequent drowsiness events in drivers”. *Workshop CENIA 2024*.

BIBLIOGRAFÍA

- [1] M. Gonçalves, R. Amici, R. Lucas, T. Åkerstedt, F. Cirignotta, J. Horne, D. Léger, W. T. McNicholas, M. Partinen, J. Téran-Santos, et al., Sleepiness at the wheel across europe: a survey of 19 countries, *Journal of sleep research* 24 (3) (2015) 242–253.
- [2] M. E. Shaik, M. M. Islam, Q. S. Hossain, A review on neural network techniques for the prediction of road traffic accident severity, *Asian Transport Studies* 7 (2021) 100040.
- [3] T. Stewart, Overview of motor vehicle traffic crashes in 2021 (report no. dot hs 813 435), technical report, U.S. Department of Transportation (2023).
- [4] Comisión Nacional de Seguridad de Tránsito (CONASET), Libro del nuevo conductor: Automovilistas, CONASET 2 (2018) 181.
- [5] European Commission (2021) Road safety thematic report – Fatigue, European road safety observatory, European Commission, Directorate General for Transport (2021).
- [6] K. Vitols, E. Voss, Driver fatigue in european road transport, European Transport Workers' Federation (2021).
- [7] Comisión Nacional de Seguridad de Tránsito (CONASET), Costo social de los siniestros de tránsito en Chile 2020, Observatorio de datos. Comisión Nacional de Seguridad de Tránsito (2021).
URL <https://www.conaset.cl/wp-content/uploads/2021/05/Costos-accidentes-2020.pdf>
- [8] Comisión Nacional de Seguridad de Tránsito (CONASET), Informe nacional de siniestros de tránsito en Chile 2023, CONASET 2 (2018) 181.
- [9] A. M. Williamson, A.-M. Feyer, Moderate sleep deprivation produces impairments in cognitive and motor performance equivalent to legally prescribed levels of alcohol intoxication, *Occupational and environmental medicine* 57 (10) (2000) 649–655.
- [10] M. E. Shaik, A systematic review on detection and prediction of driver drowsiness, *Transportation research interdisciplinary perspectives* 21 (2023) 100864.
- [11] C. J. de Naurois, C. Bourdin, A. Stratulat, E. Diaz, J.-L. Vercher, Detection and prediction of driver drowsiness using artificial neural network models, *Accident Analysis & Prevention* 126 (2019) 95–104.

-
- [12] A. Altameem, A. Kumar, R. C. Poonia, S. Kumar, A. K. J. Saudagar, Early identification and detection of driver drowsiness by hybrid machine learning, *IEEE Access* 9 (2021) 162805–162819.
- [13] A. W. Cai, J. E. Manousakis, B. Singh, E. Francis-Pester, J. Kuo, K. J. Jeppe, S. M. Rajaratnam, M. G. Lenné, M. E. Howard, C. Anderson, Subjective awareness of sleepiness while driving in younger and older adults, *Journal of sleep research* (2023) e13933.
- [14] J. S. Higgins, J. Michael, R. Austin, T. Åkerstedt, H. P. Van Dongen, N. Watson, C. Czeisler, A. I. Pack, M. R. Rosekind, Asleep at the wheel—the road to addressing drowsy driving, *Sleep* 40 (2) (2017) zsx001.
- [15] S. Scarpelli, V. Alfonsi, M. Gorgoni, M. Camaioni, A. M. Giannini, L. De Gennaro, Age-related effect of sleepiness on driving performance: a systematic-review, *Brain sciences* 11 (8) (2021) 1090.
- [16] A. Chaudhuri, A. Routray, Driver fatigue detection through chaotic entropy analysis of cortical sources obtained from scalp eeg signals, *IEEE transactions on intelligent transportation systems* 21 (1) (2019) 185–198.
- [17] P. Bashivan, M. Yeasin, G. M. Bidelman, Single trial prediction of normal and excessive cognitive load through eeg feature fusion, in: *2015 IEEE signal processing in medicine and biology symposium (SPMB)*, IEEE, 2015, pp. 1–5.
- [18] F. Dehais, A. Duprès, S. Blum, N. Drougard, S. Scannella, R. N. Roy, F. Lotte, Monitoring pilot’s mental workload using erps and spectral power with a six-dry-electrode eeg system in real flight conditions, *Sensors* 19 (6) (2019) 1324.
- [19] P. Zhang, J. Beernaerts, N. Van de Weghe, A hybrid approach combining the multi-temporal scale spatio-temporal network with the continuous triangular model for exploring dynamic interactions in movement data: A case study of football, *ISPRS International Journal of Geo-Information* 7 (1) (2018) 31.
- [20] Z. Gao, X. Wang, Y. Yang, C. Mu, Q. Cai, W. Dang, S. Zuo, Eeg-based spatio-temporal convolutional neural network for driver fatigue evaluation, *IEEE transactions on neural networks and learning systems* 30 (9) (2019) 2755–2763.
- [21] S.-Y. Han, N.-S. Kwak, T. Oh, S.-W. Lee, Classification of pilots’ mental states using a multimodal deep learning network, *Biocybernetics and Biomedical Engineering* 40 (1) (2020) 324–336.
- [22] N. S. Karuppusamy, B.-Y. Kang, Multimodal system to detect driver fatigue using eeg, gyroscope, and image processing, *IEEE Access* 8 (2020) 129645–129667.
- [23] S. Ahn, T. Nguyen, H. Jang, J. G. Kim, S. C. Jun, Exploring neuro-physiological correlates of drivers’ mental fatigue caused by sleep deprivation using simultaneous eeg, ecg, and fnirs data, *Frontiers in human neuroscience* 10 (2016) 219.
- [24] Y. Liu, H. Ayaz, P. A. Shewokis, Multisubject “learning” for mental workload classification using concurrent eeg, fnirs, and physiological measures, *Frontiers in human neuroscience* 11 (2017) 389.

-
- [25] Y. Liu, Y. Wang, Y. Liu, S. Xu, M. Chen, S. Du, J. Gong, Solubility of l-histidine in different aqueous binary solvent mixtures from 283.15 k to 318.15 k with experimental measurement and thermodynamic modelling, *The Journal of Chemical Thermodynamics* 105 (2017) 1–14.
- [26] C.-T. Lin, R.-C. Wu, S.-F. Liang, W.-H. Chao, Y.-J. Chen, T.-P. Jung, Eeg-based drowsiness estimation for safety driving using independent component analysis, *IEEE Transactions on Circuits and Systems I: Regular Papers* 52 (12) (2005) 2726–2738.
- [27] S. A. El-Nabi, W. El-Shafai, E.-S. M. El-Rabaie, K. F. Ramadan, F. E. Abd El-Samie, S. Mohsen, Machine learning and deep learning techniques for driver fatigue and drowsiness detection: a review, *Multimedia Tools and Applications* 83 (3) (2024) 9441–9477.
- [28] I.-H. Kim, J.-W. Kim, S. Haufe, S.-W. Lee, Detection of braking intention in diverse situations during simulated driving based on eeg feature combination, *Journal of neural engineering* 12 (1) (2014) 016001.
- [29] H. Zhang, M. Zhao, C. Wei, D. Mantini, Z. Li, Q. Liu, Eegdenoisenet: a benchmark dataset for deep learning solutions of eeg denoising, *Journal of Neural Engineering* 18 (5) (2021) 056057.
- [30] J. Halomoan, K. Ramli, D. Sudiana, T. S. Gunawan, M. Salman, A new eeg data processing approach to developing an accurate driving fatigue detection framework with heart rate variability analysis and ensemble learning, *Information* 14 (4) (2023) 210.
- [31] P. Zhang, X. Wang, J. Chen, W. You, Feature weight driven interactive mutual information modeling for heterogeneous bio-signal fusion to estimate mental workload, *Sensors* 17 (10) (2017) 2315.
- [32] A. Amidei, S. Spinsante, G. Iadarola, S. Benatti, F. Tamarin, P. Pavan, L. Rovati, Driver drowsiness detection: a machine learning approach on skin conductance, *Sensors* 23 (8) (2023) 4004.
- [33] S. Priyanka, S. Shanthi, A. S. Kumar, V. Praveen, Data fusion for driver drowsiness recognition: A multimodal perspective, *Egyptian Informatics Journal* 27 (2024) 100529.
- [34] C. C. Liu, S. G. Hosking, M. G. Lenné, Predicting driver drowsiness using vehicle measures: Recent insights and future challenges, *Journal of safety research* 40 (4) (2009) 239–245.
- [35] T. Kunding, A. Riener, N. Sofra, A robust drowsiness detection method based on vehicle and driver vital data, *Mensch und Computer 2017-Workshopband: Spielend einfach interagieren* 13 (2018) 433.
- [36] Y. LeCun, Y. Bengio, et al., Convolutional networks for images, speech, and time series, *The handbook of brain theory and neural networks* 3361 (10) (1995) 1995.
- [37] M. Dua, Shakshi, R. Singla, S. Raj, A. Jangra, Deep cnn models-based ensemble approach to driver drowsiness detection, *Neural Computing and Applications* 33 (2021) 3155–3168.
- [38] A. Murata, Y. Fujii, K. Naitoh, Multinomial logistic regression model for predicting driver's drowsiness using behavioral measures, *Procedia manufacturing* 3 (2015) 2426–2433.
- [39] H. Su, G. Zheng, A partial least squares regression-based fusion model for predicting the trend in drowsiness, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 38 (5) (2008) 1085–1092.

-
- [40] A. Murata, Y. Ohta, M. Moriwaka, Multinomial logistic regression model by stepwise method for predicting subjective drowsiness using performance and behavioral measures, in: *Advances in Physical Ergonomics and Human Factors: Proceedings of the AHFE 2016 International Conference on Physical Ergonomics and Human Factors*, July 27-31, 2016, Walt Disney World®, Florida, USA, Springer, 2016, pp. 665–674.
- [41] G. S. Larue, Predicting effects of monotony on driver’s vigilance, Ph.D. thesis, Queensland University of Technology (2010).
- [42] M. Schuster, K. K. Paliwal, Bidirectional recurrent neural networks, *IEEE transactions on Signal Processing* 45 (11) (1997) 2673–2681.
- [43] I. E. Livieris, E. Pintelas, P. Pintelas, A cnn–lstm model for gold price time-series forecasting, *Neural computing and applications* 32 (2020) 17351–17360.
- [44] V. Brodbeck, A. Kuhn, F. von Wegner, A. Morzelewski, E. Tagliazucchi, S. Borisov, C. M. Michel, H. Laufs, Eeg microstates of wakefulness and nrem sleep, *Neuroimage* 62 (3) (2012) 2129–2139.
- [45] A. Shahid, K. Wilkinson, S. Marcu, C. M. Shapiro, Karolinska sleepiness scale (kss), STOP, THAT and one hundred other sleep scales (2012) 209–210.
- [46] S. Otmani, T. Pebayle, J. Roge, A. Muzet, Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers, *Physiology & behavior* 84 (5) (2005) 715–724.
- [47] P. M. Forsman, B. J. Vila, R. A. Short, C. G. Mott, H. P. Van Dongen, Efficient driver drowsiness detection at moderate levels of drowsiness, *Accident Analysis & Prevention* 50 (2013) 341–350.
- [48] B.-C. Yin, X. Fan, Y.-F. Sun, Multiscale dynamic features based driver fatigue detection, *International Journal of Pattern Recognition and Artificial Intelligence* 23 (03) (2009) 575–589.
- [49] M. Akin, M. B. Kurt, N. Sezgin, M. Bayram, Estimating vigilance level by using eeg and emg signals, *Neural Computing and Applications* 17 (2008) 227–236.
- [50] A. Kokonozi, E. Michail, I. Chouvarda, N. Maglaveras, A study of heart rate and brain system complexity and their interaction in sleep-deprived subjects, in: *2008 Computers in Cardiology*, IEEE, 2008, pp. 969–971.
- [51] W. C. Liang, J. Yuan, D. C. Sun, M. H. Lin, Changes in physiological parameters induced by indoor simulated driving: Effect of lower body exercise at mid-term break, *Sensors* 9 (9) (2009) 6913–6933.
- [52] A. Ziebinski, R. Cupek, D. Grzechca, L. Chruszczyk, Review of advanced driver assistance systems (ADAS), in: *AIP Conference Proceedings*, Vol. 1906, AIP Publishing, 2017, p. 120002.
- [53] M. Lu, K. Wevers, R. Van Der Heijden, Technical feasibility of advanced driver assistance systems (adas) for road traffic safety, *Transportation Planning and Technology* 28 (3) (2005) 167–187.
- [54] V. A. Butakov, P. Ioannou, Personalized driver/vehicle lane change models for adas, *IEEE Transactions on Vehicular Technology* 64 (10) (2014) 4422–4431.

-
- [55] F. Sharbrough, American Electroencephalographic Society guidelines for standard electrode position nomenclature, *J clin Neurophysiol* 8 (1991) 200–202.
URL <https://ci.nii.ac.jp/naid/1000918324655>.
- [56] A. T. Satti, J. Kim, E. Yi, H.-y. Cho, S. Cho, Microneedle array electrode-based wearable emg system for detection of driver drowsiness through steering wheel grip, *Sensors* 21 (15) (2021) 5091.
- [57] I. Stancin, M. Cifrek, A. Jovic, A review of eeg signal features and their application in driver drowsiness detection systems, *Sensors* 21 (11) (2021) 3786.
- [58] S.-J. Jung, H.-S. Shin, W.-Y. Chung, Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel, *IET Intelligent Transport Systems* 8 (1) (2014) 43–50.
- [59] K. Van Benthem, S. Cebulski, C. M. Herdman, J. Keillor, An EEG brain–computer interface approach for classifying vigilance states in humans: A gamma band focus supports low misclassification rates, *International Journal of Human–Computer Interaction* 34 (3) (2018) 226–237. doi:<https://doi.org/10.1080/10447318.2017.1342942>.
- [60] M. Vidal, M. Rosso, A. M. Aguilera, P-spline smoothed functional ICA of EEG data, arXiv preprint arXiv:2101.05769 (2021).
URL <https://arxiv.org/abs/2101.05769>
- [61] A. A. Borbély, et al., A two process model of sleep regulation, *Hum neurobiol* 1 (3) (1982) 195–204.
- [62] S. Cateni, V. Colla, M. Vannucci, A. Borselli, et al., Fuzzy inference systems applied to image classification in the industrial field, *Fuzzy Inference System: Theory and Applications* (2012).
- [63] M. S. Devi, P. R. Bajaj, Fuzzy based driver fatigue detection, in: 2010 IEEE International Conference on Systems, Man and Cybernetics, IEEE, 2010, pp. 3139–3144.
- [64] F. O. Karray, C. W. De Silva, *Soft computing and intelligent systems design: theory, tools, and applications*, Pearson Education, 2004.
- [65] R. Sayed, A. Eskandarian, Unobtrusive drowsiness detection by neural network learning of driver steering, *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* 215 (9) (2001) 969–975.
- [66] M. Patel, S. K. Lal, D. Kavanagh, P. Rossiter, Applying neural network analysis on heart rate variability data to assess driver fatigue, *Expert systems with Applications* 38 (6) (2011) 7235–7242.
- [67] M. Ngxande, J.-R. Tapamo, M. Burke, Driver drowsiness detection using behavioral measures and machine learning techniques: A review of state-of-art techniques, 2017 pattern recognition Association of South Africa and Robotics and mechatronics (PRASA-RobMech) (2017) 156–161.
- [68] M. M. Hasan, C. N. Watling, G. S. Larue, Physiological signal-based drowsiness detection using machine learning: Singular and hybrid signal approaches, *Journal of safety research* 80 (2022) 215–225.

-
- [69] S. Soares, T. Monteiro, A. Lobo, A. Couto, L. Cunha, S. Ferreira, Analyzing driver drowsiness: From causes to effects, *Sustainability* 12 (5) (2020) 1971.
- [70] V. Pilinkienė, Market demand forecasting models and their elements in the context of competitive market, *Engineering economics* . (5 (60)) (2008) 24–31.
- [71] S. Regnier, D. B. Ridley, Forecasting market share in the us pharmaceutical market, *nature reviews drug discovery* 14 (9) (2015) 594–595.
- [72] M. M. Kumbure, C. Lohrmann, P. Luukka, J. Porras, Machine learning techniques and data for stock market forecasting: A literature review, *Expert Systems with Applications* 197 (2022) 116659.
- [73] T. Kim, H. Y. Kim, Forecasting stock prices with a feature fusion lstm-cnn model using different representations of the same data, *PloS one* 14 (2) (2019) e0212320.
- [74] N. Mbatha, H. Bencherif, Time series analysis and forecasting using a novel hybrid lstm data-driven model based on empirical wavelet transform applied to total column of ozone at buenos aires, argentina (1966–2017), *Atmosphere* 11 (5) (2020) 457.
- [75] F. Kaytez, M. C. Taplamacioglu, E. Cam, F. Hardalac, Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines, *International Journal of Electrical Power & Energy Systems* 67 (2015) 431–438.
- [76] S. Siami-Namini, N. Tavakoli, A. S. Namin, A comparison of arima and lstm in forecasting time series, in: *2018 17th IEEE international conference on machine learning and applications (ICMLA)*, IEEE, 2018, pp. 1394–1401.
- [77] F. Schorfheide, D. Song, Real-time forecasting with a mixed-frequency var, *Journal of Business & Economic Statistics* 33 (3) (2015) 366–380.
- [78] G. Zhang, B. E. Patuwo, M. Y. Hu, Forecasting with artificial neural networks:: The state of the art, *International journal of forecasting* 14 (1) (1998) 35–62.
- [79] S. Palani, S.-Y. Liong, P. Tkalic, An ann application for water quality forecasting, *Marine pollution bulletin* 56 (9) (2008) 1586–1597.
- [80] S. Mehtab, J. Sen, Analysis and forecasting of financial time series using cnn and lstm-based deep learning models, in: *Advances in Distributed Computing and Machine Learning: Proceedings of ICADCML 2021*, Springer, 2022, pp. 405–423.
- [81] J. Cao, Z. Li, J. Li, Financial time series forecasting model based on ceemdan and lstm, *Physica A: Statistical mechanics and its applications* 519 (2019) 127–139.
- [82] H. Abbasimehr, M. Shabani, M. Yousefi, An optimized model using lstm network for demand forecasting, *Computers & industrial engineering* 143 (2020) 106435.
- [83] S. Siami-Namini, N. Tavakoli, A. S. Namin, The performance of lstm and bilstm in forecasting time series, in: *2019 IEEE International conference on big data (Big Data)*, IEEE, 2019, pp. 3285–3292.
- [84] R. McKendrick, B. Feest, A. Harwood, B. Falcone, Theories and methods for labeling cognitive workload: Classification and transfer learning, *Frontiers in human neuroscience* 13 (2019) 295.

-
- [85] R. Yan, C. Wu, Y. Wang, Exploration and evaluation of individual difference to driving fatigue for high-speed railway: a parametric svm model based on multidimensional visual cue, *IET Intelligent Transport Systems* 12 (6) (2018) 504–512.
- [86] M. Awais, N. Badruddin, M. Drieberg, A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability, *Sensors* 17 (9) (2017) 1991.
- [87] R. Liu, B. Reimer, S. Song, B. Mehler, E. Solovey, Unsupervised fnirs feature extraction with cae and esn autoencoder for driver cognitive load classification, *Journal of Neural Engineering* 18 (3) (2021) 036002.
- [88] R. Elul, The genesis of the eeg, *International review of neurobiology* 15 (1972) 227–272.
- [89] R. R. Whelan, Neuroimaging of cognitive load in instructional multimedia, *Educational Research Review* 2 (1) (2007) 1–12.
- [90] S. Murugan, J. Selvaraj, A. Sahayadhas, Detection and analysis: Driver state with electrocardiogram (ecg), *Physical and engineering sciences in medicine* 43 (2) (2020) 525–537.
- [91] S. Arefnezhad, A. Eichberger, M. Frühwirth, C. Kaufmann, M. Moser, I. V. Koglbauer, Driver monitoring of automated vehicles by classification of driver drowsiness using a deep convolutional neural network trained by scalograms of ecg signals, *Energies* 15 (2) (2022) 480.
- [92] A. Chowdhury, R. Shankaran, M. Kavakli, M. M. Haque, Sensor applications and physiological features in drivers' drowsiness detection: A review, *IEEE sensors Journal* 18 (8) (2018) 3055–3067.
- [93] N. Li, D. Yang, L. Jiang, H. Liu, H. Cai, Combined use of fsr sensor array and svm classifier for finger motion recognition based on pressure distribution map, *Journal of Bionic Engineering* 9 (1) (2012) 39–47.
- [94] A. Sadun, J. Jalani, J. Sukor, Force sensing resistor (fsr): a brief overview and the low-cost sensor for active compliance control, in: *First international workshop on pattern recognition*, Vol. 10011, SPIE, 2016, pp. 222–226.
- [95] T. Nguyen, S. Ahn, H. Jang, S. C. Jun, J. G. Kim, Utilization of a combined eeg/nirs system to predict driver drowsiness, *Scientific reports* 7 (1) (2017) 43933.
- [96] S. Yaacob, N. A. I. Affandi, P. Krishnan, A. Rasyadan, M. Yaakop, F. Mohamed, Drowsiness detection using eeg and ecg signals, in: *2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAJET)*, IEEE, 2020, pp. 1–5.
- [97] I. A. Fouad, A robust and efficient eeg-based drowsiness detection system using different machine learning algorithms, *Ain Shams engineering journal* 14 (3) (2023) 101895.
- [98] S. Chaabene, B. Bouaziz, A. Boudaya, A. Hökelmann, A. Ammar, L. Chaari, Convolutional neural network for drowsiness detection using eeg signals, *Sensors* 21 (5) (2021) 1734.
- [99] C. B. S. Maior, M. J. das Chagas Moura, J. M. M. Santana, I. D. Lins, Real-time classification for autonomous drowsiness detection using eye aspect ratio, *Expert Systems with Applications* 158 (2020) 113505.

-
- [100] A. Murata, Proposal of a method to predict subjective rating on drowsiness using physiological and behavioral measures, *IIE Transactions on Occupational Ergonomics and Human Factors* 4 (2-3) (2016) 128–140.
- [101] D. Sommer, M. Golz, Evaluation of perclos based current fatigue monitoring technologies, in: *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, IEEE, 2010, pp. 4456–4459.
- [102] H. Zhang, C. Wu, X. Yan, T. Z. Qiu, The effect of fatigue driving on car following behavior, *Transportation research part F: traffic psychology and behaviour* 43 (2016) 80–89.
- [103] J. Shen, J. Barbera, C. M. Shapiro, Distinguishing sleepiness and fatigue: focus on definition and measurement, *Sleep medicine reviews* 10 (1) (2006) 63–76.
- [104] A. Watson, G. Zhou, Microsleep prediction using an ekg capable heart rate monitor, in: *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, IEEE, 2016, pp. 328–329.
- [105] C. J. de Naurois, C. Bourdin, C. Bougard, J.-L. Vercher, Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness, *Accident Analysis & Prevention* 121 (2018) 118–128.
- [106] S. Haufe, M. S. Treder, M. F. Gugler, M. Sagebaum, G. Curio, B. Blankertz, EEG potentials predict upcoming emergency brakings during simulated driving, *Journal of Neural Engineering* 8 (5) (2011) 056001. doi:10.1088/1741-2560/8/5/056001.
- [107] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, *Journal of Neuroscience Methods* 134 (1) (2004) 9–21. doi:https://doi.org/10.1016/j.jneumeth.2003.10.009.
- [108] The MathWorks Inc., Matlab version: 9.13.0 (r2022b), . (2022).
URL <https://www.mathworks.com>
- [109] B. Mandhouj, M. A. Cherni, M. Sayadi, An automated classification of EEG signals based on spectrogram and CNN for epilepsy diagnosis, *Analog Integrated Circuits and Signal Processing* (2021) 1–10doi:https://doi.org/10.1007/s10470-021-01805-2.
- [110] S. Nisar, O. U. Khan, M. Tariq, An efficient adaptive window size selection method for improving spectrogram visualization, *Computational Intelligence and Neuroscience* 2016 (2016). doi:https://doi.org/10.1155/2016/6172453.
- [111] Y. Ren, Y. Wu, Convolutional deep belief networks for feature extraction of EEG signal, in: *2014 International joint conference on neural networks (IJCNN)*, IEEE, 2014, pp. 2850–2853. doi:10.1109/IJCNN.2014.6889383.
- [112] R. Song, Driver intention prediction using model-added bayesian network, *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* 235 (5) (2021) 1236–1244. arXiv:https://doi.org/10.1177/0954407020968967, doi:10.1177/0954407020968967.
URL <https://doi.org/10.1177/0954407020968967>

-
- [113] R. A. Gougeh, T. Y. Rezaii, A. Farzamnia, An automatic driver assistant based on intention detecting using eeg signal, in: Z. Md Zain, H. Ahmad, D. Pebrianti, M. Mustafa, N. R. H. Abdullah, R. Samad, M. Mat Noh (Eds.), Proceedings of the 11th National Technical Seminar on Unmanned System Technology 2019, Springer Singapore, Singapore, 2021, pp. 617–627.
- [114] X. Zhang, Driver mental states detection during highly automated driving by decoding brain signals, Doctoral dissertation, Technische Universität Berlin (2021).
URL <https://www.depositonce.tu-berlin.de/handle/11303/12787>
- [115] L. Pang, L. Guo, J. Zhang, X. Wanyan, H. Qu, X. Wang, Subject-specific mental workload classification using eeg and stochastic configuration network (scn), Biomedical Signal Processing and Control 68 (2021) 102711.
- [116] T. G. Monteiro, C. Skourup, H. Zhang, A task agnostic mental fatigue assessment approach based on eeg frequency bands for demanding maritime operation, IEEE Instrumentation & Measurement Magazine 24 (4) (2021) 82–88.
- [117] C. Diaz-Piedra, M. V. Sebastián, L. L. Di Stasi, EEG theta power activity reflects workload among army combat drivers: an experimental study, Brain Sciences 10 (4) (2020) 199. doi:
<https://doi.org/10.3390/brainsci10040199>.
- [118] R. T. Schirrneister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggenesperger, M. Tangermann, F. Hutter, W. Burgard, T. Ball, Deep learning with convolutional neural networks for eeg decoding and visualization, Human Brain Mapping 38 (11) (2017) 5391–5420. doi:
<https://doi.org/10.1002/hbm.23730>.
- [119] Y. R. Tabar, U. Halici, A novel deep learning approach for classification of eeg motor imagery signals, Journal of Neural Engineering 14 (1) (2016) 016003. doi:
<https://doi.org/10.1088/1741-2560/14/1/016003>.
- [120] R. Palaniappan, S. Mouli, E. Fringi, H. Bowman, I. Mcloughlin, Incandescent Bulb and LED Brake Lights: Novel Analysis of Reaction Times, IEEE Access 9 (2021) 29143–29152. doi:
[10.1109/ACCESS.2021.3058579](https://doi.org/10.1109/ACCESS.2021.3058579).
- [121] A. Kumar, G. Vashishtha, C. Gandhi, Y. Zhou, A. Glowacz, J. Xiang, Novel convolutional neural network (NCNN) for the diagnosis of bearing defects in rotary machinery, IEEE Transactions on Instrumentation and Measurement 70 (2021) 1–10. doi:
[10.1109/TIM.2021.3055802](https://doi.org/10.1109/TIM.2021.3055802).
- [122] M. A. Nielsen, Neural Networks and Deep Learning, Vol. 2018, Determination press San Francisco, CA, 2015.
URL <http://neuralnetworksanddeeplearning.com>
- [123] Laboratorio Gesitran, Driving simulator manual in the "Laboratorio Gesitran", University of Concepcion, 2nd Edition (2022).
- [124] H. J. Mora, E. J. Pino, Simplified prediction method for detecting the emergency braking intention using eeg and a cnn trained with a 2d matrices tensor arrangement, International Journal of Human–Computer Interaction 39 (3) (2023) 587–600.
- [125] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea, M. E. Lopez, Real-time system for monitoring driver vigilance, IEEE Transactions on intelligent transportation systems 7 (1) (2006) 63–77.

-
- [126] A. Kuzdeuov, D. Koishigarina, D. Aubakirova, S. Abushakimova, H. A. Varol, Sf-tl54: A thermal facial landmark dataset with visual pairs, in: 2022 IEEE/SICE International Symposium on System Integration (SII), IEEE, 2022, pp. 748–753.
- [127] R. Schleicher, N. Galley, S. Briest, L. Galley, Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired?, *Ergonomics* 51 (7) (2008) 982–1010.
- [128] L. K. McIntire, R. A. McKinley, C. Goodyear, J. P. McIntire, Detection of vigilance performance using eye blinks, *Applied ergonomics* 45 (2) (2014) 354–362.
- [129] W. Wierwille, L. Ellsworth, S. Wreggit, R. Fairbanks, C. Kirn, Tresearch on vehiclebased driver status/performance monitoring: development, validation, and refinement of algorithms for detection of driver drowsiness, National Highway Traffic Safety Administration. Final Report: DOT HS 808 247 (1994).
- [130] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer, Smote: synthetic minority over-sampling technique, *Journal of artificial intelligence research* 16 (2002) 321–357.
- [131] I. Tomek, Two modifications of cnn., *IEEE Transactions on systems, Man, and Cybernetics* 6 (1976) 769–772.
- [132] G. E. Batista, A. L. Bazzan, M. C. Monard, et al., Balancing training data for automated annotation of keywords: a case study., *Wob* 3 (2003) 10–8.
- [133] R. M. Pereira, Y. M. Costa, C. N. Silla Jr, Mtl: A multi-label approach for the tomek link undersampling algorithm, *Neurocomputing* 383 (2020) 95–105.
- [134] S. Sawangarreerak, P. Thanathamthee, Random forest with sampling techniques for handling imbalanced prediction of university student depression, *Information* 11 (11) (2020) 519.
- [135] Z. Li, F. Liu, W. Yang, S. Peng, J. Zhou, A survey of convolutional neural networks: analysis, applications, and prospects, *IEEE transactions on neural networks and learning systems* (2021).
- [136] S. Targ, D. Almeida, K. Lyman, Resnet in resnet: Generalizing residual architectures, *arXiv preprint arXiv:1603.08029* (2016).
- [137] R. Wightman, H. Touvron, H. Jégou, Resnet strikes back: An improved training procedure in timm, *arXiv preprint arXiv:2110.00476* (2021).
- [138] Z. Wu, C. Shen, A. Van Den Hengel, Wider or deeper: Revisiting the resnet model for visual recognition, *Pattern Recognition* 90 (2019) 119–133.
- [139] C. Szegedy, S. Ioffe, V. Vanhoucke, A. Alemi, Inception-v4, inception-resnet and the impact of residual connections on learning, in: *Proceedings of the AAAI conference on artificial intelligence*, Vol. 31, 2017, p. .
- [140] L. Zhang, Understanding resnet from a discrete dynamical system perspective, accessed on March 25, 2024 (2021).
URL <http://openreview.net/>
- [141] A. Zhang, Z. C. Lipton, M. Li, A. J. Smola, Dive into deep learning, *arXiv preprint arXiv:2106.11342* (2021).

-
- [142] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: International conference on machine learning, pmlr, 2015, pp. 448–456.
- [143] D. Hu, C. Wang, F. Nie, X. Li, Dense multimodal fusion for hierarchically joint representation, in: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2019, pp. 3941–3945.
- [144] W. S. Parker, Ensemble modeling, uncertainty and robust predictions, Wiley interdisciplinary reviews: Climate change 4 (3) (2013) 213–223.
- [145] P. Baldi, S. Brunak, P. Frasconi, G. Soda, G. Pollastri, Exploiting the past and the future in protein secondary structure prediction, Bioinformatics 15 (11) (1999) 937–946.
- [146] E. Bisong, Google Colaboratory, Apress, Berkeley, CA, 2019, Ch. II, pp. 59–64. doi:10.1007/978-1-4842-4470-8_7.
URL https://doi.org/10.1007/978-1-4842-4470-8_7
- [147] Python Software Foundation, Python, software available from python.org (2019).
URL <https://python.org/>
- [148] A. Martín, A. Ashish, B. Paul, B. Eugene, C. Zhifeng, C. Craig, C. Greg S., D. Andy, D. Jeffrey, D. Matthieu, G. Sanjay, G. Ian, H. Andrew, I. Geoffrey, I. Michael, Y. Jia, J. Rafal, K. Lukasz, K. Manjunath, L. Josh, M. Dandelion, M. Rajat, M. Sherry, M. Derek, O. Chris, S. Mike, S. Jonathon, S. Benoit, S. Ilya, T. Kunal, T. Paul, V. Vincent, V. Vijay, V. Fernanda, V. Oriol, W. Pete, W. Martin, W. Martin, Y. Yuan, Z. Xiaoqiang, TensorFlow: Large-scale machine learning on heterogeneous systems, URL <http://tensorflow.org/>. Software available from tensorflow.org Software available from tensorflow.org (2015).
URL <https://www.tensorflow.org/>
- [149] F. Chollet, et al., Keras, software available from keras.io (2015).
URL <https://keras.io/>
- [150] K. Utsumi, K. Takano, Y. Okahara, T. Komori, O. Onodera, K. Kansaku, Operation of a p300-based brain-computer interface in patients with duchenne muscular dystrophy, Scientific Reports 8 (1) (2018) 1753.
- [151] R. Johnson Jr, E. Donchin, On how p300 amplitude varies with the utility of the eliciting stimuli, Electroencephalography and Clinical Neurophysiology 44 (4) (1978) 424–437.
- [152] L. T. DeCarlo, On the meaning and use of kurtosis., Psychological methods 2 (3) (1997) 292.
- [153] C.-H. Chuang, C.-S. Huang, L.-W. Ko, C.-T. Lin, An eeg-based perceptual function integration network for application to drowsy driving, Knowledge-Based Systems 80 (2015) 143–152.
- [154] S. G. Horowitz, M. Fukunaga, J. A. de Zwart, P. van Gelderen, S. C. Fulton, T. J. Balkin, J. H. Duyn, Low frequency bold fluctuations during resting wakefulness and light sleep: A simultaneous eeg-fmri study, Human brain mapping 29 (6) (2008) 671–682.
- [155] M. A. Kamran, M. M. N. Mannan, M. Y. Jeong, Drowsiness, fatigue and poor sleep’s causes and detection: a comprehensive study, Ieee Access 7 (2019) 167172–167186.

- [156] P. Philip, J. Taillard, C. Guilleminault, S. Quera, B. Bioulac, M. Ohayon, Long distance driving and self-induced sleep deprivation among automobile drivers, *Sleep* 22 (4) (1999) 475–480.
- [157] P. Thiffault, J. Bergeron, Monotony of road environment and driver fatigue: a simulator study, *Accident Analysis & Prevention* 35 (3) (2003) 381–391.
- [158] S. S. Smith, M. S. Horswill, B. Chambers, M. Wetton, Hazard perception in novice and experienced drivers: The effects of sleepiness, *Accident Analysis & Prevention* 41 (4) (2009) 729–733.
- [159] F. Lucidi, P. M. Russo, L. Mallia, A. Devoto, M. Lauriola, C. Violani, Sleep-related car crashes: risk perception and decision-making processes in young drivers, *Accident Analysis & Prevention* 38 (2) (2006) 302–309.
- [160] C. Schwarz, J. Gaspar, R. Yousefian, Multi-sensor driver monitoring for drowsiness prediction, *Traffic injury prevention* 24 (sup1) (2023) S100–S104.

Anexo B: Encuesta de aptitud, salud y experiencia del participante

El siguiente es el formulario utilizado para identificar si el participante cumple con los requisitos para participar en el experimento. Se relacionan preguntas sobre información básica, su condición de salud, experiencia al conducir y el tiempo dedicado a conducir cada semana.

Cuestionario de identificación del conductor

Por favor ingrese los siguientes datos.

* Indica que la pregunta es obligatoria

1. Correo *

2. Nombre completo *

3. Edad *

4. Género *

Marcar solo un óvalo.

Mujer

Hombre

Otro:

5. Seleccione el tipo de licencia vigente que tiene. *

Marca solo un óvalo.

Clase A1

Clase A2

Clase A3

Clase A4

Clase A5

Clase B

Clase D

6. ¿Cuál es su estado civil actual? *

7. Si usted es independiente o vive solo ¿Cuál es su ingreso promedio mensual?

Cuestionario de salud pre-simulación

Este cuestionario consta de 5 preguntas con respuestas SI/NO.

8. En las últimas 24 horas, ¿usted consumió algún producto que contenga cafeína (café, té, coca-cola), chocolate, tabaco o marihuana antes de la simulación? *

Marca solo un óvalo.

Sí

No

9. ¿Usted tiene indicado el uso de lentes al momento de conducir? *

Marca solo un óvalo.

Sí

No

10. ¿Usted tiene indicado el uso de lentes frente a luz artificial? *

Marca solo un óvalo.

Sí

No

11. ¿Usted sufre o ha sufrido alguna vez vértigo (sensación de que el alrededor esta en movimiento cuando no lo está)? *

Marca solo un óvalo.

- Sí
 No

12. ¿Usted tiene alguna enfermedad crónica (cáncer, diabetes, etc.)? Si la respuesta es sí, ¿está en tratamiento y control? *

Marca solo un óvalo.

- Sí
 No

Preguntas complementarias

Datos de información básica.

13. ¿Usted sufre de migraña constante? *

Marca solo un óvalo.

- Sí
 No

14. ¿Usted sufre de claustrofobia? *

Marca solo un óvalo.

- Sí
 No

Cuestionario de experiencia al conducir

Este cuestionario consta de cinco preguntas y está orientado a identificar la experiencia del participante en la conducción de vehículos.

15. Cada vez que usted necesita conducir, ¿por cuántas horas SEGUIDAS conduce un vehículo? *

Marca solo un óvalo.

- Menos de 1 hora.
- Entre 1 y 3 horas.
- Entre 3.1 y 5 horas.
- Más de 5 horas.

16. En los últimos 6 meses, ¿usted ha conducido un vehículo por al menos 3 veces a la semana? *

Marca solo un óvalo.

- Sí
- No
- Otro: _____

17. En los últimos seis meses, ¿por cuántas horas conducía un vehículo en CARRETERA a la semana? *

Marca solo un óvalo.

- Menos de 1 hora.
- Entre 1 y 3 horas.
- Entre 3.1 y 5 horas.
- Más de 5 horas.

18. En los últimos 6 meses, ¿usted ha conducido un vehículo al menos 3 veces a la semana en **CARRETERA**? *

Marca solo un óvalo.

- Sí
 No

19. En los últimos 6 meses, ¿usted ha conducido un vehículo al menos 3 veces a la semana en **la CIUDAD**? *

Marca solo un óvalo.

- Sí
 No

20. En la siguiente pregunta, seleccione solo una respuesta por fila. ¿Qué tipo de vehículos y **por cuanto tiempo** ha conducido usted? *

Selecciona todos los que correspondan.

	No	Menos de 1 año	Entre 1 y 2 años	Entre 2.1 y 4 años	Más de 4 años
Automóvil con capacidad de hasta 5 plazas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Camioneta	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Camión	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Furgón	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Otro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Anexo C: Encuesta post-conducción

Este es el formulario utilizado para recabar información sobre el estado anímico y de salud del participante durante el tiempo del experimento en el simulador.

Encuesta post-conducción

Por favor seleccione una opción en cada una de las siguientes preguntas.

* Indica que la pregunta es obligatoria

1. Correo *

2. Nombre completo *

3. Usted presentó incomodidad general *

Marca solo un óvalo.

No presentó el síntoma

Bajo

Intermedio

Alto

4. Usted presentó fatiga *

Marca solo un óvalo.

No presentó el síntoma

Bajo

Intermedio

Alto

5. Usted presentó dolor de cabeza *

Marca solo un óvalo.

- No presentó el síntoma
- Bajo
- Intermedio
- Alto

6. Usted presentó fatiga visual *

Marca solo un óvalo.

- No presentó el síntoma
- Bajo
- Intermedio
- Alto

7. Usted presentó dificultad al focalizar *

Marca solo un óvalo.

- No presentó el síntoma
- Bajo
- Intermedio
- Alto

8. Usted presentó náuseas *

Marca solo un óvalo.

- No presentó el síntoma
- Bajo
- Intermedio
- Alto

9. Usted presentó cabeza pesada *

Marca solo un óvalo.

- No presentó el síntoma
- Bajo
- Intermedio
- Alto

10. Usted presentó visión borrosa *

Marca solo un óvalo.

- No presentó el síntoma
- Bajo
- Intermedio
- Alto

11. Usted presentó mareo con los ojos abiertos *

Marca solo un óvalo.

- No presentó el síntoma
- Bajo
- Intermedio
- Alto

12. Usted presentó mareo con los ojos cerrados *

Marca solo un óvalo.

- No presentó el síntoma
- Bajo
- Intermedio
- Alto

13. Usted presentó complicación estomacal *

Marca solo un óvalo.

- No presentó el síntoma
- Bajo
- Intermedio
- Alto

14. Usted presentó eructos *

Marca solo un óvalo.

- No presentó el síntoma
- Bajo
- Intermedio
- Alto

Anexo D: Encuesta sobre el estilo de conducción del participante

En este formulario se relacionan preguntas que buscan identificar el estilo de conducción que el participante cree tener al conducir un vehículo.

Cuestionario de estilo de conducción

Por favor, responde a las siguientes preguntas.

* Indica que la pregunta es obligatoria

1. Edad *

2. Género *

Marca solo un óvalo.

- Mujer
 Hombre
 Otro

3. Aunque tenga prioridad de paso reduzco la velocidad en todos los cruces. *

Marca solo un óvalo.

- No me identifico
 Me identifico levemente
 Me identifico moderadamente
 Me identifico frecuentemente
 Me identifico completamente

4. A menudo, me siento presionado a elevar la velocidad en carretera por los conductores que van detrás mío. *

Marca solo un óvalo.

- No me identifico
 Me identifico levemente
 Me identifico moderadamente
 Me identifico frecuentemente
 Me identifico completamente

5. En general, baso mi conducta en carretera en el lema "mejor prevenir que curar". *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

6. Conduzco estando pendiente de las maniobras inesperadas de los otros conductores que van más adelante. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

7. Cuando conduzco en carretera, tengo la capacidad de pensar en otras cosas para pasar el tiempo, sin distraerme de la conducción. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

8. Cuando se pone en verde el semáforo y vehículo de adelante no parte, espero con ^{*} paciencia hasta que lo haga.

Marca solo un óvalo.

- No me identifico
 Me identifico levemente
 Me identifico moderadamente
 Me identifico frecuentemente
 Me identifico completamente

9. Discuto o peleo con otros conductores o peatones. ^{*}

Marca solo un óvalo.

- No me identifico
 Me identifico levemente
 Me identifico moderadamente
 Me identifico frecuentemente
 Me identifico completamente

10. Disfruto del paisaje mientras conduzco. ^{*}

Marca solo un óvalo.

- No me identifico
 Me identifico levemente
 Me identifico moderadamente
 Me identifico frecuentemente
 Me identifico completamente

11. Disfruto de la sensación de acelerar y pasar los cambios de forma rápida. ^{*}

Marca solo un óvalo.

- No me identifico
 Me identifico levemente
 Me identifico moderadamente
 Me identifico frecuentemente
 Me identifico completamente

12. Disfruto la sensación de manejar al límite de la capacidad del vehículo y de las condiciones de la carretera. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

13. Al intentar ingresar a la carretera por una intersección rural en la que no tengo derecho de paso, espero pacientemente a que terminen de pasar los otros vehículos. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

14. Me enojo con aquellos que conducen lento en la pista rápida (pista izquierda). *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

15. En carretera, manejo un poco más encima del límite de velocidad sugerido. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

16. Intento conducir con precaución. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

17. Mientras conduzco trato de relajarme. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

18. Insulto a otros conductores. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

19. Reflexiono y pienso mientras conduzco. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

20. Si quiero realizar un viaje largo con un vehículo, lo planifico con antelación. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

21. Si estoy a tiempo, acelero para pasar un semáforo que justo cambió a luz roja. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

22. Me agrada conversar cuando conduzco, para disminuir el estrés que provoca la congestión vehicular. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

23. Generalmente me pasa que planifico mal una ruta, y en consecuencia, me encuentro en un "taco" que podría haber evitado. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

24. Suelo olvidar que llevo las luces altas hasta que otro conductor me hace luces advirtiéndome de ello. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

25. A veces, quiero encender el limpiaparabrisas y, en su lugar, enciendo las luces. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

26. Generalmente, me siento estresado mientras manejo en carretera, incluso más que en ciudad. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

27. Hay días que me siento frustrado o incapaz de conducir. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

28. Me siento nervioso cuando manejo en carretera. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

29. A veces, tengo que llegar a un lugar y dar más vueltas de las necesarias, sobretodo en la ciudad. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

30. Toco la bocina o hago luces al vehículo de delante como señal de enojo. *

Marca solo un óvalo.

- No me identifico
- Me identifico levemente
- Me identifico moderadamente
- Me identifico frecuentemente
- Me identifico completamente

Anexo E: Formulario de Consentimiento Informado

Documento del formulario de consentimiento informado aprobado por el Comité de Ética, Bioética y Bioseguridad de la Universidad, el cual fue leído y aprobado por cada participante antes de iniciar en el experimento.



Universidad de Concepción
Facultad de Ingeniería
Departamento de Ingeniería Eléctrica



Formulario de consentimiento informado para participar en estudio

Título del estudio	Predicción de eventos tempranos de somnolencia mediante un conjunto de datos multifactoriales en conductores de vehículos.
Tipo de investigación	Proyecto de tesis doctoral
Investigador Responsable	Hermes Javier Mora
Profesor guía de tesis	Dr. Esteban Pino Quiroga
Lugar en el cual se realizará el estudio	Universidad de Concepción, Facultad de Ingeniería, Departamento de Ingeniería civil, Laboratorio de Transportes.
Medio de contacto	h

Si en este formulario de consentimiento informado usted encuentra alguna palabra que no entienda, por favor, solicite al investigador responsable del estudio que le clarifique cualquier palabra o duda que se le presente. Usted recibirá una copia de este formulario de consentimiento informado para discutirlo con quien usted desee.

El objetivo de este documento es proporcionarle la mejor y mayor cantidad de información posible que le ayude a tomar una decisión libre y consiente para decidir si participa o no en el estudio científico "Predicción de eventos tempranos de somnolencia mediante un conjunto de datos multifactoriales en conductores de vehículos".

1. Introducción

Soy Hermes Javier Mora, estudiante del programa de Doctorado en Ciencias de la Ingeniería con mención en Ingeniería Eléctrica de la Facultad de Ingeniería de la Universidad de Concepción. Esta investigación es parte de la propuesta de investigación de mi tesis doctoral la cual busca implementar un nuevo método para detectar la somnolencia en conductores de vehículos.

Antes de que usted decida participar en el estudio, debe leer cuidadosamente todo el formulario. Sienta absoluta libertad para pedir información adicional sobre cualquier duda que tenga para asegurarse de que entiende los procedimientos, riesgos y beneficios de participar de este estudio.

Una vez que haya leído y comprendido la información reportada en este formulario y si usted desea participar voluntariamente en el estudio, se le solicitará firmar el consentimiento informado.

2. Propósito del estudio

La somnolencia en conductores de vehículos causa más de un millón de muertes cada año en el mundo. Los dispositivos que se usan actualmente para detectar la somnolencia no son tan buenos como nos gustaría que fueran. No obstante, para diseñar nuevas técnicas que permitan detectar somnolencia, se requiere obtener señales que muestren cuando el conductor se está durmiendo mientras conduce. Obtener estas señales en las calles mientras el conductor conduce un vehículo real es peligroso. Por esta razón, existe la posibilidad de realizar estos estudios en simuladores de conducción para averiguar si es posible detectar la somnolencia de una manera más efectiva. El simulador de conducción permite conducir un vehículo virtual en un computador





que tiene dos pantallas y accesorios similares a los de un vehículo real. Con este estudio buscamos obtener un conjunto de señales del sueño mientras está realizando la conducción del vehículo en el simulador.

3. Selección de los participantes

Se está invitando a conductores de vehículos (hombres y mujeres) mayores de edad y con experiencia en conducir vehículos para participar de manera voluntaria en la investigación científica sobre detección de somnolencia mientras conduce un vehículo virtual. Para participar en el estudio, se debe cumplir con los siguientes requisitos:

- Ser mayor de edad.
- Tener una licencia de conducir vigente.
- Tener experiencia en la conducción de vehículos.
- Responder a la encuesta sobre su estado de salud y su experiencia en la conducción de vehículos.
- En las últimas 24 horas antes de iniciar la conducción del vehículo en el simulador, no haber consumido ningún tipo de bebidas que contengan cafeína (café, té, bebidas energizantes), no haber consumido ningún tipo de bebidas alcohólicas ni sustancias psicoactivas.

4. Parámetros de exclusión

Usted NO podrá participar de este estudio si presenta alguna de las siguientes condiciones:

- Presenta alguna condición de salud como migraña permanente o claustrofobia.
- Alguna condición de salud que le impida permanecer sentado frente a una pantalla de computador por más de 30 minutos.



5. Procedimiento del estudio

Se necesita obtener un conjunto de datos técnicos relacionados con la conducción y un conjunto de señales fisiológicas de los participantes mientras están conduciendo un vehículo virtual en un simulador de conducción. Buscamos detectar las señales eléctricas del corazón, cerebro, de los músculos y de la piel. Para obtener estas señales se utiliza:

- Un equipo de Electroencefalografía: este dispositivo tiene 32 electrodos con cables que se ubican en un gorro semitransparente que luego se colocan en su cabeza. En cada uno de los 32 electrodos se aplica una sustancia de gel. El equipo no causa dolor ni representa peligro alguno para su salud física o mental.
- Un equipo de Electrocardiografía: este dispositivo tiene tres electrodos que se ubican en su pecho y registran la actividad eléctrica de su corazón. No causa dolor ni representa riesgo alguno para su salud.
- Un equipo de Electromiografía: este dispositivo tiene dos sensores que se ubican sobre los músculos de su pierna derecha y registran la actividad eléctrica de esa pierna. No presenta ningún riesgo para su salud.
- Equipo de resistencia galvánica para registrar la actividad sobre la piel en su mano izquierda.

5.1. Para colocar los sensores que se han mencionado, se procederá con los siguientes pasos:

- Con anticipación se le informará el día y la hora en la cual se le solicitará presentarse en el laboratorio.
- Usted deber haber leído completamente el presente formulario y si decide participar, firmar el consentimiento informado.
- Si usted acepta participar, debe responder a una encuesta de 10 preguntas sobre su salud física y su experiencia al conducir vehículos.



- Si en este punto usted sigue participando, se colocará sobre su cabeza un gorro que tiene 32 sensores con cables.
- En cada uno de los 32 sensores se aplicará una cantidad (aproximadamente 2 gramos) de gel conductor que no genera ningún riesgo para su salud física ni mental.
- En su pecho se colocarán tres sensores que registran la actividad eléctrica de su corazón.
- En los músculos de su pierna derecha se colocarán dos sensores que registran la actividad de esta pierna.
- En su antebrazo izquierdo se colocarán dos sensores que registran la actividad de su piel.
- Se le solicitará sentarse en la silla del simulador para conducir un vehículo virtual en un computador siguiendo una ruta virtual previamente diseñada.
- Al terminar la simulación, se procederá a retirar cada uno de los sensores colocados en su cuerpo.

5.2. Si usted acepta participar del estudio, debe cumplir los siguientes requisitos:

- Presentarse en el lugar, fecha y hora previamente indicadas.
- 24 horas antes de que usted llegue al laboratorio, no debe consumir ningún tipo de bebidas alcohólicas, sustancias psicoactivas y ninguna clase de bebidas que contengan cafeína (café, té, bebidas energizantes).
- Antes de llegar al laboratorio, se le solicitará lavar y secar completamente su cabello.

Todo el tiempo que permanezca en el simulador de conducción estará únicamente usted, pero su actividad y seguridad van a monitorearse a través de una cámara conectada a un monitor en la sala de control del simulador. También debe conocer que todas las señales que se obtengan con estos dispositivos se guardarán en un computador y serán parte exclusiva del proyecto de investigación. Para garantizar el anonimato de todos los participantes del estudio, los datos personales obtenidos no se entregarán a terceras personas. Las señales y datos fisiológicos obtenidos tampoco podrán entregarse a ninguno de los participantes, incluso si usted decide retirarse del estudio en cualquier momento.

6. Riesgos asociados con el estudio

A juicio de los investigadores y la información asociada a experimentos similares, el estudio con estos dispositivos no tiene ningún tipo de riesgo físico o psicológico para usted. Sus respuestas a la encuesta no le ocasionarán ningún tipo de riesgo ni tendrán consecuencias de ninguna naturaleza.

Durante la ubicación de los sensores y la realización de este estudio, es posible que usted pueda presentar incomodidad:

- El gel que se aplicará en cada uno de los 32 sensores colocados en su cabeza no causa ningún tipo de lesión física ni daño psicológico. Al terminar el estudio y quitar todos los sensores, usted puede necesitar lavar su cabello para quitar algún residuo del gel aplicado.
- Ninguno de los sensores ni dispositivos que se utilizarán durante el estudio causan daño a su salud física y mental, pero pueden resultarle incómodos o causar irritación leve en la piel.
- Las pantallas del computador en el simulador de conducción pueden resultarle molestas en algún momento durante la simulación.

7. Beneficios del estudio científico

En estudios similares realizados por otros investigadores, se ha observado que es posible detectar somnolencia mientras se conduce un vehículo en un computador. Con su participación es posible realizar este experimento y obtener una nueva y eficiente base de datos para detectar somnolencia en los conductores para ayudar a reducir el número de accidentes de tránsito. Además, su participación puede contribuir a producir conocimiento científico en la Universidad de Concepción.



8. Compensación

No se dará ninguna suma de dinero o regalo por tomar parte en esta investigación. Recuerde que, en cualquier momento del estudio, usted puede dejar de participar, aun si usted ha firmado el formulario de consentimiento informado.

9. Confidencialidad

La identidad de aquellas personas que participen en esta investigación científica no se compartirá con nadie. La información personal que se recoja de este proyecto de investigación se pondrá fuera del alcance del público y solo el investigador responsable tendrá acceso a ella. Cualquier información acerca de usted tendrá un número en vez de su nombre. Sus datos de la encuesta se manejarán de manera confidencial y estarán almacenados en un computador que requiere de una clave de acceso. No habrá ningún dato público que lo pueda identificar a usted.

10. Participación voluntaria

Su participación en esta investigación es totalmente voluntaria. Usted puede cambiar de idea más tarde y dejar de participar en cualquier momento, aun cuando haya aceptado y firmado el consentimiento informado. Su decisión de participar o no en este estudio no le afectará de ninguna forma en su relación con la Universidad de Concepción.

Usted podrá solicitar información relacionada con el proyecto de investigación en el momento que lo estime pertinente al investigador Hermes Javier Mora h [@udec.cl](mailto:hj.mora@udec.cl) o al profesor guía Dr. Esteban Pino e [@biomedica.udec.cl](mailto:epino@biomedica.udec.cl)

11. Formulario de consentimiento informado

El Comité de Ética, Bioética y Bioseguridad de la Vicerrectoría de Investigación y desarrollo de la Universidad de Concepción (decreto U. DE C. Nº. 2017-083) es el encargado de revisar y avalar la realización de este estudio para que cumpla con todos los lineamientos éticos. Este Comité está conformado por un grupo de personas independientes del estudio, que evalúa el cumplimiento de la normativa ética nacional e internacional para asegurar la protección de los derechos, la seguridad y bienestar de los seres humanos involucrados en una investigación.

Si usted tiene dudas o consultas sobre sus derechos como participante de la investigación, puede comunicarse con el Comité de Ética, Bioética y Bioseguridad de la Universidad de Concepción: dirección Victoria 490 Concepción, teléfono 412204302.

Con el propósito de velar por el fiel cumplimiento del consentimiento informado, es posible que usted sea contactado por algún integrante del Comité de Ética, Bioética y Bioseguridad para realizar una breve entrevista o encuesta, en forma personal, telefónica o electrónica. El investigador responsable del estudio estará al tanto de la comunicación entre usted y el Comité. Usted puede negarse a participar de la encuesta si así lo desea y sin que afecte su participación en el estudio. *Tenga presente que, la firma de este formulario de consentimiento informado significa su aceptación consiente y voluntaria en el estudio de investigación. Por favor lea detenida y cuidadosamente la siguiente información.*





Formulario:

Yo _____, RUT _____, he sido invitado a participar de una investigación científica que involucra conducir un vehículo virtual en un simulador de conducción durante aproximadamente 2.5 horas. Se me ha informado que en algunas partes de mi cuerpo se colocaran sensores no invasivos para registrar la actividad eléctrica de mi cerebro, corazón, movimiento de la pierna derecha y de la piel. Se me ha informado de los posibles riesgos y consecuencias que pudiera traer para mi persona. Se me ha dado a entender que no se me compensará más allá de los gastos de viaje. Se me ha proporcionado los nombres de dos personas del estudio que pueden ser fácilmente contactadas usando el nombre y dirección de correo electrónico suministrados. Se me ha informado sobre mi derecho de consultar al Comité de Ética, Bioética y Bioseguridad de la Universidad de Concepción si llego a tener alguna duda con mis derechos como participante de esta investigación científica.

Todas mis dudas se han resuelto de manera satisfactoria y he leído y comprendido toda la información reportada en el formulario de consentimiento informado. Entiendo que mi participación en este estudio es voluntaria y podré renunciar a participar en cualquier momento. Comprendo y acepto que los datos no personales obtenidos del estudio pueden publicarse o difundirse con fines científicos y educativos sin que puedan identificarme de ninguna manera. Entiendo que mi información personal no se compartirá con nadie y solo los investigadores del estudio conocen los datos que yo les he proporcionado de manera voluntaria. Entiendo y acepto que los datos obtenidos no me pueden devolverse incluso si decido retirarme del estudio en cualquier momento. De manera voluntaria, libre y consiente acepto participar en el estudio que se enmarca en la investigación titulada "Predicción de eventos tempranos de somnolencia mediante un conjunto de datos multifactoriales en conductores de vehículos".

Entiendo que recibiré una copia firmada de este formulario de consentimiento informado.

Nombre del participante: _____ **Teléfono:** _____
Firma del participante: _____ **E-mail:** _____
Año _____ **mes** _____ **día** _____



De manera atenta he explicado el propósito de la investigación. He explicado acerca de los riesgos y beneficios que implica su participación. He contestado a cada una de las preguntas en la medida de lo posible y he preguntado si tiene dudas adicionales. Una vez se terminó la lectura del formulario y se respondió a todas las preguntas, se procedió a firmar el presente documento.

Nombre del Investigador Responsable: _____ **RUT:** _____
Firma del investigador: _____ **E-mail:** _____
día _____ **mes** _____ **año** _____ **Teléfono: (+56)** _____

**Nombre del Director del centro o su delegado/
Ministro de Fe:** _____ **RUT:** _____
Firma del ministro de fe: _____ **E-mail:** _____
día _____ **mes** _____ **año** _____ **Teléfono:** _____



Universidad de Concepción
Facultad de Ingeniería
Departamento de Ingeniería Eléctrica

Formulario para revocar el consentimiento informado y retirarse del estudio

Al Investigador Responsable Hermes Javier Mora, yo _____,
RUT _____, por este medio manifiesto de forma libre y consiente mi decisión de retirarme de
este estudio de investigación.

Estoy informado de que no necesito justificar esta mi decisión y que, de parte del Investigador Responsable y
en el marco de este estudio, no habrá sanción o reproche hacia mi persona.

Entiendo que recibiré una copia firmada de este formulario.

día _____ mes _____ año _____

Nombre del Participante: _____

Firma del Participante: _____ **Teléfono:** _____



Nombre del Investigador Responsable: _____

Firma del investigador: _____ **E-mail:** _____

Teléfono: (+56) _____

Hermes Javier Mora Pantoja.
Universidad de Concepción, 2024.
