

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE INGENIERÍA EN SISTEMAS

**DISEÑO Y APLICACIÓN DE UN MODELO DE RECONOCIMIENTO
DE 11 GESTOS DE LA MANO USANDO SEÑALES EMG, IMU, DTW
Y KNN**

**DISEÑO DE UN MODELO DE RECONOCIMIENTO DE 11 GESTOS
DE LA MANO QUE FUNCIONE EN TIEMPO REAL USANDO DTW Y
KNN**

**TRABAJO DE INTEGRACIÓN CURRICULAR PRESENTADO COMO
REQUISITO PARA LA OBTENCIÓN DEL TÍTULO DE INGENIERO EN
CIENCIAS DE LA COMPUTACIÓN**

ÓSCAR MAURICIO RIVERA CAJÍA

oscar.rivera@epn.edu.ec

DIRECTOR: ÁNGEL LEONARDO VALDIVIESO CARAGUAY

angel.valdivieso@epn.edu.ec

DMQ, SEPTIEMBRE 2022

CERTIFICACIONES

Yo, ÓSCAR MAURICIO RIVERA CAJÍA declaro que el trabajo de integración curricular aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.



ÓSCAR MAURICIO RIVERA CAJÍA

Certifico que el presente trabajo de integración curricular fue desarrollado por ÓSCAR MAURICIO RIVERA CAJÍA, bajo mi supervisión.



ÁNGEL LEONARDO VALDIVIESO CARAGUAY
DIRECTOR

DECLARACIÓN DE AUTORÍA

A través de la presente declaración, afirmamos que el trabajo de integración curricular aquí descrito, así como el (los) producto(s) resultante(s) del mismo, son públicos y estarán a disposición de la comunidad a través del repositorio institucional de la Escuela Politécnica Nacional; sin embargo, la titularidad de los derechos patrimoniales nos corresponde a los autores que hemos contribuido en el desarrollo del presente trabajo; observando para el efecto las disposiciones establecidas por el órgano competente en propiedad intelectual, la normativa interna y demás normas.

ÓSCAR MAURICIO RIVERA CAJÍA

ÁNGEL LEONARDO VALDIVIESO CARAGUAY

DEDICATORIA

Dedico este trabajo principalmente a mis padres Óscar Rivera y Guadalupe Cajía, a quienes debo en gran medida haber llegado a este punto de mi vida. Dedico este trabajo también a mi hermana como símbolo de referencia para su avance. Finalmente, este trabajo es dedicado a la memoria de una amiga.

AGRADECIMIENTO

Agradezco a todos los profesores que intervinieron en mi formación académica y en especial, a aquellos que además de su labor como profesores destacaron, a mi juicio, como inspiración y complemento a mis principios.

ÍNDICE DE CONTENIDO

CERTIFICACIONES.....	I
DECLARACIÓN DE AUTORÍA.....	II
DEDICATORIA.....	III
AGRADECIMIENTO.....	IV
ÍNDICE DE CONTENIDO.....	V
RESUMEN	VI
ABSTRACT	VII
3 DESCRIPCIÓN DEL COMPONENTE DESARROLLADO.....	1
3.1 Objetivo general.....	2
3.2 Objetivos específicos	2
3.3 Alcance	2
3.4 Marco teórico	3
4 METODOLOGÍA.....	6
4.1 FASE DE DISEÑO.....	6
4.2 FASE DE IMPLEMENTACIÓN.....	7
4.3 FASE DE EVALUACIÓN DE RESULTADOS.....	13
4.4 FASE DE INTEGRACIÓN.....	13
5 RESULTADOS, CONCLUSIONES Y RECOMENDACIONES.....	16
5.1 Resultados	16
5.2 Conclusiones.....	21
5.3 Recomendaciones.....	21
6 REFERENCIAS BIBLIOGRÁFICAS	23
7 ANEXOS.....	25

RESUMEN

En el presente trabajo se desarrolló un modelo de reconocimiento de gestos de la mano en tiempo real. Las entradas para este modelo son señales de movimiento de la mano medidas por electromiografía de superficie (EMG) y por una unidad de medida inercial (IMU). Ambos tipos de señales son obtenidas a partir de los sensores comerciales Myo Armband y G-Force Pro. Por su parte, la salida del modelo es la etiqueta clasificada del gesto que el usuario ejecuta en cada momento y el instante de tiempo en donde se realizó el gesto. Este modelo puede reconocer 11 gestos basado en los algoritmos KNN y DTW. La evaluación del modelo se llevó a cabo midiendo el tiempo de respuesta y la exactitud de clasificación y reconocimiento.

PALABRAS CLAVE: Machine Learning, Hand Gesture Recognition (HGR), K-Nearest Neighbors (KNN), Dynamic Time Warping (DTW), Electromyography (EMG), Inertial Measurement Unit (IMU)

ABSTRACT

In this work, a Real-Time Hand Gesture Recognition model was developed. The inputs of this model are hand movement signals measured by surface electromyography (EMG) and by an inertial measurement unit (IMU). Both types of signals are obtained from the commercial sensors Myo Armband and G-Force Pro. For its part, the output is the classified label and the instant of the gesture that the user executes any moment. This model can recognize 11 gestures based on the KNN and DTW algorithms. The model was evaluated by measuring the response time and the classification and recognition accuracy.

KEYWORDS: Machine Learning, Hand Gesture Recognition (HGR), K-Nearest Neighbors (KNN), Dynamic Time Warping (DTW), Electromyography (EMG), Inertial Measurement Unit (IMU)

3 DESCRIPCIÓN DEL COMPONENTE DESARROLLADO

En el presente componente se desarrolló un Modelo de Reconocimiento de Gestos de la Mano (HGR, por sus siglas en inglés) basado en señales electromiográficas (EMG) y señales del movimiento del brazo obtenidas de una Unidad de Medida Inercial (IMU, por sus siglas en inglés); el mismo fue desarrollado para reconocer 11 gestos de la mano en tiempo real. En este modelo, los módulos de extracción de características y clasificación fueron construidos usando los algoritmos Deformación dinámica del tiempo (DTW, por sus siglas en inglés) y K vecinos más cercanos (KNN, por sus siglas en inglés), respectivamente. DTW se usó como función de distancia con el fin de calcular la similitud entre señales, las mismas que son utilizadas en el proceso de clasificación utilizando el algoritmo de Machine Learning (ML, por sus siglas en inglés). Por su parte, KNN fue seleccionado por ser uno de los algoritmos más simples de implementar y su eficacia al resolver problemas de clasificación [1]. El funcionamiento en tiempo real implica que el modelo da una respuesta en un tiempo menor a 300 ms [2]. El desempeño del modelo desarrollado fue evaluado en tiempo de respuesta y términos de exactitud de clasificación y reconocimiento.

El abordaje del presente componente fue llevado a cabo a través de una metodología general de ML en algoritmos de aprendizaje supervisado [3] dividida en las fases de Diseño, Implementación y Evaluación. El modelo, tiene la siguiente estructura: adquisición de datos, preprocesamiento, extracción de características y clasificación.

Finalmente, el modelo HGR a través de una última y cuarta fase fue integrado al segundo componente del proyecto con el fin de darle al modelo HGR un uso útil mediante una aplicación de interfaz humano-máquina. La aplicación fue definida e implementada en el segundo componente del TIC.

3.1 Objetivo general

Desarrollar un modelo de reconocimiento, en tiempo real, de 11 gestos de la mano usando DTW y KNN y que tengan como entrada señales EMG y señales IMU.

3.2 Objetivos específicos

1. Diseñar un modelo de reconocimiento de 11 gestos de la mano que funcione en tiempo real usando DTW y KNN y que tengan como entrada señales EMG e IMU.
2. Validar el modelo propuesto en términos de su exactitud de clasificación, reconocimiento y tiempo de respuesta.
3. Integrar el modelo HGR con una aplicación de interfaz humano-máquina que utilice gestos de la mano, realizada en el segundo componente del proyecto.

3.3 Alcance

El desarrollo del presente trabajo abarca la aplicación de una metodología general de ML en donde se contemplan las etapas de Diseño, Implementación, Evaluación e Integración de un modelo HGR para el reconocimiento de 11 gestos de la mano los cuales están ilustrados en la Figura 1:

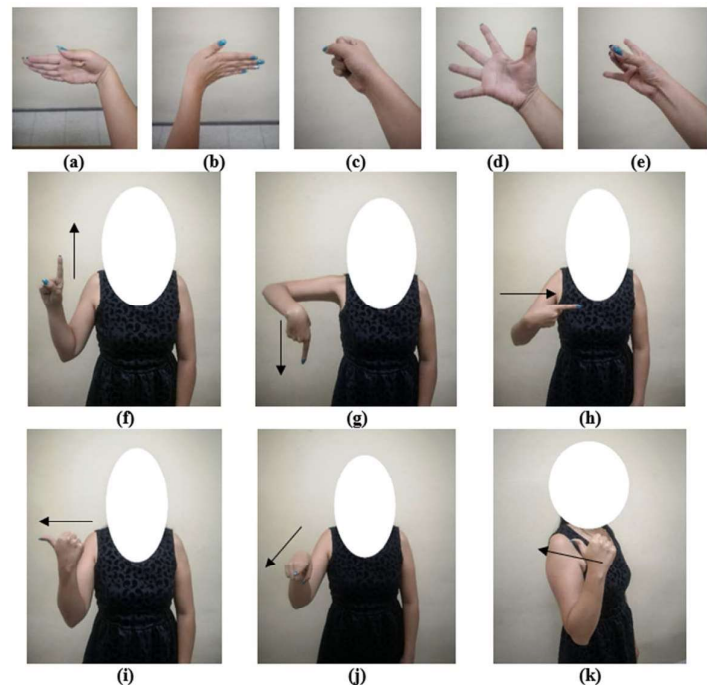


Figura 1. (a) waveIn, (b) waveOut, (c) fist, (d) open, (e) pinch, (f) up, (g) down, (h) left, (i) right, (j) forward, y (k) backward.

Para llevar a cabo la evaluación del modelo se realiza, por una parte, la tarea de medir el tiempo de procesamiento, el cual hace referencia al tiempo que el modelo tarda en devolver una respuesta. Adicionalmente, la evaluación contempla medir las tasas de exactitud en cuanto a reconocimiento y clasificación. La tasa de exactitud de reconocimiento hace referencia a la capacidad del modelo para detectar en qué instante se está realizando el gesto. Por otra parte, la tasa de exactitud de clasificación se refiere a la capacidad del modelo para clasificar correctamente o no la etiqueta el gesto.

La integración del modelo HGR con el segundo componente del presente proyecto se llevará a cabo mediante un seguimiento con el fin de adaptar y ajustar el modelo para un funcionamiento en línea adecuado. Esto conlleva que el modelo HGR debe de ser adaptado a la clasificación de señales transmitidas en línea mediante un sensor.

3.4 Marco teórico

3.4.1 Modelo HGR, sensores EMG e IMU

Los modelos HGR son sistemas humano-computador que determinan: qué gesto es realizado y el momento en el que el gesto es realizado [2]. La adquisición de datos para estos modelos puede llevarse a cabo a través de diferentes herramientas como, por ejemplo, el uso de guantes, IMUs, sensores de visión o sensores de EMG de superficie.

Para el presente componente, la adquisición de datos se lleva a cabo usando una combinación de sensores los cuales son: sensores de EMG de superficie y una IMU. Los sensores de EMG de superficie son dispositivos que, a través de electrodos integrales, miden la actividad eléctrica producida por contracción de un músculo [4]. Por su parte, la IMU es un dispositivo que combina 3 sensores: un giroscopio de 3 ejes que mide el cambio de velocidad de ángulo, un acelerómetro de 3 ejes que mide la inclinación y un magnetómetro de 3 ejes, el cual mide la orientación de la IMU [5].

3.4.2 Algoritmo K-Nearest Neighbors

KNN es un algoritmo de clasificación no paramétrico dentro de los algoritmos de aprendizaje supervisado de Machine Learning. KNN resulta ser sencillo de ejecutar y óptimo en ejecución, pues al carecer de proceso de aprendizaje, su funcionamiento se basa en almacenar un conjunto de datos para posteriormente realizar la clasificación comparando una nueva muestra con el conjunto de datos.

La clasificación se establece determinando un número de datos que presenten características más similares a la nueva muestra y asignándole la categoría de la mayoría de esos datos más cercanos. El número de datos más cercanos K desempeña un papel muy importante en el rendimiento de la clasificación pues decide qué número de vecinos se deben de considerar. [6]. En el componente, el algoritmo K-NN tiene la función de clasificar la señal de un gesto en base al cálculo de su similitud con las categorías de gestos previamente almacenadas.

3.4.3 Algoritmo Dynamic Time Warping

DTW es un algoritmo ampliamente utilizado como una medida de similitud entre series temporales en varios dominios [7]. Esta medida de distancia, en comparación con otras usadas comúnmente en ML (Minkowski, Euclidiana o Manhattan), presenta la notable ventaja de invariancia frente a la condición de deformación de una señal [8]. Por ejemplo, dos frecuencias de muestreo diferentes podrían generar dos piezas de señales, aunque la primera señal es solo una versión comprimida de la otra. En este caso, será muy diferente y desviado de la verdad usar la distancia euclidiana puntual. Por el contrario, DTW capturaría bien esa escala y generaría una distancia muy pequeña entre ellos.

Dada la tarea de clasificación, KNN debe calcular la similitud entre señales de gestos de la mano representadas como series temporales, las cuales pueden presentar deformaciones, tales como desplazamiento o escalado en el eje del tiempo. Dado este escenario de deformación en señales, DTW resulta ser una medida de distancia conveniente para efectuar la tarea de coincidencia de patrones.

3.4.4 Clasificación y Reconocimiento de gestos

Al desarrollo de modelos HGR se le puede dar dos enfoques: la clasificación y el reconocimiento del gesto. La diferencia entre ambos enfoques radica en que la clasificación de gestos identifica tan solo la etiqueta del gesto mientras que el reconocimiento de gestos además de la etiqueta del gesto identifica el instante de tiempo en el que el gesto es ejecutado[9]. Esta diferencia de conceptos se ilustra en la Figura 2.

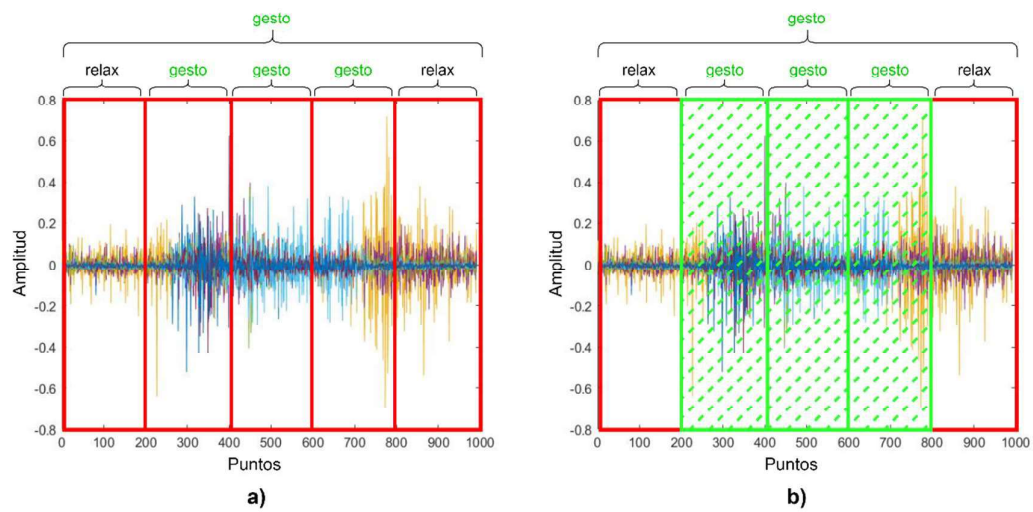


Figura 2. (a) Clasificación y (b) Reconocimiento de gestos

Es importante destacar que la revisión de varios trabajos en la literatura científica muestra que generalmente se centran en la clasificación del gesto mas no en el reconocimiento [10-13]. En el presente trabajo se busca cubrir ambos enfoques: clasificación y reconocimiento de gestos.

4 METODOLOGÍA

4.1 FASE DE DISEÑO

Se realizó una revisión del estado del arte en relación con modelos HGR que hayan trabajado con señales EMG, IMU y ML. La revisión tuvo como fin recopilar información de valor útil para construir la estructura que constituiría el modelo HGR a desarrollar. La Figura 3 presenta una descripción gráfica de la estructura del modelo, el mismo que se divide en 4 etapas: adquisición de datos, preprocesamiento, extracción de características y clasificación.

El modelo tiene como entrada las señales EMG e IMU correspondientes al gesto a clasificar más los correspondientes datasets para la etapa de clasificación. En la etapa de adquisición de datos se usa la técnica de ventana deslizante; a continuación, se realiza un preprocesamiento para realizar la extracción de características. Finalmente se realiza la clasificación usando algoritmos KNN y DTW dando como salida del modelo la predicción del gesto de entrada.

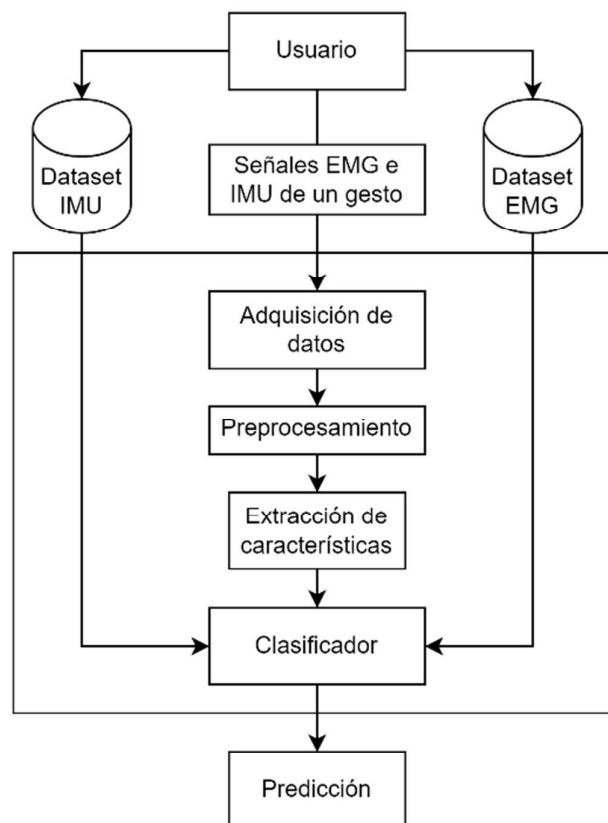


Figura 3. Estructura del modelo

Los gestos planteados, mostrados en la Figura 1, fueron divididos según su tipo de ejecución como se muestra en la Tabla 1. Aquellos gestos en los que la contracción de los músculos del antebrazo es responsable de los movimientos de las diferentes partes de la mano se denominaron gestos estáticos [14]. Los gestos en los que la trayectoria del movimiento del brazo impera en su ejecución se denominaron gestos dinámicos.

Tabla 1. Organización de gestos según su tipo de ejecución

Gestos estáticos	Gestos dinámicos
a) waveIn (mano a la izquierda)	f) up (apuntar hacia arriba)
b) waveOut (mano a la derecha)	g) down (apuntar hacia abajo)
c) fist (puño)	h) left (apuntar hacia izquierda)
d) open (mano abierta)	i) right (apuntar hacia derecha)
e) pinch (doble golpe de los dedos)	j) forward (apuntar al frente)
	k) backward (apuntar hacia atrás)

4.2 FASE DE IMPLEMENTACIÓN

4.2.1 DESCRIPCIÓN DEL CONJUNTO DE DATOS

Para el desarrollo del presente trabajo se utilizó un conjunto de datos formado por señales EMG e IMU de 85 usuarios, el cual puede encontrarse en Anexo I. Estas señales fueron recopiladas durante la ejecución de 12 gestos: 11 gestos indicados en la Figura 1 más el gesto relax el cual representa un gesto en reposo. La duración de ejecución de cada gesto fue de 5 segundos. Para cada gesto, se ejecutaron 15 repeticiones, componiendo un total de 180 muestras para cada usuario. Nótese que cada uno de los 12 gestos de un usuario cuenta con su correspondiente par de señales EMG e IMU como puede apreciarse en la Tabla 2.

Tabla 2. Organización de gestos para un usuario

n de gesto	Gestos estáticos			n de gesto	Gestos dinámicos		
	Repeticiones	Nombre del gesto	Señales		Repeticiones	Nombre del gesto	Señales
1	1	relax	EMG	7	91	up	EMG
	15		IMU		105		IMU
2	16	waveIn	EMG	8	106	down	EMG
	30		IMU		120		IMU
3	31	waveOut	EMG	9	121	left	EMG
	45		IMU		135		IMU
4	46	fist	EMG	10	136	right	EMG
	60		IMU		150		IMU
5	61	open	EMG	11	151	forward	EMG
	75		IMU		165		IMU
6	76	pinch	EMG	12	166	backward	EMG
	90		IMU		180		IMU

El conjunto de datos fue dividido aleatoriamente en dos grupos: grupo de entrenamiento con 43 usuarios y grupo de prueba con 42 usuarios. Dentro de cada grupo, cada usuario dividió sus repeticiones: grupo de entrenamiento, con 180 repeticiones y grupo de prueba con 180 repeticiones.

Las señales EMG e IMU se recopilaron usando 2 sensores de tipo brazalete: Myo Armband y G-Force Pro. Cada uno posee 8 sensores EMG de superficie y una IMU. En la Tabla 3 se muestra la frecuencia de muestreo para cada sensor de acuerdo con el tipo de señal grabada. Por otro lado, la Figura 4 ilustra la muestra de un gesto la cual está constituida por ambas señales EMG e IMU. Se puede notar que la señal EMG presenta un aspecto bastante irregular a diferencia de la señal IMU, señales presentadas en las Figuras 4a y 4b respectivamente.

Tabla 3. Frecuencia de muestreo para señales EMG e IMU

	Señales EMG	Señales IMU
Myo Armband	200 Hz	50 Hz
G-Force Pro	500 Hz	50 Hz

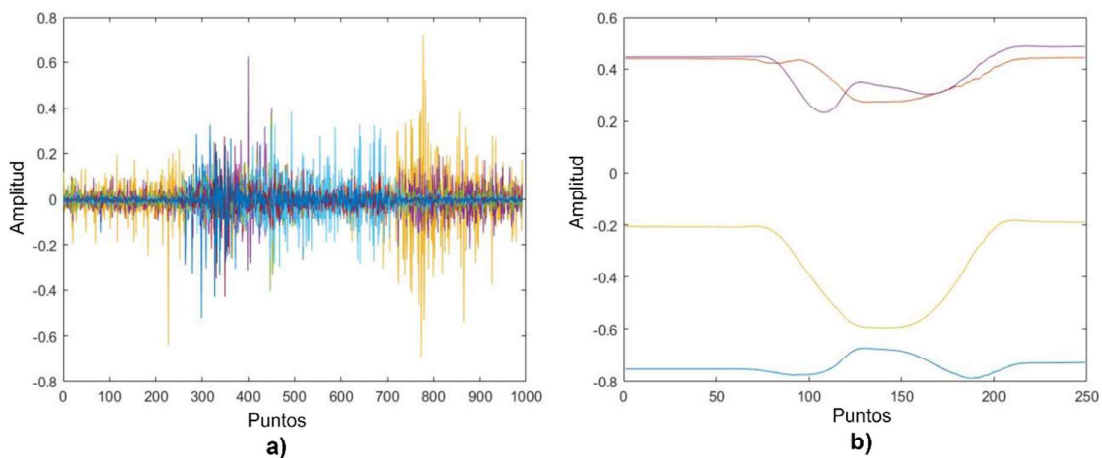


Figura 4. Señales de la ejecución del gesto dinámico Right (a) señal EMG (b) señal IMU

4.2.2 ADQUISICIÓN DE DATOS

Para adquirir los datos de entrada para el clasificador se aplicó la técnica de ventana deslizante, definida por una longitud de ventana y un paso entre ventanas consecutivas. Esta técnica se usa para ambas señales EMG e IMU, tal como se ilustra en la Figura 5. Se definió el mismo valor de paso en el dominio del tiempo para señales EMG e IMU, por lo que el número de iteraciones para analizar la señal completa es la misma.

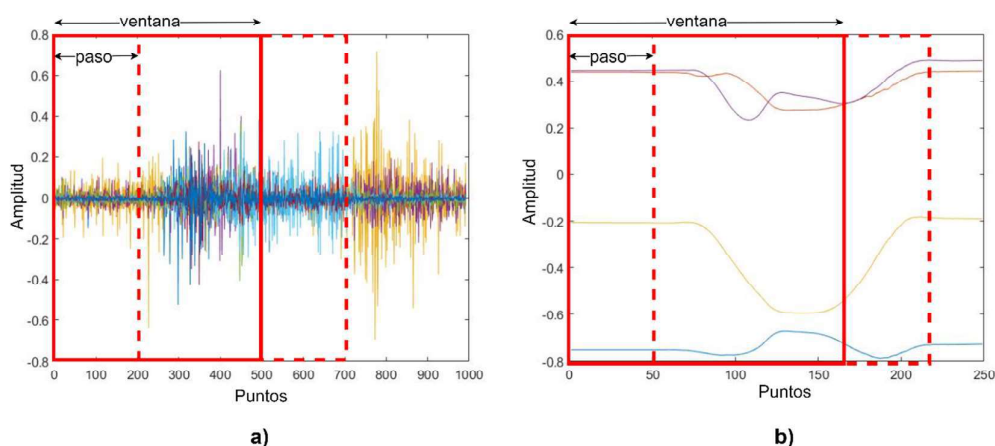


Figura 5. Ventana deslizando aplicada a (a) señal EMG y (b) señal IMU.

Con el proceso de ventana deslizando, las señales se dividen en segmentos más pequeños, los cuáles serán las entradas para la posterior etapa de clasificación. La longitud de ventana para señales EMG e IMU puede variar. De forma experimental se determinó las longitudes óptimas de acuerdo con los resultados de clasificación arrojados por el modelo con cada prueba, según se detalla en la sección de Resultados. Los valores de longitud de ventana pueden variar de acuerdo con el sensor que se haya usado para tomar las señales. A continuación, en la Tabla 2 se muestran los hiper parámetros hallados para la aplicación de la técnica de ventana deslizando.

Tabla 4. Valores de longitudes de ventana

Sensor	paso [s]	ventanaEMG [s]	ventanaIMU [s]
MYO	1	3.4	4.1
GFORCE	1	3.4	3.5

4.2.3 PREPROCESAMIENTO

El preprocesamiento se basa en extraer la envolvente de la señal y tiene como fin facilitar las posteriores fases de extracción de características y clasificación. Fue aplicado solo a señales EMG pues como puede apreciarse en la Figura 6a presentan un aspecto irregular. El preprocesamiento consistió en rectificar la señal, para ello se calculó el valor absoluto de la señal. Luego, para reducir el ruido y de esa forma obtener una señal más suave, se aplicó el filtro Butterworth con cuarto orden y una frecuencia de corte de 5 Hz. El resultado del preprocesamiento de una señal EMG es una señal mucho más suave lo cual se puede apreciar en la Figura 6b. La aplicación del preprocesamiento a las señales IMU no fue necesaria pues estas no se presentan como señales irregulares que puedan complicar posteriores fases del modelo.

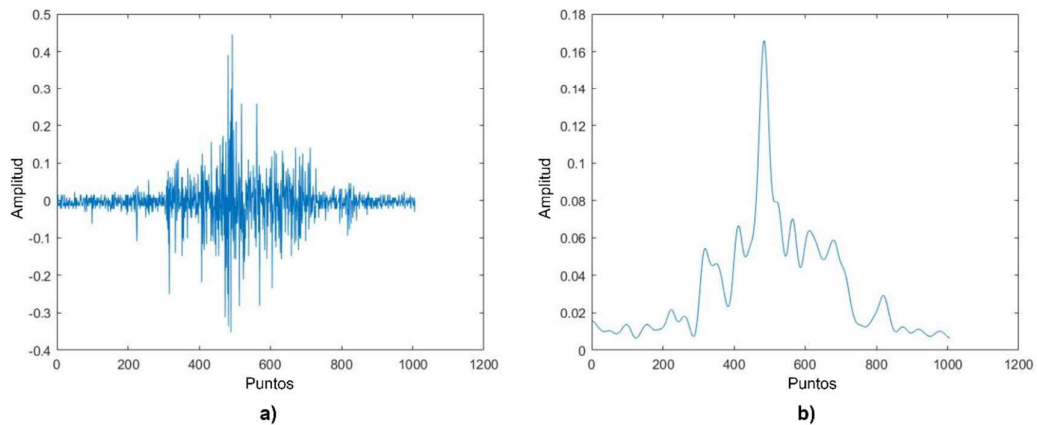


Figura 6. Preprocesamiento (a) señal EMG cruda (b) señal EMG preprocesada

4.2.4 EXTRACCIÓN DE CARACTERÍSTICAS

La extracción de características se basa en el análisis de las ventanas EMG e IMU con el fin de seleccionar la ventana con la que se va a realizar la fase de clasificación. Para este propósito, se usó la función de extracción de características “energía” la cual mide la distribución de energía de ambas señales [15]. La entrada de esta función es la ventana EMG procesada junto a la ventana IMU como se puede apreciar en la Figura 7a.

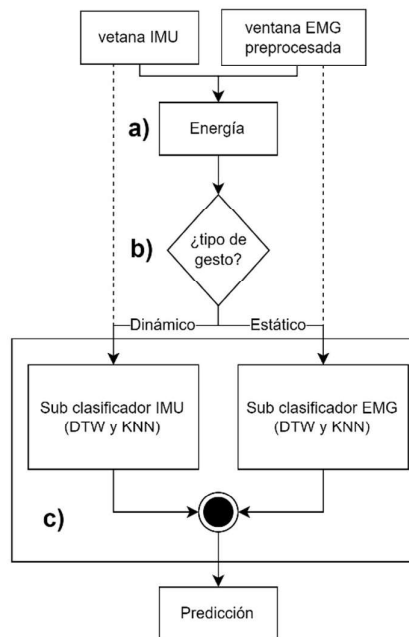


Figura 7. (a) Función “energía” (b) Switch gesto estático o dinámico (c) Clasificador

La salida de esta función es un vector de características que se utiliza como entrada a un switch el cual determina a que tipo de gesto corresponden ambas ventanas EMG e IMU.

El switch está basado en un modelo de clasificación lineal logístico. Este modelo determina si un par de ventanas EMG e IMU de una misma señal corresponden a un gesto estático o dinámico. El modelo fue entrenado usando la misma función “energía” para el dataset de entrenamiento. Su evaluación alcanzó una tasa en promedio del 93.02% de exactitud de clasificación.

4.2.5 CLASIFICADOR

El módulo de clasificación trabaja de manera exclusiva con señales EMG o con señales IMU por separado. En consecuencia, el clasificador se constituye de 2 sub clasificadores como se muestra en la Figura 7c. El primer sub clasificador, al trabajar solo solo con señales EMG, clasifica gestos estáticos. De manera similar, el segundo sub clasificador al trabajar solo con señales IMU, clasifica gestos dinámicos. La razón por la que se trabaja con señales EMG para realizar la clasificación de gestos estáticos es que en la ejecución de gestos estáticos impera la actividad muscular, la cual es medida a través de señales EMG. De manera similar, en la ejecución de gestos dinámicos impera el movimiento del brazo el cual es medido a través de señales IMU. Para determinar con que sub clasificador se debe de llevar a cabo la tarea de clasificación se parte de la predicción de gesto estático o dinámico explicada en la subsección anterior. Como muestran las Figura 7b y 7c, a través de un switch se elige una ventana de características de acuerdo con el tipo de gesto determinado y, en consecuencia, se elige también el sub clasificador correspondiente a la naturaleza de la ventana elegida (EMG o IMU).

Ambos sub clasificadores se construyeron usando el algoritmo KNN. El enfoque en el que KNN se basa en este trabajo es la estimación de la probabilidad condicional basada en la frecuencia relativa de los vecinos más cercanos a la ventana que se quiere clasificar. Para determinar los vecinos más cercanos, son necesarios: (a) datos previamente almacenados y (b) una función de distancia.

(a) Datos previamente almacenados/Datos entrenados

Para el sub clasificador EMG se construyó un conjunto de datos denominado datasetEMG, el cual contiene 90 señales EMG correspondientes a 15 repeticiones de cada uno de los 5 gestos estáticos más el gesto relax. De manera similar, para el segundo sub clasificador se construyó un conjunto de datos denominado datasetIMU, el cual contiene 105 señales IMU correspondientes a 15 repeticiones de cada uno de los 6

gestos dinámicos más el gesto relax. La organización de ambos datasets se ilustra en la Tabla 2.

Tabla 5. Organización de datasets

datasetEMG				datasetIMU			
n de gesto	Repeticiones	Nombre del gesto	Señales	n de gesto	Repeticiones	Nombre del gesto	Señales
1	1	relax	EMG	1	15	relax	IMU
	15						
2	16	waveIn	EMG	7	91	up	IMU
	30						
3	31	waveOut	EMG	8	106	down	IMU
	45						
4	46	fist	EMG	9	121	left	IMU
	60						
5	61	open	EMG	10	136	right	IMU
	75						
6	76	pinch	EMG	11	151	forward	IMU
	90						
				12	166	backward	IMU
					180		

(b) Función de distancia DTW

Para la función de distancia se usó el algoritmo DTW con la distancia Manhattan. Este algoritmo tiene como entradas en el caso del sub clasificador EMG: la ventana EMG y el datasetEMG. En el caso del sub clasificador IMU: la ventana IMU y el datasetIMU. De este modo, DTW calcula la distancia de similitud entre la ventana de entrada con cada señal contenida en el dataset correspondiente a la naturaleza de la ventana que se requiera clasificar.

4.2.6 Hiper parámetros de los sub clasificadores

Para ambos sub clasificadores el valor de k vecinos más cercanos se determinó con la formula $k = \text{ceil}[\log_2 N]$, donde N es el número de muestras que forman el dataset correspondiente al sub clasificador. El umbral se estableció en 80%, para evitar falsos positivos. Los valores de longitudes de ventana, sin embargo, varían para ambos sub clasificadores y fueron determinados de forma experimental. En la Tabla 5 y 6 se muestran los hiper parámetros para cada sub clasificador correspondiente. Es importante destacar que el tamaño del paso entre ventanas consecutivas seleccionado fue de 1

segundo debido a los recursos computacionales disponibles para el desarrollo del presente trabajo.

Tabla 6. Hiper parámetros para los sub clasificadores EMG e IMU

	Sensor	Paso [s]	Ventana [s]	kNN	Umbral
sub clasificador EMG	MYO	1	3.4	7	80%
	GFORCE	1	3.4	7	80%
sub clasificador IMU	MYO	1	4.1	7	80%
	GFORCE	1	4.9	7	80%

Recapitulando, ambos sub clasificadores siguen los siguientes pasos:

1. Elección de k vecinos más cercano a la señal a clasificar usando la función de distancia DTW.
2. Estimación de las probabilidades condicionales colocando la señal a clasificar en la categoría más similar a las categorías disponibles en el dataset.
3. Evaluación de la validez de predicción en base a un umbral. Si se supera el umbral se tiene como salida la etiqueta de ese gesto predicho, de lo contrario, se asigna la etiqueta relax.

La implementación del presente modelo se encuentra disponible en el Anexo II.

4.3 FASE DE EVALUACIÓN DE RESULTADOS

La fase de evaluación de resultados tiene como objetivos realizar pruebas y medir varios aspectos con relación al rendimiento del modelo. Aspectos sobre cómo se realizó el ajuste de hiper parámetros, la validación del modelo en términos de exactitud de reconocimiento y clasificación como también sus correspondientes resultados son explicados con detalle en el capítulo 5.1.

4.4 FASE DE INTEGRACIÓN

La integración con el segundo componente del presente proyecto requirió del ajuste del modelo para un funcionamiento online. A diferencia del desarrollo del modelo offline, el modelo online trabaja recibiendo señales del sensor Myo Armband. El modelo online clasifica una ventana de señal y devuelve una etiqueta para que la aplicación ejecute una respectiva acción por cada periodo de tiempo. En la Figura 8 se describe como se integra el funcionamiento del modelo online y la aplicación .



Figura 8. Integración del modelo online y el segundo componente del proyecto

El modelo online mantiene la estructura del modelo offline, por lo que a través del método de ventana deslizante una ventana se moverá sobre la transmisión de datos que el sensor proporcione. Por cada periodo de tiempo definido por el paso entre ventanas consecutivas, se llama a la transmisión de datos del sensor llenando la ventana con esos datos para las primeras iteraciones. En las siguientes iteraciones, la ventana usa datos de iteraciones anteriores para mantener la ventana llena, al mismo tiempo que va añadiendo nuevas muestras de datos, tal como se ilustra en la Figura 9.

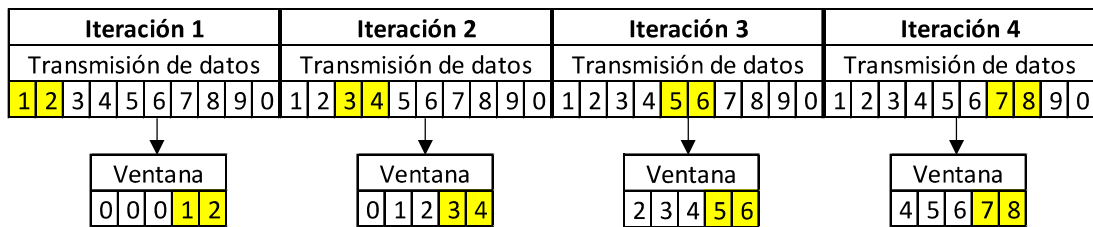


Figura 9. Ventana deslizante aplicada a retransmisión de datos en línea

Hiper parámetros para adquisición de datos

La clasificación se calcula sobre los datos de la ventana correspondientes a cada iteración. Para alcanzar una precisión en tiempo real, el tamaño del paso entre ventanas consecutivas se estableció en 0.25 segundos. Es decir, el tiempo de procesamiento de cada ventana consecutiva se debe ejecutar por debajo de ese paso. Para lograr este objetivo, tras un seguimiento de la fase de integración, se tomaron en cuenta las siguientes consideraciones:

- Recursos computacionales. Se buscó trabajar con un computador que contase con mayores recursos computacionales, es decir, que tuviera más de 2 núcleos con el fin de aprovechar tareas de procesamiento en paralelo lo cual brindaba tiempos de respuesta más cortos.
- El número de muestras que constituyen el datasetEMG. Se estableció trabajar solo con 5 repeticiones de cada gesto. Esto se debe al tiempo de procesamiento que el algoritmo DTW toma para realizar el cálculo de similitudes entre un número de señales almacenadas en el datasetEMG y una señal de entrada.

La longitud de ventana óptima para el paso (0.25 segundos) se halló experimentalmente para cada uno de los sub clasificadores. En el caso del sub clasificador EMG se estableció una longitud de ventana de 3.4 s y para el sub clasificador IMU una longitud de ventana de 5 segundos. Para ambos sub clasificadores el valor de vecinos más cercanos se determinó con la formula $k = \text{ceil}[\log_2 N]$, donde N es el número de muestras que forman el dataset correspondiente al sub clasificador. El umbral con el fin de evitar falsos positivos se estableció en 80%. Los hiper parámetros para el modelo online se presentan en la Tablas 7.

Tabla 7. Hiper parámetros para los sub clasificadores EMG e IMU

	Paso [s]	Ventana [s]	kNN	Umbral
sub clasificador EMG	0.25	3.4	7	80%
sub clasificador IMU	0.25	5	5	80%

Aplicación de uso del modelo desarrollado

La aplicación desarrollada en el segundo componente consistió en un visor de modelos 3D en realidad virtual. La aplicación permite el control de opciones, visualización e interacción con modelos 3D a través de gestos de la mano. La estructura del funcionamiento integrado de ambos componentes se ilustra en la Figura 8 y se explica de la siguiente manera: el presente componente, modelo HGR, clasifica una señal y devuelve una etiqueta correspondiente a un gesto. La etiqueta devuelta por el primer componente es la entrada al segundo componente, visor de modelos 3D, e indica la acción correspondiente a realizar por la aplicación. Una video demostración del sistema íntegro en funcionamiento puede verse en el enlace del Anexo III.

5 RESULTADOS, CONCLUSIONES Y RECOMENDACIONES

5.1 Resultados

Los resultados que se analizan en este presente subcapítulo se organizan de la siguiente manera. En primer lugar, se exponen tanto las pruebas realizadas para la estimación de las longitudes de ventana óptimas para cada sub clasificador como también los resultados obtenidos. En segundo lugar, se explican las pruebas realizadas para estimar las longitudes de ventana en la fase de Integración. Finalmente se exponen los resultados de tiempo de respuesta, exactitud en términos de reconocimiento y clasificación del modelo desarrollado en este trabajo.

5.1.1 Estimación de longitudes de ventana óptimas

Las estimaciones de la longitud óptima de la ventana deslizante tanto para el sub clasificador EMG como para el sub clasificador IMU se realizaron de manera experimental. Las pruebas se realizaron a los sub clasificadores de manera independiente, es decir, pruebas para estimar una óptima longitud de ventana para la clasificación de señales EMG y aparte, pruebas para estimar otra óptima longitud de ventana para la clasificación de señales IMU.

Las pruebas consistieron en ejecutar el sub clasificador con distintos tamaños de ventana a 5 usuarios elegidos aleatoriamente dentro del dataset de entrenamiento. Se probaron con longitudes de ventana desde 1 hasta 5 segundos (con un paso de 1 segundo). Para ambos sub clasificadores, el valor de k vecinos más cercanos fue establecido en 7 y el umbral de clasificación en 80%. Estas pruebas fueron evaluadas comparando las tasas de precisión de clasificación promedio obtenidas en cada prueba de acuerdo con cada longitud de ventana probada. Para una mejor apreciación de resultados, estos se ilustran en figuras que acotan longitudes de ventana en donde se aprecia con más rigor el valor máximo obtenido de las pruebas realizadas en cada caso.

Para las muestras tomadas por el sensor Myo Armband, en la Figura 10a se puede observar que para el sub clasificador EMG la precisión promedio de clasificación llega a un valor máximo con una longitud de ventana de 3.4 segundos. La Figura 10b ilustra que para el sub clasificador IMU la precisión promedio de clasificación llega a un valor máximo con una longitud de ventana de 4.1 segundos.

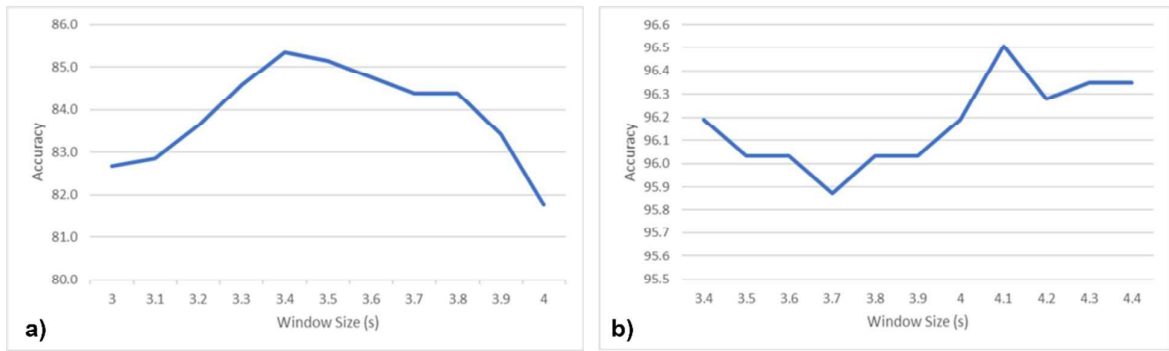


Figura 10. Resultados de pruebas de (a) ventana EMG y (b) ventana IMU

Para las muestras tomadas por el sensor G-Force Pro, en la Figura 11a se puede observar que para el sub clasificador EMG la precisión promedio de clasificación llega a un valor máximo con una longitud de ventana de 3.4 segundos. La Figura 11b muestra que para el sub clasificador IMU la precisión promedio de clasificación llega a un valor máximo con una longitud de ventana de 4.9 segundos.

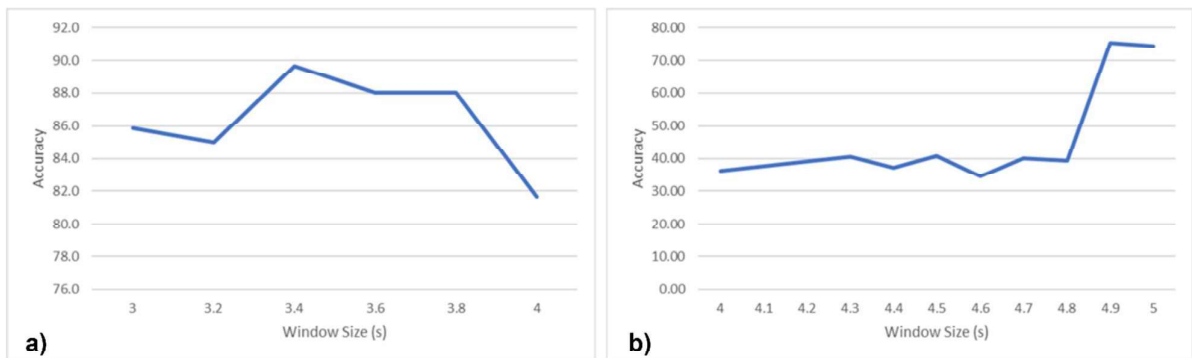


Figura 11. Resultados de pruebas de (a) ventana EMG y (b) ventana IMU

5.1.2 Estimación de longitudes de ventana óptimas fase integración

Las pruebas que se hicieron para determinar las longitudes de ventana óptimas EMG e IMU para la fase de integración fueron similares a las que han sido descritas en la subsección anterior. Se conservaron los mismos parámetros de clasificación para cada sub clasificador, pero se usó un distinto valor de paso entre ventanas consecutivas: 0.25 segundos. Este cambio se debe a que la respuesta en tiempo real del modelo para la fase de Integración implica un valor de paso menor a 0.3 segundos. Para realizar estas pruebas solo se trabajó con muestras obtenidas con el sensor Myo Armband.

En la Figura 12a se puede observar que para el sub clasificador EMG la precisión promedio de clasificación llega a un valor máximo con una longitud de ventana de 3.4 segundos. Por otro lado, en la Figura 12b se puede observar que para el sub clasificador IMU la precisión promedio de clasificación se mantiene entre valores de 4.6 y 5 segundos.

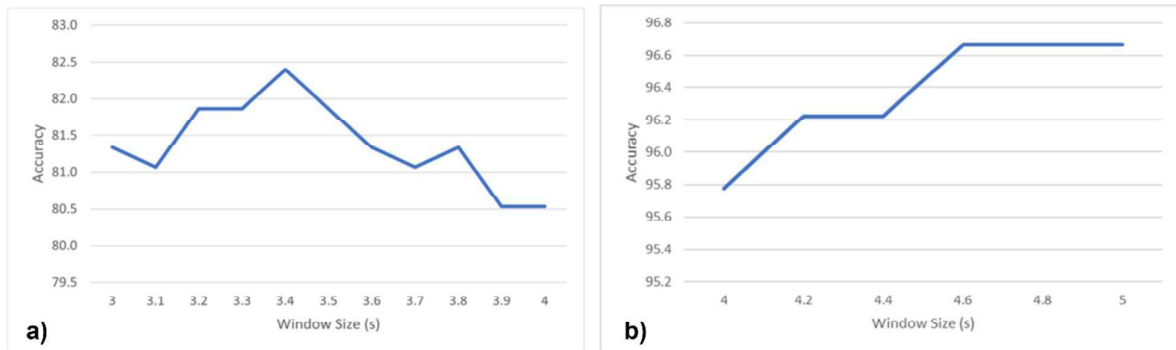


Figura 12. Resultados de pruebas de (a) ventana EMG y (b) ventana IMU

5.1.3 Resultados de evaluación del modelo

Con respecto a la evaluación del modelo en términos de tiempo de respuesta, clasificación y reconocimiento, el modelo fue evaluado con 42 usuarios del dataset de prueba a través de un proceso externo. Observando la matriz de confusión en la Tabla se puede observar los resultados de clasificación a detalle. El gesto que presenta mayor sensibilidad es Up (93.2%) mientras que el gesto con menor sensibilidad es WaveIn (67.3%).

Por otro lado, los gestos con mayor y menor precisión respectivamente son WaveIn (97.1%) y Down (77.2%). Se ilustra también que la exactitud de clasificación general del modelo es del 80% mientras que la precisión media es del 84.45%. Es importante entender que los resultados en términos de clasificación llegan al 80% de precisión tomando en cuenta que el elevado número de gestos a clasificar (11 gestos) lo cual resulta bastante satisfactorio.

Tabla 8. Matriz de confusión del modelo

		Confusion Matrix												
Output Class	forward	516 6.8%	7 0.1%	21 0.3%	9 0.1%	11 0.1%	12 0.2%	1 0.0%	13 0.2%	21 0.3%	7 0.1%	19 0.3%	20 0.3%	78.5% 21.5%
	fist	4 0.1%	463 6.1%	5 0.1%	2 0.0%	0 0.0%	1 0.0%	0 0.0%	11 0.1%	5 0.1%	1 0.0%	7 0.1%	0 0.0%	92.8% 7.2%
	waveIn	1 0.0%	5 0.1%	424 5.6%	0 0.0%	9 0.1%	6 0.1%	0 0.0%	9 0.1%	0 0.0%	0 0.0%	4 0.1%	0 0.0%	92.6% 7.4%
	right	1 0.0%	1 0.0%	8 0.1%	510 6.7%	12 0.2%	38 0.5%	1 0.0%	3 0.0%	3 0.0%	6 0.1%	1 0.0%	1 0.0%	87.2% 12.8%
	waveOut	0 0.0%	3 0.0%	4 0.1%	0 0.0%	432 5.7%	0 0.0%	0 0.0%	5 0.1%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	97.1% 2.9%
	pinch	15 0.2%	18 0.2%	11 0.1%	21 0.3%	8 0.1%	461 6.1%	2 0.0%	7 0.1%	8 0.1%	1 0.0%	3 0.0%	13 0.2%	81.2% 18.8%
	relax	52 0.7%	111 1.5%	104 1.4%	24 0.3%	119 1.6%	77 1.0%	626 8.3%	91 1.2%	42 0.6%	19 0.3%	50 0.7%	15 0.2%	47.1% 52.9%
	open	4 0.1%	1 0.0%	2 0.0%	1 0.0%	5 0.1%	5 0.1%	0 0.0%	447 5.9%	0 0.0%	0 0.0%	0 0.0%	3 0.0%	95.5% 4.5%
	backward	16 0.2%	0 0.0%	29 0.4%	7 0.1%	0 0.0%	1 0.0%	0 0.0%	14 0.2%	525 6.9%	0 0.0%	16 0.2%	20 0.3%	83.6% 16.4%
	up	7 0.1%	0 0.0%	0 0.0%	17 0.2%	28 0.4%	4 0.1%	0 0.0%	3 0.0%	2 0.0%	587 7.8%	6 0.1%	2 0.0%	89.5% 10.5%
	down	8 0.1%	20 0.3%	8 0.1%	39 0.5%	6 0.1%	24 0.3%	0 0.0%	3 0.0%	19 0.3%	9 0.1%	522 6.9%	18 0.2%	77.2% 22.8%
	left	6 0.1%	1 0.0%	14 0.2%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	24 0.3%	5 0.1%	0 0.0%	1 0.0%	538 7.1%	91.2% 8.8%
			81.9% 18.1%	73.5% 26.5%	67.3% 32.7%	81.0% 19.0%	68.6% 31.4%	73.2% 26.8%	99.4% 0.6%	71.0% 29.0%	83.3% 16.7%	93.2% 6.8%	82.9% 17.1%	85.4% 14.6%
		forward	fist	waveIn	right	waveOut	pinch	relax	open	backward	up	down	left	
		Target Class												

La exactitud de reconocimiento del modelo fue de 66.12% con una desviación estándar de $\pm 18.30\%$ lo cual puede apreciarse de forma gráfica en la Figura 13. Este bajo resultado de la evaluación del modelo encuentra su explicación al tamaño del paso entre ventanas consecutivas seleccionado. El modelo fue desarrollado con un paso de un 1 segundo lo cual quiere decir que la señal iba a ser barrida por 5 iteraciones dando menos oportunidad de estimar con mayor exactitud en qué momento el gesto es ejecutado. El resultado de exactitud de reconocimiento evidencia la importancia de contar con unos recursos computacionales para realizar pruebas en menos tiempo.

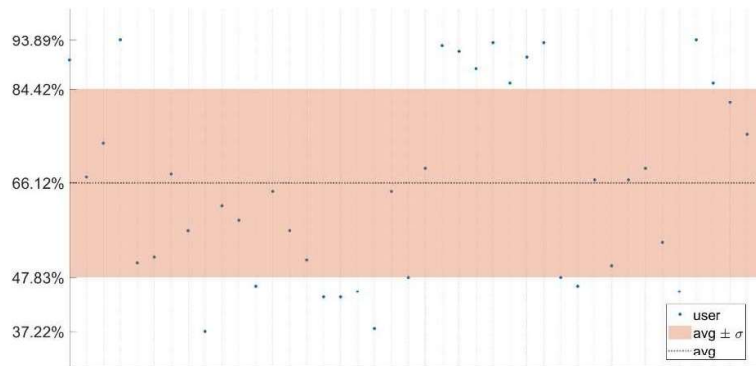


Figura 13. Exactitud de reconocimiento.

El tiempo de respuesta es el tiempo que cada sub clasificador (EMG o IMU) tarda en procesar y clasificar cada una de las ventanas de la señal a clasificar. El procesamiento de cada ventana fue cronometrado y almacenado para un posterior análisis. En la Figura 14 se muestran los tiempos de procesamiento para cada ventana procesada. Se puede observar que el mayor tiempo de procesamiento es de 0.22 segundos mientras que la mayoría de los tiempos de procesamiento de ventana se mantienen por debajo de 0.16 segundos. Con esto se puede afirmar que el modelo devuelve una respuesta en tiempo real, es decir por debajo de 0.3 segundos.

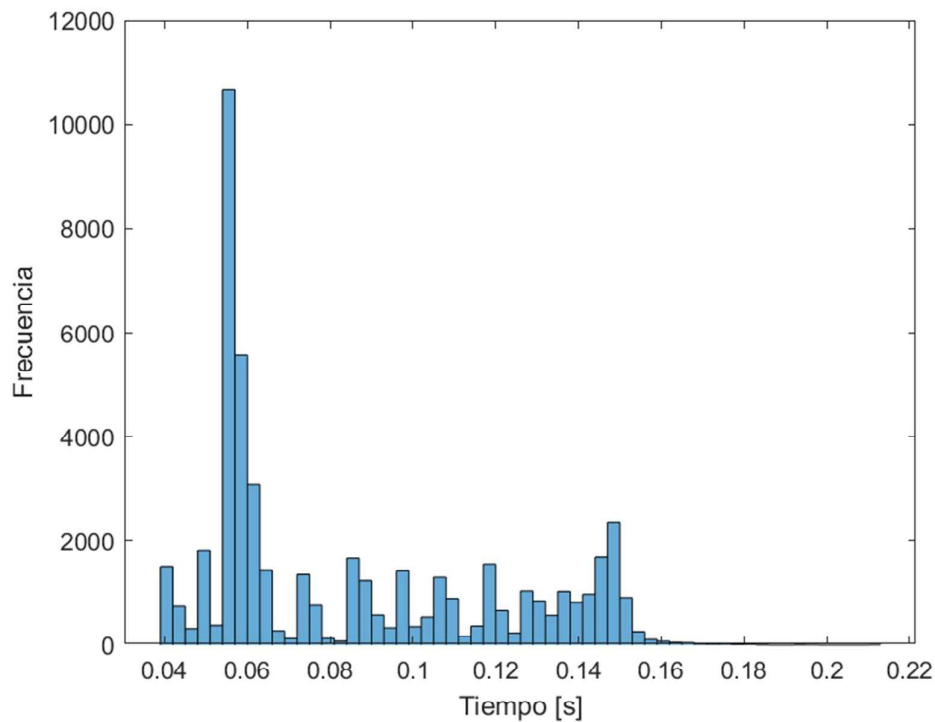


Figura 14. Histograma de tiempos de procesamiento por cada ventana

5.2 Conclusiones

Se ha desarrollado un modelo de reconocimiento de gestos del mano basado en señales EMG e IMU obtenidas con sensores de tipo brazaletes. La estructura del modelo HGR desarrollado se basa en las etapas de adquisición de datos, preprocesamiento, extracción de características y clasificación. La etapa de clasificación se dividió en dos sub clasificadores de acuerdo con el tipo de gesto que se va a clasificar (estático o dinámico). Para ambos sub clasificadores se usó el mismo valor de k vecinos más cercano junto con el algoritmo DTW, y umbral de clasificación, sin embargo, se usaron distintas longitudes de ventana deslizante para cada respectivo sub clasificador. Los recursos computacionales tuvieron un papel muy importante en el tiempo de la realización de pruebas. Trabajar con un computador con mejores prestaciones para la fase de integración permitió ajustar el delay del tiempo de respuesta.

El modelo fue evaluado con los datos del conjunto de prueba y los resultados muestran que el modelo presenta una exactitud de clasificación del $80.04\% \pm 13.66\%$ y una exactitud de reconocimiento del $66.12\% \pm 18.30\%$. El tiempo de respuesta del modelo está por debajo de 0.22 segundos permitiendo al modelo funcionar en tiempo real. Finalmente, el modelo fue exitosamente integrado a una aplicación de interfaz humano-máquina.

5.3 Recomendaciones

A continuación, se plantean las siguientes recomendaciones para el trabajo futuro del presente proyecto:

- Realizar una revisión de los datos adquiridos antes de formar el conjunto de datos definitivo. En este sentido, los datos con defectos que no hayan sido revisados provocarían problemas en el desarrollo del modelo.
- Disminuir la duración de la toma de las muestras de los gestos, es decir, tomar muestras durante una duración menor a 5 segundos, como también disminuir el número de repeticiones por muestras. Esta mejora elevaría en gran medida el rendimiento de la tarea de clasificación.
- Usar un equipo de cómputo que cuente con recursos computacionales enfocados para tareas de ML destacando el recurso del número de núcleos mayor a 2. Este número de núcleos resulta crucial para el desempeño óptimo de tareas que requieran de procesamiento en paralelo.

- Se recomienda probar con un paso entre ventanas consecutivas menor a 1 segundo, cuando los recursos computacionales sean los suficientes. Un paso menor brindaría otro enfoque que bien podría brindar mejores resultados de reconocimiento.
- Probar a clasificar ambas señales EMG e IMU a la vez, es decir, eliminar el switch de la estructura del modelo y por lo tanto usar técnicas de extracción de características alternativas .

6 REFERENCIAS BIBLIOGRÁFICAS

- [1] Guo, G., Wang, H., Bell, D., Bi, Y., Greer, K. (2003). KNN Model-Based Approach in Classification. In: Meersman, R., Tari, Z., Schmidt, D.C. (eds) On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE. OTM 2003. Lecture Notes in Computer Science, vol 2888. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-39964-3_62
- [2] A. Jaramillo-Yáñez, M. E. Benalcázar, and E. Mena-Maldonado, "Real-Time Hand Gesture Recognition Using Surface Electromyography and Machine Learning: A Systematic Literature Review," *Sensors*, vol. 20, no. 9, p. 2467, Apr. 2020, doi: 10.3390/s20092467.
- [3] Praveena, M., & Jaiganesh, V. A Literature Review on Supervised Machine Learning Algorithms and Boosting Process. *International Journal of Computer Applications*, 169(8), 32–35. <https://doi.org/10.5120/IJCA2017914816>
- [4] Kilby, J., Prasad, K., & Mawston, G. (2016). Multi-channel surface electromyography electrodes: a review. *IEEE Sensors Journal*, 16(14), 5510-5519.
- [5] Faisal, Ilham & Purboyo, Tito & Ansori, Anton. (2019). A Review of Accelerometer Sensor and Gyroscope Sensor in IMU Sensors on Motion Capture. *Journal of Engineering and Applied Sciences*. 15. 826-829. 10.36478/jeasci.2020.826.829.
- [6] Taunk, Kashvi & De, Sanjukta & Verma, Srishti & Swetapadma, Aleena. (2019). A Brief Review of Nearest Neighbor Algorithm for Learning and Classification. 1255-1260. 10.1109/ICCS45141.2019.9065747.
- [7] Senin, Pavel. (2009). Dynamic Time Warping Algorithm Review.
- [8] Cai, Xingyu, Xu, Tingyang, Yi, Jinfeng, Huang, Junzhou, & Rajasekaran, Sanguthevar.. DTWNet: a Dynamic Time Warping Network. *Advances in neural information processing systems*, 32. Retrieved from <https://par.nsf.gov/biblio/10181828>.
- [9] M. E. Benalcázar, Á. L. Valdivieso Caraguay, and L. I. Barona López, "A User-Specific Hand Gesture Recognition Model Based on Feed-Forward Neural Networks, EMGs, and Correction of Sensor Orientation," *Applied Sciences*, vol. 10, no. 23, p. 8604, Dec. 2020, doi: 10.3390/app10238604.

- [10] E. A. Chung and M. E. Benalcázar, "Real-Time Hand Gesture Recognition Model Using Deep Learning Techniques and EMG Signals," 2019 27th European Signal Processing Conference (EUSIPCO), 2019, pp. 1-5, doi: 10.23919/EUSIPCO.2019.8903136.
- [11] Qi, J., Jiang, G., Li, G., Sun, Y., & Tao, B. (2019). Surface EMG hand gesture recognition system based on PCA and GRNN. *Neural Computing and Applications*, 32, 6343-6351.
- [12] Zhang, & Shih, Yang-Hsin & Qian, Jinwu. (2019). Real-Time Surface EMG Pattern Recognition for Hand Gestures Based on an Artificial Neural Network. *Sensors*. 19. 3170. 10.3390/s19143170.
- [13] Tavakoli, Mahmoud & Benussi, Carlo & Lopes, Pedro & Osorio, Luis & de Almeida, Aníbal. (2018). Robust hand gesture recognition with a double channel surface EMG wearable armband and SVM classifier. *Biomedical Signal Processing and Control*. 46. 121-130. 10.1016/j.bspc.2018.07.010.
- [14] M. E. Benalcázar, A. G. Jaramillo, Jonathan, A. Zea, A. Páez and V. H. Andaluz, "Hand gesture recognition using machine learning and the Myo armband," 2017 25th European Signal Processing Conference (EUSIPCO), 2017, pp. 1040-1044, doi: 10.23919/EUSIPCO.2017.8081366.
- [15] L. I. B. Barona López et al., "An Energy-Based Method for Orientation Correction of EMG Bracelet Sensors in Hand Gesture Recognition Systems," *Sensors*, vol. 20, no. 21, p. 6327, Nov. 2020, doi: 10.3390/s20216327.

7 ANEXOS

ANEXO I. Conjunto de datos

El conjunto de datos usado para el desarrollo del presente trabajo se encuentra disponible a través del siguiente enlace:

<https://laboratorio-ia.epn.edu.ec/en/resources/dataset/emg-imu-epn-100>

ANEXO II. Implementación del modelo

El siguiente enlace contiene el código desarrollado durante la implementación del presente trabajo:

https://epnecuador-my.sharepoint.com/:f:/g/personal/laboratorio_ia_epn_edu_ec/EuAu9Z0zkp1Fqh35KuLzqg_wBHfQ2hcT6xXle8s22sB_Rcg?e=MvVokg

ANEXO III. Video demostración

El siguiente enlace muestra el funcionamiento de la integración del primer y segundo componente:

<https://youtu.be/NQcEOSXeYqc>