

The HouseBender: Sistema de reconocimiento de gestos para el control de dispositivos domóticos

Daniel Rivera Arroyave, Juan Jaramillo Tobón, Raúl Arcila Rodríguez and Danny Múnera Ramírez

Facultad de Ingeniería, Universidad de Antioquia, Medellín, Colombia

Email: {daniel.riveraa} {juan.jaramillo62} {randres.arcila} {danny.munera}@udea.edu.co

Resumen—La domotización de componentes en el hogar es algo cada vez más común por los avances tecnológicos y su fácil implementación. En este artículo se presenta el desarrollo de un control domótico basado en un clasificador de gestos haciendo uso de tecnologías de inteligencia artificial. Nuestro sistema propone la captura de los gestos por medio de un acelerómetro y giroscopio para su posterior clasificación mediante una CNN, y mediante el uso de los servicios MQTT se controlarán los dispositivos domóticos y/o componentes IoT.

Index Terms—Domótica, Redes neuronales artificiales, Internet de las Cosas (IoT), Machine learning.

I. INTRODUCCIÓN

El reconocimiento de gestos se refiere al proceso de entendimiento y clasificación de movimientos significativos realizados por las manos, brazos, cara y alguna vez por la cabeza de las personas [1]. El reconocimiento de gestos se está volviendo cada vez más popular como una forma de entrada para la electrónica de consumo como parte de la interacción entre humanos y computadoras *HCI* por sus siglas en inglés de Human Computer Interaction, un ejemplo de esto son los sistemas de realidad aumentada, los smartwatches y hasta los controles de video consolas [2]. En los últimos años, con el rápido desarrollo de la tecnología de sensores, surge un nuevo tipo de tecnología de interacción que permite reconocer el movimiento de los usuarios. Este reconocimiento se realiza usualmente a través de un acelerómetro, el cual puede medir la cantidad de aceleración de un dispositivo en movimiento. Por lo tanto, se puede lograr un reconocimiento de gestos al analizar las señales de aceleración entregadas por el dispositivo [3].

El reconocimiento de gestos puede ser una interfaz de interacción muy útil que puede ser usada en diferentes contextos, por ejemplo uno de los más adecuados puede ser la domótica. Los sistemas domóticos son tecnologías aplicadas al control y la automatización inteligente de la vivienda, es de gran utilidad en el hogar, permite la gestión eficiente del uso de la energía, brindan mayor seguridad y confort al detectar fallas en el hogar, como por ejemplo fugas de gas o de agua, además de automatizar procesos o tareas en el hogar permite la comunicación entre el usuario y el sistema. Estos sistemas pueden llegar a ser muy complejos de utilizar para algunos usuarios. Hoy en día muchas empresas proponen soluciones para el control de sistemas domóticos, estas son las más populares en reconocimiento de voz: Google Home, Amazon Echo, Apple HomePad, entre otras; pero las soluciones pueden llegar a ser costosas y complicadas de utilizar.

En este artículo se propone un sistema de reconocimiento de gestos para el control de dispositivos domóticos llamado “The HouseBender”. Este sistema ofrece una solución de bajo costo y alta precisión, el cual está basado en el uso de Redes Neuronales Convoluciones. The HouseBender implementa un acelerómetro y un giroscopio para monitorear los gestos del usuario, además posee un módulo ESP-015 que se comunica con el servidor vía wifi en tiempo real, mandando así los valores censados para su debido procesamiento, para así hacer uso de los sistemas domóticos de la casa.

En el resto del artículo presentaremos los trabajos relacionados en la Sección II. La Sección III presenta la adquisición de las señales de movimiento y el acelerómetro. La Sección IV expone la construcción de la base de datos, el análisis de la calidad de señal y el procesamiento que se les realizó a las señales. En el apartado V tenemos el algoritmo de aprendizaje de máquina (CNN). En la Sección VI se presenta el sistema de forma general y su integración con la domótica además de las evidencias de funcionamiento del sistema mientras que las conclusiones sobre el proyecto en general se encuentran en la Sección VII.

II. TRABAJOS RELACIONADOS

El *machine learning* como ciencia de datos de datos no es algo nuevo y sus fundamentos teóricos se remontan a los años 50 pero no es sino hasta los últimos años donde se ha experimentado un boom tecnológico a nivel de hardware que ha posibilitado buscar soluciones a nuevos problemas e ingeniar nuevas soluciones a los problemas ya existentes mediante técnicas de machine learning y deep learning en diversas áreas dentro de las cuales se encuentra el reconocimiento de gestos realizados por un humanos. A continuación se plantean algunos trabajos relevantes donde han aplicado técnicas de machine learning como complemento para el uso de dispositivos del hogar.

Reconocedor de Movimientos Low-Cost mediante Redes Neuronales [4]

En este trabajo se hace uso del control remoto WiiMote para el control de dispositivos plug and play, haciendo uso de los acelerómetros con los que viene equipado el control remoto, basan su reconocimiento en red neuronal (red de kohonen), plug and play “ofrece las siguientes características: 1) provee un descubrimiento automático de cualquier dispositivo conectado a la red; 2) permite una configuración automática de la red sin intervención del usuario; 3) es independiente tanto

de los dispositivos como de las redes; 4) utiliza protocolos estándares; y 5) ofrece un sistema sencillo de extensibilidad”, los datos son captados por el wiimote, enviados por bluetooth a un módulo encargado de realizarles un tratamiento de datos; luego esta información es enviada a la red neuronal entrenada para clasificar el gesto y dependiente de esta realiza la acción sobre el dispositivo plug and play; los resultados fueron aceptables para 6 gestos, entrenados con 300 muestras y mejoraron al hacer una detección de outliers mediante la mediana y la varianza.

Sistema de Reconocimiento de Gestos de la Mano basado en Acelerómetro para Interacción en TV Digital [1]

Como su título lo indica se plantea exclusivamente para control de televisión, al ser un problema de clasificación de patrones, se hace uso de multilayer perceptron y máquina de soporte de vectorial para 8 tipos de gestos diferentes y plantean que la mayor parte de la literatura sólo utiliza acelerómetros de 3 ejes y que los gestos son almacenados en vectores que contienen las variaciones de los niveles de aceleración en función del tiempo en el espacio tridimensional. Las dificultades que documentan son que los gestos: “En primer lugar, presentan ambigüedad en la segmentación, es decir, no se conocen los límites de la realización del gesto. En segundo lugar, presentan variabilidad espacio-temporal ya que el gesto varía dinámicamente en forma y duración, incluso para los mismos gestos y mismas personas, hacen uso de una base de datos anteriormente habilitada para este tipo de proyectos. El gesto es activado mediante un botón y termina cuando se suelta y además plantean una ecuación para convertir la frecuencia de muestreo obtenida a una frecuencia de muestreo deseada.

Sistema de reconocimiento de Gestos Utilizando Smartwatches [5]

En este trabajo hacen uso de *Smartwatches*, los cuales se han vuelto muy populares durante los últimos años y es más conveniente para los usuarios realizar una integración con el sistema utilizando, el reloj utilizado es un Samsung Gear S2 encargado de publicar los datos captados por el acelerómetro interno de 3 ejes mediante MQTT en un formato JSON. Es necesario resaltar que hicieron un estudio donde encontraron que los gestos normales de una persona difícilmente superan los 9 metros/seg cuadrado, así que cuando se supera este umbral es cuando se comienza a capturar un gesto a una frecuencia de 26 hz durante un segundo (26 muestras por gesto) y es publicado, para finalizar comentan que se obtuvieron mejores resultados utilizando 7 gestos normalizando los datos.

En esta revisión bibliográfica se encontraron diversos paradigmas para realizar el reconocimiento de gestos mediante acelerómetros, algunos reflejan modelos de *machine learning* tradicionales y otros el uso de técnicas de *deep learning* que son las que finalmente se usaron en el sistema propuesto. Además en todos variaron el dispositivo de uso siendo [5] con su *Smartwatch* el más similar a nuestro proyecto, pero ampliamente diferenciado debido a que nuestro sistema utiliza

un dispositivo diseñado específicamente para la captura de la señal de movimiento y ninguna otra función.

III. ADQUISICIÓN SEÑAL DE MOVIMIENTO-ACELERÓMETRO

Para este proyecto usaremos un módulo IMU [6] o unidad de medición inercial el cual usa sistemas microelectromecánicos los cuales por medio de lecturas de capacitancia pueden determinar bajo qué fuerzas está el módulo y así entregarnos con alta precisión información sobre el acelerómetro y el giroscopio con 6 grados de libertad (DoF). Un sistema micro electromecánico (SMEM) [7] se refiere a la tecnología electromecánica de dispositivos microscópicos, sobre todo los que tienen partes móviles y que realizan diferentes funciones como percepción, procesamiento de datos, comunicación y actuación sobre el entorno.

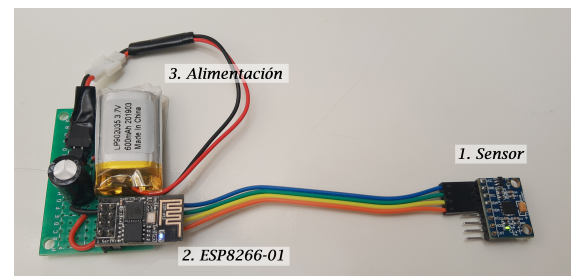


Figura 1: Dispositivo desarrollado para la recolección de datos

Como elemento esencial para funcionamiento de este sistema fue necesario desarrollar un dispositivo portable y adaptable que pueda ser usado en la mano por el usuario y que es encargado de la medición y transmisión de los datos de interés en la realización de los gestos y a su vez pueda funcionar de manera autónoma similar a como se realizó en [5] el SmartWatch. El dispositivo en cuestión es el que se aprecia en la figura 1 y sus elementos fundamentales que se encuentran señalados son descritos a continuación

1. **Sensor acelerómetro y giroscopio:** El sensor usado es una MPU6050, un sensor ampliamente reconocido en múltiples proyectos como aeromodelismo e IoT, dado a que tiene una buena sensibilidad, buenos tiempos de respuesta y es altamente configurable, este sensor es capaz de medir la aceleración en los 3 ejes cartesianos, además de eso es capaz de medir también la velocidad angular en cada uno de estos mismos ejes.
2. **Módulo ESP:** El módulo ESP-015 es un módulo wifi equipado con un procesador ESP el cual posee varios pines multipropósito, además de que tiene una interfaz I2C la cual es una característica fundamental para el uso de nuestro sensor, este módulo tiene la capacidad de mandar los valores censados en tiempo real por wifi a un servidor.
3. **Alimentación:** La alimentación del dispositivo viene dada por una batería lipo de 400 mAh la cual se conecta a una circuito que rectifica los 3.7 voltios de la batería a 3.3 voltios que usa el modulo ESP, esto lo hace

un regulador lineal de voltaje (MCP1700) y usa 2 capacitores como protección.

Todos estos elementos en conjunto permiten la adquisición de la señal de movimiento de manera inalámbrica en bajo tiempo de ejecución, en la figura 2 se presenta la señal correspondiente a la realización y captura de un gesto durante un tiempo de 5 segundos.

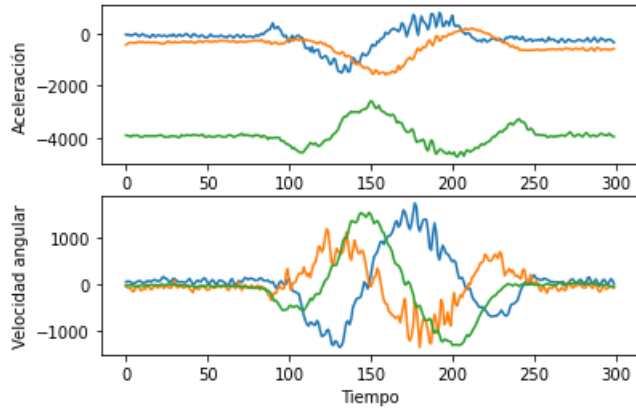


Figura 2: Señal de gesto X en ventana de 5 seg a 60HZ, arriba aceleración en 3 ejes y abajo velocidad angular.

IV. CREACIÓN DE LA BASE DE DATOS

Componentes de la base de datos

Para el Proyecto se recolectó un dataset de 1600 muestras donde cada una posee 300 filas y 8 columnas. Las muestras representan los datos enviados por el dispositivo durante 5 segundos que fueron recopiladas en archivos CSV, la estructura de las columnas en los archivos es la siguiente:

Tiempo: esta columna se usa para determinar si los mensajes se están enviando en los tiempos adecuados para garantizar que la frecuencia de muestreo sea siempre de 60Hz

Contador: esta columna se usa para determinar si los mensajes llegan de forma ordenada y no se perdió ningún dato.

Acx, Acy y Acz: Corresponde la aceleración en los 3 ejes cartesianos X, Y y Z medida por el acelerómetro.

Gx, Gy y Gz: Corresponde a la velocidad angular medida por el giroscopio.

Tipos de movimiento y protocolo de adquisición

Las muestras fueron tomadas para 4 tipos gestos realizadas por dos personas diferentes siguiendo un protocolo donde el dispositivo estaba encendido enviando datos en transmisión continua mientras la persona iniciaba el software implementado para la creación de la base de datos, una vez iniciado el software carga un contador de 5 segundos para que extienda la mano al frente y luego comience a hacer el movimiento en otros 5 segundos terminando con la mano extendida en la misma posición inicial un ejemplo de los movimientos se puede ver en la figura 3.

Mientras que los tipos de gesto diseñados junto con su descripción a grandes rasgos son:

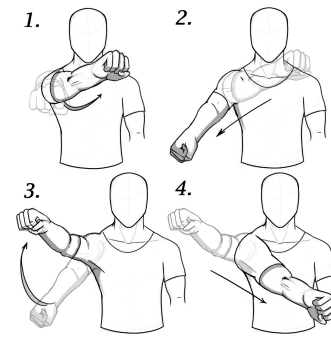


Figura 3: Secuencia de movimientos para el gesto X.

1. **Activación:** Parte del brazo estirado y consiste en un golpe recto al pecho y luego volver a la posición inicial.
2. **Encender:** Con el brazo estirado, realizar un círculo hacia la derecha.
3. **Alimentación:** Se realiza un círculo con la muñeca y termina con un puño hacia el frente.
4. **X:** Se dibuja un lazo o un símbolo de infinito con el brazo estirado.

Análisis de calidad de la señal

Dentro de los archivos CSV, además de guardar las medidas del sensor, se incluyeron dos columnas correspondientes al tiempo y a un contador que son utilizados para evaluar la calidad de cada muestra analizada desde la transmisión además de probar la viabilidad del protocolo para la implementación final del sistema. Todo esto debido a que lo medido es enviado a través de wifi y protocolos de transferencia donde existe una posibilidad de que los paquetes de datos enviados por el dispositivo se pierdan o lleguen en desorden. El tiempo es usado para garantizar que la frecuencia a la que llegan los datos al servidor corresponda o esté en un rango de tiempo similar a la frecuencia en que el dispositivo toma los datos del acelerómetro y envía durante el transcurso de toda la señal midiendo el tiempo transcurrido entre paquete y paquete, mientras que el contador es usado para verificar que los paquetes de información llegaron en todos en orden. Este análisis permitió descartar las muestras “no confiables” por el retraso presente entre sus datos y repetir su ejecución, es pertinente aclarar que nunca se presentó una situación en la que los datos llegasen en desorden.

Preprocesamiento de señal

Como se puede ver en la figura 2 Las señales recolectadas poseen toda la información del gesto, pero además se recopila información extra que parece provenir de vibraciones producidas por el pulso de la persona o un mal montaje del dispositivo de recolección, esta información no es crítica para el proceso de clasificación de los gestos.

Haciendo un análisis superficial, se puede ver que el ruido tiene la apariencia de una señal de baja frecuencia compuesta por diversas señales de alta frecuencia sumadas a ella que no son de interés para nuestro sistema, por lo que decidimos

en optar por pasar la señal por un filtro pasa bajas. Dentro de la teoría del procesamiento de señales se encuentra la Transformada de Fourier [8] que es una transformación matemática que toma una serie de tiempo y la descompone en una representación de las frecuencias que la componen como se puede ver en la figura 4, esta transformación también puede ser revertida haciendo uso de su transformada inversa que devuelve de la representación hacia la señal inicial sin perder información, entonces se procedió a hallar la transformada de Fourier de las señales entrantes se buscó la forma de eliminar las altas frecuencias que allí se presentaban.

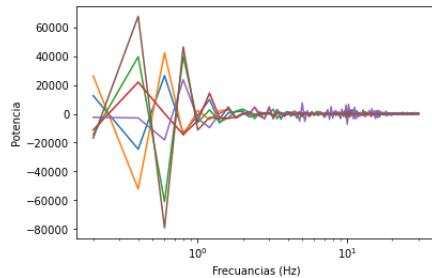


Figura 4: Ejemplo transformada de Fourier en señal captada en escala logarítmica.

La figura 4 es la representación en frecuencia de la señal anteriormente mostrada en la figura 2, se determinó que la mejor forma de eliminar frecuencias específicas que se presentaran en ella es multiplicando la representación en frecuencias por una señal sigmoideal que no interfiere con las bajas frecuencias mientras que las altas son reducidas a cero.

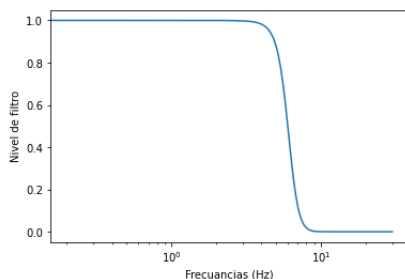


Figura 5: Nivel de filtro vs frecuencia en señal sigmoideal en escala logarítmica.

Esta señal sigmoideal es fácilmente manipulable para cambiar su punto de inflexión y de esta manera graduar la alteración sobre las frecuencias que deseamos reducir como se ve en la figura 6 donde los picos de mayor frecuencia dejan de ser apreciables.

Después de realizar la multiplicación de la transformada de Fourier por la señal sigmoideal y posteriormente aplicar la transformada inversa se obtiene la señal que aparece en la figura 7.

V. ALGORITMO DE APRENDIZAJE DE MÁQUINA

Se inició haciendo experimentos con *redes neuronales artificiales* sencillas las cuales no tuvieron buenos resultados

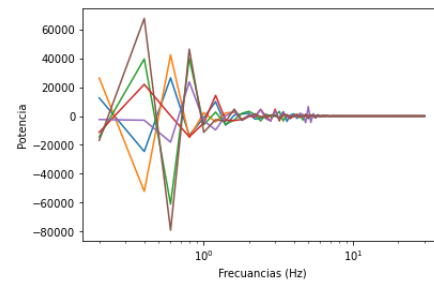


Figura 6: Análisis en frecuencia alterado por sigmoide.

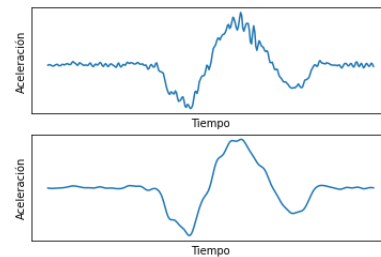


Figura 7: Señal de aceleración en eje antes y después de ser transformada.

respecto a *Accuracy*, dado que sin importar las diferentes configuraciones usadas siempre se tenía el mismo resultado, el modelo llegaba a la solución de determinar que todos los gestos pertenecían a la misma clase. Dado que a simple vista se puede ver una diferencia clara entre los distintos tipos de gestos en sus señales, optamos por usar una red neuronal convolucional [9]. Una variación de la red neuronal convencional que consta en la aplicación de una serie de filtros que se le aplican sobre toda la señal y posteriormente una red neuronal procesa la activación de los filtros para hacer la clasificación de las muestras. Durante la fase de entrenamiento se actualizan los pesos de la red neuronal y además se modifican los valores de los filtros para que puedan identificar las características más importantes que se presentan en cada una de las señales. Esta red neuronal permite identificar la realización de un gesto dentro de una muestra con una duración mucho más larga sin importar tampoco la duración del gesto dentro de la muestra.

Para encontrar los mejores hiperparámetros se dividió el conjunto de muestras en una proporción de 80 % para entrenamiento y 20 % para validación, y se crearon distintos modelos en los cuales se fue variando sus hiperparámetros dentro de los que destacan:

1. **Tamaño de los filtros:** Este hiperparámetro determina la longitud de los filtros que se le aplicaran a las muestras en la capa convolucional de la red neuronal antes de entrar a la clasificación
2. **Número de filtros:** El número de filtros determina cuántas características diferentes buscaremos para clasificar las muestras.
3. **Batch size:** se refiere a la cantidad de ejemplos o muestras de entrenamiento utilizados en cada iteración del entrenamiento.

Esto concluye en 24 modelos diferentes los cuales tuvieron resultados variados, pero solo 16 de estos modelos convergieron en un *accuracy* deseable, pero cada modelo internamente seguía presentando problemas para la clasificación de ciertos gestos, así que para tener un modelo que presentara mejores métricas de desempeño se decidió implementar un *Voting classifier* [10] con los mejores 6 que al llegar a un consenso entre ellos mejoran así el *accuracy*, la arquitectura de estos modelos puede verse en la figura 8.

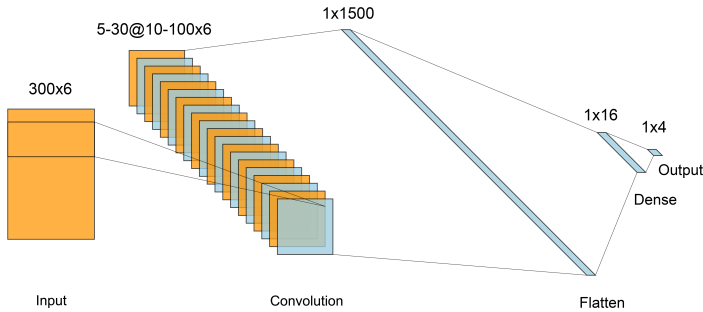


Figura 8: Arquitectura red neuronal convolucional

Para tomar la decisión final sobre las métricas de desempeño del modelo resultante se tenía un conjunto de datos extra con 100 muestras las cuales corresponden a gestos realizados por una tercera persona que no participó en la recolección del primer conjunto de datos, utilizando este conjunto se obtuvieron muy buenos resultados con una clasificación casi perfecta salvo por el gesto de *Activación III* con la etiqueta 0, la matriz de confusión resultante del ensamble de modelos se presenta en la figura 9.

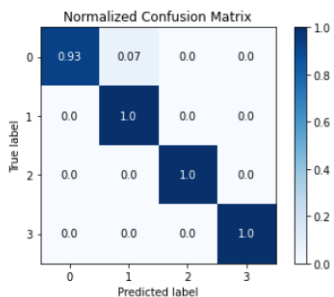


Figura 9: Matriz de confusión del modelo final.

Luego de tener un modelo aceptable procedimos a escoger la forma de implementar el modelo y lograr tener una predicción continua sobre los datos transmitidos por el dispositivo, dado que el objetivo propuesto era que el dispositivo fuera lo más simple y que para realizar una predicción no fuera necesario nada más que los movimientos del usuario sin interacción con botones de activación o similar. Partiendo del funcionamiento de una CNN optamos por un funcionamiento similar en nuestra implementación usando una ventana móvil la cual tiene un tamaño determinado y se mueve a través

de un número n de paquetes de datos, esta implementación fue probada realizando un nuevo conjunto de muestras con duración de 20 segundos donde se encontraban varios gestos separados por unos segundos de quietud similar a como sería el comportamiento de un usuario del sistema, la señal captada por el acelerómetro y giroscopio está en la figura 10 y en ella se pueden identificar en qué instantes está sucediendo un gesto o no.

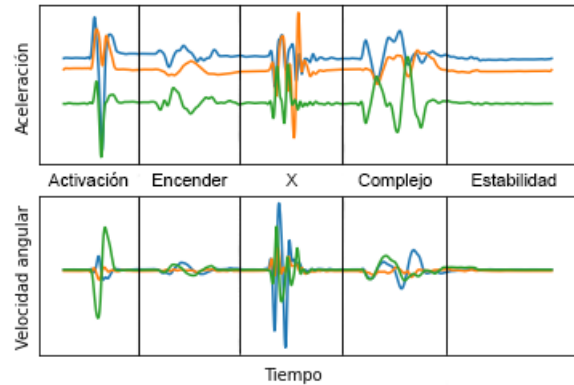


Figura 10: Señal captada durante 20 segundos con varios gestos realizados en ella.

Lo más importante para escoger el tamaño de la ventana y del salto es darle oportunidad al gesto a aparecer en una de las ventanas seleccionadas sin que sea cortada y pierda información para maximizar la posibilidad de tener mejores resultados al intentar hacer una predicción. El comportamiento del modelo bajo esta simulación de tiempo real, o con muestras de larga duración, utilizando la ventana móvil está presente en la figura 11, donde cada color representa la probabilidad de cada clase en la predicción del modelo a medida que la ventana se desplaza sobre los datos de la muestra. De esta simulación se concluye que el modelo clasifica todos los momentos de quietud como el gesto de *Encender* o etiqueta 1. Un ejemplo de esto es el segmento final de la señal que, como se ve en la figura 10, se encuentra en estabilidad y no se está realizando ningún gesto en ella pero en la figura 11 el modelo realizó una predicción que supera el umbral definido.

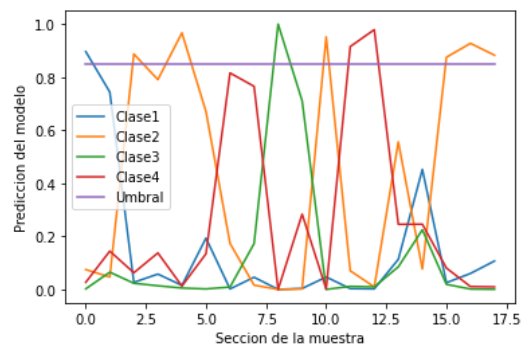


Figura 11: Probabilidad de cada clase en ventana de datos.

Esto representa un problema problema en el sistema y fue

necesario descartar el uso de esa clase y usar solamente los otros 3 gestos restantes en los que no se presenta este error de clasificación. Para finalizar se implementó una función encargada de extender los datos captados por la ventana con el fin de que tengan el tamaño de entrada del modelo después de definir que el tamaño apropiado para la ventana es de 180 datos y mientras que el modelo recibe muestras de 300 datos de longitud.

VI. APLICACIÓN E INTEGRACIÓN CON DOMÓTICA

Antes de comenzar con la creación de la base de datos se realizaron diversos experimentos para determinar cuál era el protocolo de comunicación más adecuado y se confirmó que la señal con mayor cobertura en un hogar es a través de wifi, combinado con un socket UDP en el servidor. Se escogió UDP sobre TCP debido a que este último presentaba problemas de retraso por su funcionamiento interno donde se crea la conexión y validaciones en la transmisión de paquetes lo cual era incompatible con la idea de hacer la aplicación lo más cercana a tiempo real. La arquitectura implementada está compuesta por 4 componentes correspondientes al dispositivo de captura de datos, dispositivos de domótica, servidor de aplicación y servidor mqtt como se puede apreciar en la figura 12.

El flujo de funcionamiento empieza con un usuario usando el dispositivo descrito en la sección III, este se encuentra conectado a la red wifi del hogar por donde envía la señal medida por el acelerómetro a una frecuencia de 60 Hz y esta señal es capturada por un servidor UDP desarrollado en python donde además se encuentra el comité de máquinas encargado de realizar las predicciones cada 60 datos nuevos, cuando el comité de máquinas predice que el gesto realizado corresponde a un gesto de *Activación* es iniciada una máquina de estados que funciona con los próximos dos gestos que se predigan para determinar el dispositivo domótico y la acción que se planea realizar con él, Una vez establecido esto el servidor de la aplicación publica la orden de acción en un tópico correspondiente al dispositivo domótico en un servidor mqtt correspondiente al dispositivo domótico en un servidor mqtt externo, pero dentro de la misma red hogar al que se encuentran suscritos los dispositivos domóticos que el usuario planea controlar¹.

VII. CONCLUSIONES

La arquitectura para el sistema utilizada en este proyecto permite la interacción con dispositivos domóticos mediante el uso de diversas tecnologías dentro de las cuales destaca un dispositivo diseñado para medir los movimientos de un usuario y un modelo de machine learning encargado de identificar la acción a realizar por este.

El uso de una red neuronal convolucional fue clave para el funcionamiento exitoso de este proyecto, debido a su estructura interna capaz de identificar los distintos patrones presentes en las señales de los gestos, y funcionar apropiadamente teniendo en cuenta que dentro de las muestras recolectadas

¹El video con la evidencia de la pruebas funcionales puede ser consultado en este enlace: <https://youtu.be/V9V4NW66U3Q>

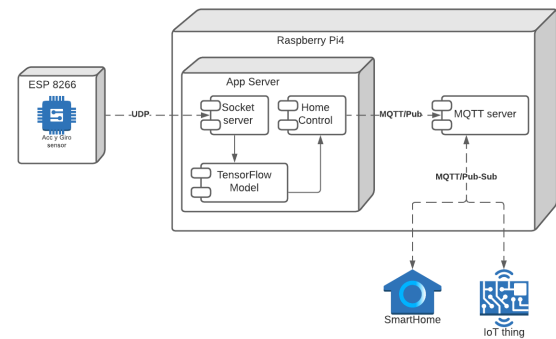
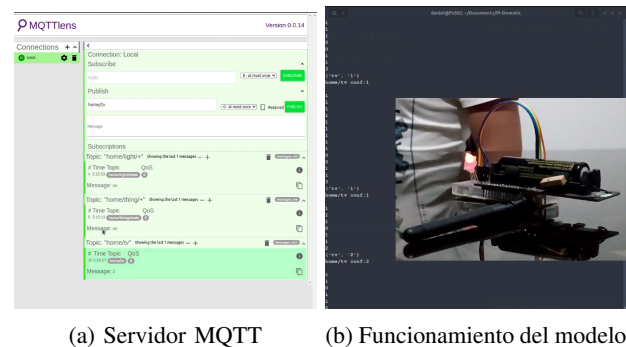


Figura 12: Arquitectura general del sistema.



(a) Servidor MQTT

(b) Funcionamiento del modelo

Figura 13: Evidencia de funcionamiento del sistema

los gestos podían tener diferente duración y tamaños, y estos a la vez estar representando un mismo gesto.

El diseño e implementación de un dispositivo encargado de la captura y transmisión de los datos para el entrenamiento y pruebas del proyecto fue de gran facilidad, gracias al poder y versatilidad del módulo ESP8266-01, que además también se pudo utilizar en el despliegue sin necesidad de hacer modificación alguna y con esto no nos quedan dudas de que un producto final amigables de usar es muy posible.

REFERENCIAS

- [1] J. Ducloux, P. Colla, P. Petrashin, W. Lancioni, and L. Toledo, "Sistema de Reconocimiento de Gestos de la Mano basado en Acelerómetro para Interacción en TV Digital," 2017. [Online]. Available: <https://rdu.iua.edu.ar/handle/123456789/461>
- [2] "14 human computer interaction examples," Nov. 2018. [Online]. Available: <https://www.getsmarter.com/blog/market-trends/14-human-computer-interaction-examples/>
- [3] S. Kallio, J. Kela, and J. Mantyjarvi, "Online gesture recognition system for mobile interaction," in *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme - System Security and Assurance (Cat. No.03CH37483)*, vol. 3, Oct. 2003, pp. 2070–2076 vol.3, iSSN: 1062-922X.
- [4] M. Rehm, Matthias, N. Bee, Nikolaus, E. Andre, and Elisabeth, "Wave like an egyptian accelerometer based gesture recognition for culture specific interactions," 01 2008.
- [5] D. Lima Dantas, C. Figueiredo, and S. Figueiredo, "Sistema de reconhecimento de gestos utilizando smartwatches," 10 2017.
- [6] L. Luengas-Contreras, B. López-Ávila, and J. Jiménez-Espinoza, "Caracterización de unidades de medición inercial (IMUs) en estática y dinámica," *Ingenio Magno*, vol. 8, pp. 92–102, Dec. 2017. [Online]. Available: <http://revistas.ustatunja.edu.co/index.php/ingeniomagno/article/view/1391>

- [7] "Nanounidades de memoria." [Online]. Available: <https://www.investigacionyciencia.es/revistas/investigacion-y-ciencia/el-proyecto-nanodrive-349/nanounidades-de-memoria-3784>
- [8] A. Papoulis, *Sistemas Digitales y Analogicos, Transformadas de Fourier, Estimacion*. Marcombo, Nov. 1985, google-Books-ID: qRPf0Zor0MUC.
- [9] "How convolutional neural networks see the world." [Online]. Available: <https://blog.keras.io/how-convolutional-neural-networks-see-the-world.html>
- [10] J. Xie, B. Xu, and Z. Chuang, "Horizontal and Vertical Ensemble with Deep Representation for Classification," *arXiv:1306.2759 [cs, stat]*, Jun. 2013, arXiv: 1306.2759. [Online]. Available: <http://arxiv.org/abs/1306.2759>