

# Tesis Doctoral

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA



Doctorado en Empresa, Internet y  
Tecnología de las Comunicaciones (EmITIC)



SISTEMA DE IDENTIFICACIÓN DE  
LENGUA DE SIGNOS MEDIANTE  
SERIES TEMPORALES

Jesús Galván Ruiz

Director  
Carlos Manuel Travieso González

Noviembre de 2023, Las Palmas de G.C.



**UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA  
ESCUELA DE DOCTORADO**

Programa de doctorado **Empresa, Internet y Tecnologías de las  
Comunicaciones (EmITIC)**

Título de la Tesis: **Sistema de Identificación de Lengua de Signos  
Mediante Series Temporales**

Tesis Doctoral presentada por **Jesús Galván Ruiz**

Dirigida por el **Dr. Carlos Manuel Travieso González**

Las Palmas de Gran Canaria, a 7 de noviembre de 2023

El Director,  
**Carlos Manuel  
Travieso González**

El Doctorando,  
**Jesús Galván Ruiz**



# Agradecimientos

*Esta tesis está dedicada a mi mujer Sonia y mis hijos Cloe y Miguel, lo más importante de mi vida.*

*Agradecer a mis padres toda la fe y amor que pusieron en mí. Eugenio y Micaela “Allí donde estés mamá, sigues en mi recuerdo”*

*A mis hermanos María, Eugenio y Saray por estar siempre a mi lado.*

*A mi director de tesis Carlos Manuel Travieso por su esfuerzo, trabajo y comprensión en todo el desarrollo de esta tesis, estoy muy agradecido.*

*Agradecer de forma especial a Alejandro García Agrelo, Alejandro Piñán Roescher y a Yeremy Santana Suárez su colaboración desinteresada, gente buena.....buena gente.*

*Agradecer a mis amigos de toda la vida empezando por Luis Laborda y por todas esas “Fieras” indomables, que han participado de forma indirecta, para que este trabajo haya llegado a su fin.*

*A todos mis amigos y colegas de la universidad, Octavio, Jorge, ...y demás gente maravillosa, que han hecho de mí el ingeniero que soy.*

*Agradecer a mis alumnos de 1º de DAW del curso 22-23 del instituto el Rincón su colaboración desinteresada.*

*A mis compañeras/os del instituto el Rincón, por su apoyo en los momentos más complicados de esta tesis, muchas gracias.*

*A todas/os alumnos, compañeros/as y amigos/as que han estado conmigo en todas las etapas de mi vida, y que gracias a ellos/as soy la persona que soy, y que una parte de esta tesis les pertenece.*

*Agradecer por último la colaboración de la asociación de personas sordas de Las Palmas ASOR por su colaboración.*

*Muchas Gracias*



# ÍNDICE

<b>1. Introducción</b> .....	3
1.1 Introducción .....	3
1.2 Motivación.....	3
1.3 Justificación de la propuesta .....	4
1.4 Evolución Tecnológica en el reconocimiento de la lengua de signos.....	4
1.5 Hipótesis y Objetivos .....	7
1.6 Estructura de la memoria .....	7
<b>2. Estado del Arte</b> .....	11
2.1 Introducción .....	11
2.2 Evolución de los Dispositivos de Reconocimiento de Gestos.....	12
2.2.1 <i>Guantes de Datos</i> .....	13
2.2.2 <i>Electrodos EMG (Electromiografía)</i> .....	15
2.2.3 <i>Ultrasonido</i> .....	17
2.2.4 <i>WiFi</i> .....	19
2.2.5 <i>RFID Identificación por Radio Frecuencia</i> .....	20
2.2.6 <i>Cámaras RGB y RGB-d</i> .....	20
2.2.7 <i>Leap Motion</i> .....	23
2.2.8 <i>VR o Realidad Virtual</i> .....	24
2.2.9 <i>UltraLeap</i> .....	25
2.2.10 <i>Kinect y Azure Kinect</i> .....	25
2.3 Conclusiones.....	27
<b>3. Materiales y métodos</b> .....	57
3.1 Introducción .....	57
3.2 Adquisición .....	57
3.2.1 <i>Sensor Leap Motion</i> .....	57
3.2.2 <i>Proceso de Adquisición</i> .....	61
3.3 Base de Datos .....	62

3.3.1 Conjunto de Palabras Usadas .....	62
3.4 Preprocesamiento de Datos .....	64
3.5 Generación de Patrones .....	65
3.5.1 Parámetros .....	65
3.5.2 Generación.....	66
3.6 Comparador.....	67
3.6.1 Introducción al Dynamic Time Warping, DTW .....	67
3.6.2 Desarrollo del DTW.....	68
3.6.3 Esquema del comparador.....	70
3.7 Implementación.....	70
<b>4. Metodología experimental.....</b>	<b>77</b>
4.1 Introducción .....	77
4.2 Validación cruzada o “Cross Validation” .....	77
4.3 Sistema implementado.....	79
4.4 Calidad del sistema.....	80
4.4.1 Matriz de confusión bi-clase .....	80
4.4.2. Precisión.....	80
4.4.3 Recuperación o “Recall” .....	81
4.4.4 F1 .....	81
4.4.5 Exactitud.....	81
4.4.2 Matriz de confusión multi-clase.....	81
4.5 Procedimiento .....	83
<b>5. Resultados.....</b>	<b>87</b>
5.1 Introducción .....	87
5.2 Validación Cruzada .....	87
5.2.1 Conjunto de palabras 1.....	87
5.2.2 Análisis de resultados 1 .....	88
5.2.3 Conjunto de palabras 2.....	90
5.2.4 Análisis de resultados 2 .....	91

5.2.5 Conjunto de palabras 3.....	94
5.2.6 Análisis de resultados 3 .....	95
5.3 Resultado global de la validación cruzada.....	98
5.4 Discusión.....	99
5.5 Eficacia del sistema.....	101
5.5.1 Coste computacional .....	102
<b>6. Conclusiones .....</b>	<b>105</b>
6.1 Valoración final.....	105
6.2 Méritos logrados.....	105
6.3 Futuros trabajos .....	105
<b>Bibliografía .....</b>	<b>109</b>
<b>Apéndice I.....</b>	<b>143</b>
<b>Apéndice II .....</b>	<b>151</b>
<b>Apéndice III .....</b>	<b>155</b>



# Capítulo 1: Introducción



## 1. Introducción

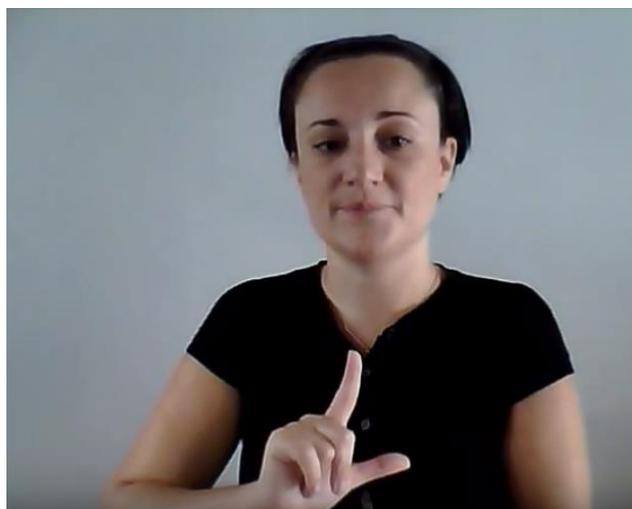
### 1.1 Introducción

En la primera parte de este capítulo, se desarrolla la motivación y justificación por la cual se ha decidido realizar esta tesis. Seguidamente se expondrán los trabajos relativos previos que se han desarrollado anteriormente. En la siguiente sección, se abordarán la hipótesis y los objetivos respectivamente. Por último, se comentará la estructura de la memoria.

### 1.2 Motivación

Desde el principio de los tiempos para poder comunicarse, la humanidad ha creado a lo largo de la historia, diferentes lenguajes para la interacción entre personas. Los humanos tenemos la capacidad no sólo de reproducir sonidos sino de escucharlos. Aun así, hoy en día, existen barreras de comunicación para aquellas personas que tienen su capacidad auditiva disminuida. Muchas de esas personas tienen una incapacidad total para oír, lo que ha hecho, que esas personas desarrollen diferentes formas de comunicación para interactuar entre ellas.

Hoy en día, habitan en el mundo, unos 70 millones de personas sordas según (Nations, s. f.). Esas personas están representadas por más de 135 asociaciones, que promueven el uso del lenguaje de signos para comunicarse. Concretamente en España según el INE, (*Utilización de la lengua de signos por sexo y edad. Población de 6 y más años con discapacidad de audición.*, s. f.), existen 1.230.000 personas que tienen una discapacidad auditiva de diferente grado y tipo. Basadas también en esa misma encuesta, unas 27.300 personas usan la lengua de signos para comunicarse entre ellas. La mayoría de estas personas están integradas en las cuatro asociaciones implantadas a nivel nacional. La lengua de signos española es única en nuestro país, cabe destacar que no existe un lenguaje de signos común a nivel global. Cada país, incluso cada región, tiene su propio lenguaje de signos (véase la Figura 1.1).



**Figura 1.1.** Persona signando una palabra

La motivación principal de esta tesis, es romper la barrera existente entre las personas sordas y parlantes creando un modelo que sirva para que, de forma fácil, económica y sin ningún tipo de componente invasivo, puedan comunicarse de forma fluida entre ellas.

### 1.3 Justificación de la propuesta

Hoy en día, el avance en el desarrollo tecnológico ha sido muy importante. Se han desarrollado muchas tecnologías no sólo en el ámbito de las telecomunicaciones, sino también en el mundo de los sensores. Este avance tecnológico, ha influido en muchas investigaciones que tienen como propósito el reconocimiento de palabras del lenguaje de signos en distintos idiomas, usando todas las tecnologías actuales existentes para lograr dicho objetivo.

Esta tesis aborda esta cuestión, primero analizando las diferentes tecnologías existentes, seguidamente proponiendo una hipótesis e intentando alcanzar una serie de objetivos para su demostración, llegando a una serie de resultados analizados de forma exhaustiva y terminando con un conjunto de conclusiones.

Es muy importante que las personas sordas se integren de la forma más natural posible, en una sociedad donde la interacción de unos con otros es muy importante para el desarrollo personal, emocional y laboral. Romper con las barreras auditivas, supone un gran avance para la inserción de este colectivo.

Actualmente no existe ningún producto o proceso que trabaje este traductor de signos. No obstante, existen numerosas investigaciones en curso, alguna de ellas desarrolladas en el estado del arte de esta tesis. Estas investigaciones abordan el problema de la interpretación de los gestos, tanto estáticos como dinámicos, con diferentes tipos de resultados.

### 1.4 Evolución Tecnológica en el reconocimiento de la lengua de signos

Desde los años 60 se han creado diferentes tipos de dispositivos que intentan ayudar a la interacción hombre-máquina. El primer dispositivo que podemos encontrar es el *Sketchpad* (véase la Figura 1.2). Donde se podía interactuar directamente sobre una pantalla. A finales de los 60, apareció el “espacio sensitivo”, que era capaz de extraer la forma de una persona sobre un fondo. A partir de los años 80 con la entrada y desarrollo de la fibra óptica, se diseñó el *DataGlove*, que podía visualizar en una pantalla los movimientos de la mano. Para el entretenimiento, Nintendo desarrolló un guante que servía como mando para alguno de sus juegos. Ya en los años 90, el uso de sensores resistivos dio la posibilidad de crear guantes, capaces de detectar el movimiento a partir de su posición. A finales de los 90 con la nueva aparición de sensores como el acelerómetro, se desarrolló el *AcceleGlove*, capaz de trasladar los movimientos de la mano gracias a ese tipo de sensores. En el año 2010 Microsoft desarrolló la cámara *Kinect*, para el entretenimiento en sus consolas *Xbox*. Esta cámara tenía la posibilidad de reconocer los movimientos de una persona y trasladarlos al interior de un juego, creando una interacción en tiempo real con el mismo. Dos años después salió al mercado el

*Leap Motion*. Este dispositivo es un sensor volumétrico con la capacidad de reconocer el movimiento de las manos en tres dimensiones. En 2014 se desarrolló un brazalete de ultrasonidos que reconoce el movimiento de las manos. Con el paso del tiempo, Microsoft dejó de desarrollar la cámara *Kinect* para uso comercial y de ocio, y decidió a partir de la misma, desarrollar la cámara *Azure Kinect*, pero esta vez con un propósito profesional e industrial. Actualmente, ya no se comercializan cámaras *Kinect* para ocio. En el año 2020, la empresa *UltraLeap* que también comercializa el *Leap Motion*, desarrolló un dispositivo capaz de detectar el movimiento de las manos añadiendo una base háptica, dando la sensación de que estás tocando el dispositivo. Para ello, utiliza una serie de ultrasonidos que simulan el tacto del dispositivo, dando la sensación de palpar un objeto.

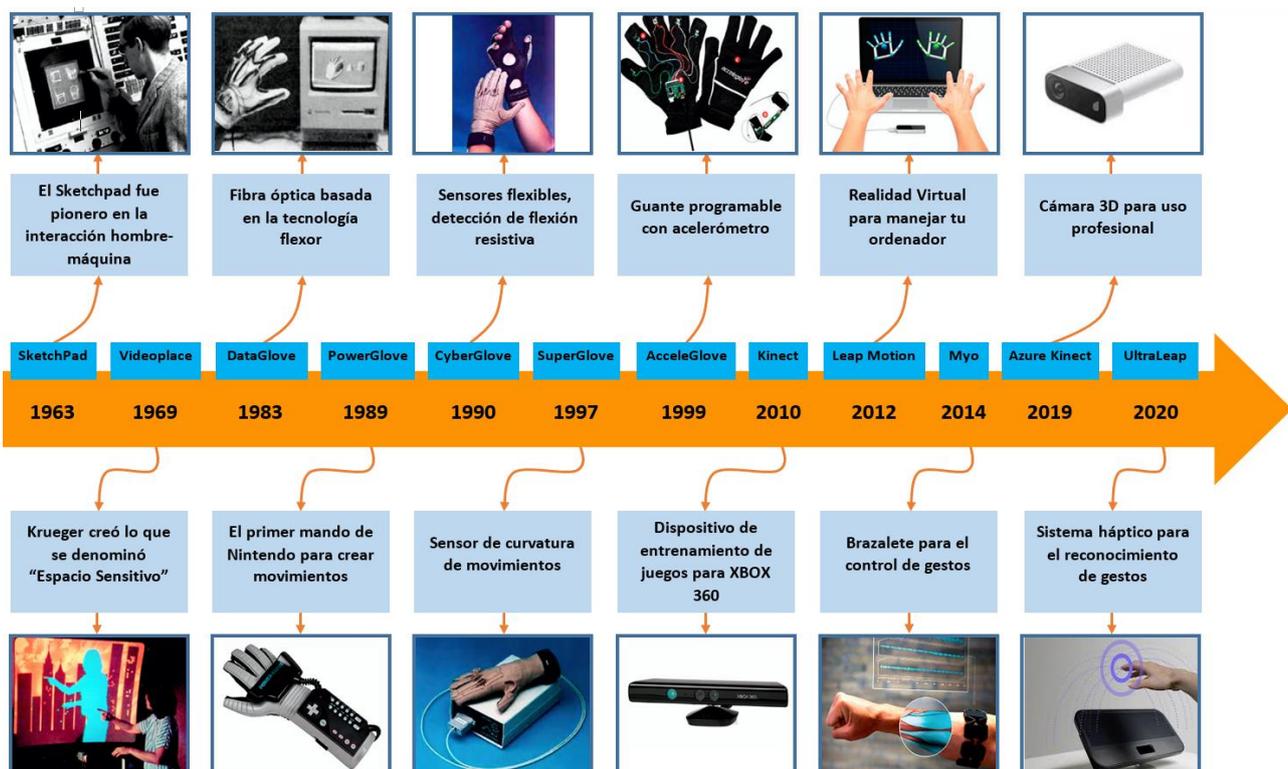


Figura 1.2. Dispositivos diseñados hasta el momento para el reconocimiento gestual

El dispositivo que se ha empleado en diferentes trabajos incluida esta tesis, es el *Leap Motion* (*API Overview — Leap Motion JavaScript SDK v3.2 Beta documentation, s. f.*). A día de hoy, se pueden encontrar diferentes trabajos de investigación que usan el *Leap Motion Controller (LMC)*, como sensor volumétrico en 3 dimensiones. Las investigaciones que se han propuesto para su estudio tienen están divididas en tres conjuntos. El primer conjunto trabaja con signos dinámicos, o sea palabras. El segundo grupo son el de los signos estáticos o alfabeto dactilo. El último conjunto es aquel que mezcla tanto los signos dinámicos como los estáticos.

En cuanto a los trabajos relativos de signos estáticos está por ejemplo, (Funasaka et al., 2015) que usó 24 signos estáticos que corresponden al alfabeto de la lengua de signos americana (ASL), usando para su reconocimiento la técnica del árbol de decisión con una precisión del 82,71%. Otro estudio que usa máquina

de soporte de vectores (SVM) y además combina *Kinect* con LMC es (Marin et al., 2014), obtuvieron una eficiencia del 80,86% para 10 signos de ASL.

Otra investigación, que trabaja con la lengua de signos griega (GSL) (Simos y Nikolaidis, 2016), usa perceptrón multicapa (MLP), para dedos y manos, sin que interactúen entre ellos. Sus resultados de exactitud estaban entre el 99,08% y 98,96%, respectivamente. Para ASL y combinando 32 números y signos de ASL (Mapari y Kharat, 2016), alcanzaron aproximadamente un 90% usando el MLP. Otro estudio como el de (Vaitkevicius et al., 2019), usó 24 letras estáticas con ASL e intentaba crear frases, que posteriormente usando regresión lineal, consiguió un acierto del 86,1%. Durante esa investigación se trabajó de forma conjunta con el *Leap Motion* y Realidad Virtual (VR).

Entre los investigadores que desarrollaron sistemas utilizando signos dinámicos y estáticos están, (Mohandes, Aliyu, et al., 2014; Mohandes, Deriche, et al., 2014), donde ambos estudios, trabajaron con 28 signos del Lenguaje de Señas Árabe (ArSL) de forma estática, excepto dos signos dinámicos. El primer experimento se usaron las técnicas de Naïve Bayes y MLP, el vecino más cercano o K-NN, y modelos ocultos de Markov o HMM. Los resultados para los dos experimentos fueron de 99,1%, del 97,7% y por último del 98,3%, respectivamente. Otra investigación a tener en cuenta, que usa redes neuronales o NN es (Hisham y Hamouda, 2018). Ellos alcanzaron una exactitud del 90,35% para 16 palabras estáticas con el lenguaje de signos árabe, usando *Leap Motion* y *Kinect*. También llegaron a obtener una eficiencia del 95,22% usando K-NN, y del 90,78% hasta el 89,12% con SVM. Al mismo tiempo, en ese trabajo, lograron una exactitud del 96,15% con 26 letras del lenguaje de signos americano usando el MLP. Por último, llegaron a obtener con DTW y 20 palabras dinámicas del lenguaje de signos árabe, una eficiencia del 96,41%. Otro trabajo a reseñar es el de (Naglot y Kulkarni, 2016) que obtuvo una eficiencia del 96,15% usando MLP y trabajando con 26 letras del lenguaje de signos americano.

Hay más trabajos de investigación que usan letras del lenguaje de signos americano como el de (Chong y Lee, 2018). Ellos trabajaron con 10 dígitos y 26 letras. Estos experimentos, sólo con letras, arrojaron unos resultados para redes neuronales profundas o DNN del 93,81% y con máquina de soporte de vectores del 80,30%. Si además se añaden los 10 dígitos para DNN, se alcanzó el 88,79% y para SVM el 72,79%. Otra investigación que trabajó con signos estáticos y dinámicos usando el lenguaje de signos americano es (C. K. M. Lee et al., 2021). Su investigación usó 2 letras dinámicas y 24 estáticas. Usando redes neuronales recurrentes o RNN lograron un 98,19%. Con memoria a largo plazo o LSTM alcanzaron una eficiencia del 97,96% y para la máquina de soporte de vectores un 98,35%. El uso de redes neuronales convolucionales los empleó (Tao et al., 2018) para 26 letras del lenguaje de signos americano, obteniendo una exactitud que variaba entre el 99,7% y el 80,1%. Aplicando técnicas como la máquina de soporte de vectores y del vecino más cercano se encuentra la investigación de (Anwar et al., 2017). Su trabajo basado en el lenguaje de signos indonesio, logró un 93,85% para SVM y del 95,15% para el K-NN. Basado en el lenguaje de signos árabe, se encuentra (Alnahhas y Alkhatib, 2020). Ellos usaron 15 palabras logrando un 96% de eficiencia usando LSTM.

Dentro del último trabajo que usa tanto signos dinámicos como estáticos está (Avola et al., 2019), donde con una exactitud del 96,41% y usando LSTM, pudo distinguir 12 signos dinámicos y 18 estáticos.

Por último, comentar los trabajos que solamente utilizaron signos dinámicos. En primer lugar está (Jenkins, 2021). Él usó 13 palabras dinámicas usando diferentes técnicas y comparando unas con otras. Con redes neuronales obtuvo una precisión del 99,9%, con el árbol de decisión del 99,7%, con SVM del 99,9%, K-NN del 98,7% y finalmente con Naïve Bayes un resultado del 96,4%. Finalmente, comentar la investigación realizada por (Elons et al., 2014), que usando MLP para 50 palabras del ArSL, alcanzó una eficiencia del 88%.

### 1.5 Hipótesis y Objetivos

Los sistemas de reconocimiento de signos han sido estudiados profundamente en muchos trabajos de investigación usando diferentes tecnologías. Esta tesis doctoral trabaja con el sensor volumétrico *Leap Motion* y el sistema de clasificación denominado deformación dinámica del tiempo (DTW), donde intenta aplicar nuevas técnicas para el reconocimiento de signos dinámicos de lenguaje de signos español. En esta tesis, se propone validar la siguiente hipótesis;

*“Es posible crear un traductor del lenguaje de signos español usando un sensor económico, fiable y fácil de manejar, traduciendo los signos con independencia de la persona que signa.”*

El objetivo general es crear un sistema de reconocimiento de palabras del lenguaje de signos español, a partir de la hipótesis planteada.

Objetivos específicos:

- Crear una base de datos extensa de patrones de palabras de uso cotidiano que demuestren la viabilidad del sistema.
- Utilización de dispositivos económicos y fáciles de manejar por cualquier usuario.
- Procesar y clasificar las señales captadas usando técnicas de *machine learning*.

### 1.6 Estructura de la memoria

Este documento está compuesto por 6 capítulos, la bibliografía y 3 anexos. Después de este primer capítulo, la estructura de la memoria es la siguiente;

- Capítulo 2: contiene el estado del arte, donde en el mismo se profundiza en los diferentes sensores que existen en la actualidad y en las técnicas usadas para tal fin.
- Capítulo 3: se especifican los materiales y métodos usados para la investigación tales como el sensor volumétrico, las técnicas usadas en el procesamiento de datos, generación de patrones y sistema de clasificación.

- Capítulo 4: en este capítulo se incluyen la metodología experimental aplicada para esta tesis, y la descripción de los parámetros de calidad, para medir los resultados.
- Capítulo 5: se muestran los resultados obtenidos de los experimentos para validar la hipótesis planteada.
- Capítulo 6: muestra una valoración final, los méritos logrados gracias a la consecución de la tesis, como las líneas futuras de trabajo para la ampliación de esta investigación.
- Bibliografía: en esta sección se incluye toda la relación de bibliográfica necesaria para el desarrollo de la investigación.
- Anexo I: en esta sección se incluyen diferentes gráficos relacionados con los resultados, que complementan la investigación. Por ejemplo, están todas las matrices de confusión de donde se han extraído los resultados expuestos en el capítulo 5.
- Anexo II: en esta parte se incluye la tabla de figuras y las referencias a las fuentes de algunas de ellas.
- Anexo III: esta parte incluye fragmentos de los diferentes artículos desarrollados y enlaces a vídeos promocionales sobre esta investigación.

# Capítulo 2:

# Estado del

# Arte



## 2. Estado del Arte

### 2.1 Introducción

Dentro de la comunicación no verbal, la parte gestual es fundamental, ya que el lenguaje corporal es capaz de enviar mensajes, sin que el habla intervenga en ellos. Estos gestos se realizan a través del movimiento de diferentes partes del cuerpo, siendo los más usuales las manos, brazos y cara.

Los gestos manuales son los más expresivos y los más usados, ya que los mismos suelen parecer, de forma general, gestos más naturales e intuitivos para las personas, impactando en el mensaje del habla. Por ejemplo, en el contexto de los mensajes emocionales en (Mehrabian, 2017), tanto las señales como los gestos, pueden llegar a representar hasta un 55% del impacto del mensaje.

La comunicación no verbal es usada por la humanidad desde hace millones de años. Los gestos clásicos como asentir con la cabeza para decir sí o girarla para negar algo, son habituales. Otros como encogerse de hombros o fruncir el ceño, se usan de forma natural por muchas personas. También existen algunos gestos que se han adquirido de los animales. Por ejemplo, enseñar los dientes para mostrar agresividad. Existen muchos ejemplos de lenguajes creados a partir de gestos, donde se utilizan en diferentes disciplinas para comunicarse. Un ejemplo puede ser los usados en buceo, primeros auxilios y aviación (véase la Figura 2.1). También destacar que las personas sordas se comunican a través del uso de este tipo de lenguajes gestuales.



Figura 2.1. Diferentes usos del Lenguaje de Signos

Hoy en día, las tecnologías son una parte natural de la vida diaria del ser humano, cubriendo un número creciente de aplicaciones. El constante aumento de la potencia y velocidad de los procesadores y sensores, han posibilitado la creación de aplicaciones cada vez más complejas, aumentando proporcionalmente la necesidad de una interacción hombre-máquina (IHM) más eficiente.

En este contexto, actualmente se consolida el uso de pantallas táctiles e interfaces de voz, mientras que, el cambio del teclado-ratón por interfaces más naturales e intuitivas aún está lejos de ser una realidad. En la actualidad, hay que contemplar que el uso de las manos para el manejo de una interfaz, puede ser un inconveniente. Los ejemplos podrían ser, un conductor que intenta consultar el sistema de navegación de un vehículo sin tener que quitar las manos del volante, o una persona en una reunión interactuando

## 2. Estado del Arte

---

discretamente con un dispositivo informático. A todo esto, hay que destacar que cuando se produce una manipulación directa y física, la propagación de microbios podrían alterar de forma crítica algunos espacios, como pueden ser las áreas de cirugía de un hospital.

En este sentido, el reconocimiento de gestos puede ser útil para incrementar aún más la interacción entre máquina y hombre. Las capturas de vídeo digital, han traído muchos intentos de reconocer gestos dinámicos. La actualización de nuevas tecnologías como sensores de profundidad o cámaras de alta resolución, permiten desarrollar formas de detección de movimientos y actuar en tiempo real.

Esas capturas son aplicadas en áreas como en domótica, interfaces de PC, control de pacientes médicos, manipulación y navegación para realidad virtual o aumentada, entre otros.

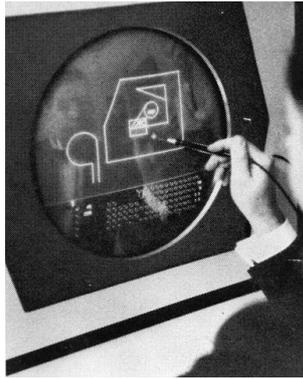
Por definición, la interacción consiste en hacer gestos comprensibles para que los equipos de computación puedan interpretarlos. Es necesario en primer lugar, captar los movimientos o sea, la velocidad y aceleración junto con los ángulos y rotaciones. Para realizar esta tarea, fundamentalmente hay dos enfoques; el primero de ellos, consiste en el uso de los dispositivos sensores a modo de guantes o pulseras (no ópticos); y el segundo, se apoya en el uso de técnicas de cámara de visión (ópticos). Existen distintas desventajas como la latencia, calibración, costo, rango de movimiento o comodidad del usuario.

Es importante diferenciar entre gestos y poses. En un principio cuando una persona realiza un gesto estático, se denomina pose, en cambio cuando el gesto es dinámico y se suceden una secuencia de poses, estamos hablando de un acto dinámico. Un ejemplo puede ser, usar la mano para decir hola. De esta forma se entiende que existen 2 niveles para reconocer los gestos. Un nivel inferior para interpretar una pose y uno superior para reconocer gestos.

Cabe reseñar que, interpretar tanto los gestos como las poses no es sencillo. Una de esas dificultades, es el hecho de que pueden existir varios significados asociados a un mismo gesto, al mismo tiempo estos significados pueden ser interpretados de forma distinta por otras culturas, contextos o idiomas. Un ejemplo, puede ser el de los lenguajes de signos, donde usando el mismo idioma, los signos pueden variar según la zona o región donde se signe.

### 2.2 Evolución de los Dispositivos de Reconocimiento de Gestos

Hoy en día, se intenta que el lenguaje gestual sea reconocido de la forma más intuitiva y natural. Para ello el avance de los sensores está siendo fundamental en el logro de dicho objetivo, haciendo que la interacción hombre-máquina, sea más sencilla. El hecho de que los sensores hayan evolucionado, ha añadido múltiples mejoras para alcanzar ese objetivo. Al principio de la década de los 60, algunos investigadores lograron capturar la escritura usando una pantalla (Sutherland, 1963). Este fue el comienzo para el uso de lápices y tabletas especiales (véase la Figura 2.2). A finales de la década de los 60, se empezó a trabajar con algunos prototipos de realidad virtual (Krueger et al., 1985). A partir de ese momento se pensó que, en un futuro no muy lejano, existiría un mundo sin pantallas, de forma que la gente podría interactuar con su entorno de una forma directa.



**Figura 2.2.** Escritura sobre Sketchpad en 1963

No sería hasta los años 80, cuando se iniciaron las capturas de movimiento mediante guantes con sensores de flexión y posición. En la década de los 90, se incrementaron los trabajos de identificación de gestos en imágenes y vídeo con métodos de visión artificial, algo que ha ido mejorando hasta la actualidad, donde se dispone de sistemas en tiempo real y de seguimiento de personas (Myers, 1998; Sharma y Verma, 2015). Por otro lado, existen sistemas electromagnéticos que son capaces de localizar la posición de un objeto midiendo dichos campos.

### 2.2.1 Guantes de Datos

Este tipo de dispositivo se coloca en las manos, y a partir de ahí se extraen los movimientos que dan como resultado la posición. Por lo general, tienen incorporados sensores táctiles o de otro tipo, capaces de reproducir una posición exacta de los movimientos realizados por las diferentes articulaciones de la falange y la muñeca. Una de las grandes ventajas que tiene a la hora de obtener los descriptores, es que no necesita ningún procesamiento de datos, como en el caso de una imagen capturada por una cámara. La principal desventaja para el usuario, es su incomodidad.

Las dos categorías existentes de este tipo de dispositivos, son pasivos (véase la figura 2.3) y activos (Premaratne et al., 2010). Los guantes que tienen algún tipo de sensor están en el conjunto de activos. Éstos pueden tener acelerómetros, potenciómetros, resistencias capacitivas u otros. Estos guantes se conectaban por cables a los ordenadores, cosa que en la actualidad han sido sustituidos por tecnología inalámbricas. Cuando se hablan de guantes pasivos, se refieren a elementos que no poseen elemento electrónico. Algunos de ellos, además, pueden tener colores que ayudan a ser identificados.



**Figura 2.3.** Guantes Pasivos que ayudan a diferenciar la posición de los dedos

El primer tipo en aparecer fue el *Sayre Glove* en 1977 y basaba su funcionamiento en tubos flexibles (en lugar de fibra óptica), para medir la flexión de los dedos emitiendo un haz de luz en un extremo, mientras que un sensor (diodo fotoeléctrico), detectaba la intensidad de ese haz emitido según la flexión y posición de los dedos.

Los laboratorios del *MIT Media Lab*, desarrollaron un dispositivo de captura de movimiento, a principios de los años 80. Su objetivo era capturar los movimientos, no obtener el control de un instrumento. Cuando intentaron llevar a cabo el proyecto, se toparon con una tecnología inmadura (Sturman & Zeltzer, 1994), por lo que pronto desecharon la idea.

Más adelante en 1983, se desarrolló una patente que era capaz de reconocer las posiciones de la mano y al mismo tiempo era capaz de gesticular caracteres alfanuméricos. Éste, además, fue conocido como el primer guante de realidad virtual (Zimmerman, 1985). Además, el mismo se comercializó posteriormente. Este *DataGlove*, se componía de fibras ópticas ligeras, que estaban pegadas en el dorso de los dedos. A medida que los dedos se doblaban, la luz se atenuaba, haciendo que los sensores recogieran la información y los procesaran de forma adecuada, determinando el ángulo del movimiento. Esto requería por parte de cada usuario una pre-calibración.

Entre el año 1989 y 1995, vieron la luz dos tipos de guantes, el *Super Glove* y el *Power Glove*. Estos guantes, usaban para medir la flexión de los dedos, materiales con resistencia eléctrica. Una de sus grandes ventajas es que eran más baratos que los anteriores, aunque menos precisos (LaViola, 1999). Posteriormente se añadieron acelerómetros y sensores magnéticos, haciendo que estos cambios le dieran más exactitud. También, a partir de ese momento comenzaron a aparecer dispositivos de este tipo para usos académicos y comerciales. Actualmente, el *CyberGlove* (*CyberGlove Systems LLC*, s. f.), (véase la Figura 2.4), es el modelo más reconocido y ha sido especialmente desarrollado para la industria de la animación gráfica y el cine. Además, el *MIT AcceleGlove* (Hernandez-Rebollar et al., 2002) está siendo ampliamente utilizado para entornos de realidad virtual.



Figura 2.4. CyberGlove III usado para el control del Robots

Los guantes de datos han tenido una evolución tan importante, que su uso se ha universalizado en muchos campos, desde la realidad virtual y el entretenimiento, pasando por la comprensión de signos, la medicina (Santos et al., 2015) o la producción industrial (Moreno et al., 2013).

### 2.2.2 Electrodo EMG (Electromiografía)

Un dispositivo diferente a los guantes explicados con anterioridad, son las pulseras. Éstas tratan de eludir todas las posibles interferencias ambientales que se producen, por ejemplo, cuando se interacciona con la voz. Estas pulseras están compuestas por sensores de electromiografía. Su funcionamiento se basa, en la medición de los potenciales producidos por las células musculares durante una actividad.

Entre estas tecnologías de reconocimiento, se está trabajando activamente en estudios basados en el reconocimiento de bioseñales; señal de electrooculografía: actividad eléctrica de los movimientos del globo ocular o “posición de la mirada” registrada alrededor de los ojos, señal de electroencefalografía: actividad eléctrica cerebral registrados desde el cuero cabelludo, señal de electrocardiograma: actividad eléctrica del corazón y otros tipos de bioseñales.

Al estudio del sistema neuromuscular, se le denomina electromiografía. Éste se encarga de analizar, detectar y procesar mediante el uso de electrodos, las señales eléctricas que producen tanto los nervios como los músculos.

La composición de los músculos se basa en células especializadas, que tienen la capacidad de relajarse o contraerse. De esta forma son capaces de producir movimientos, proporcionar una adecuada estabilización y temperatura (Reaz et al., 2006) o transportar sustancias por el cuerpo. Gracias a las propiedades, mecanismos de control y estructura, se pueden aglutinar en tres tipos de tejidos: músculos cardíacos, músculos lisos y músculos esqueléticos. Los músculos cardíacos son los que producen los latidos del corazón. Los músculos lisos se encuentran en su mayoría en diferentes órganos, como pueden ser, por ejemplo, los intestinos. Por último, los músculos esqueléticos se encargan de mover el cuerpo. De estos últimos se obtienen las señales EMG (Valentinuzzi, 2007).

El organismo es capaz de generar potenciales bioeléctricos producidos por flujos de corriente iónicas. A estos potenciales bioeléctricos se les denomina potenciales iónicos.

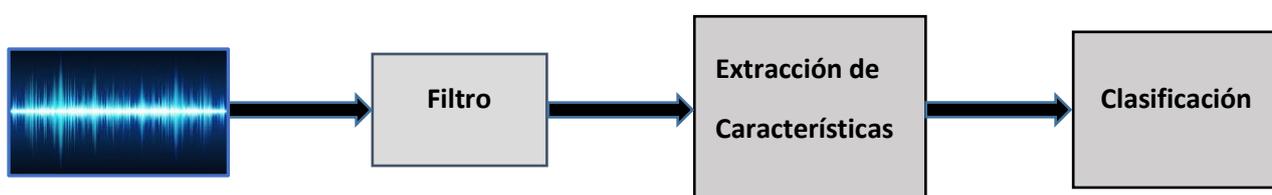
## 2. Estado del Arte

La medición eficiente de estos potenciales iónicos requiere que se conviertan en potenciales eléctricos, antes de que puedan medirse con métodos convencionales. Para obtener el voltaje tanto superficial como intramuscular, es necesario el uso de electrodos (Farina et al., 2016). Los electrodos EMG son capaces de registrar potenciales de todos los músculos a su alcance. Para la medición existe un problema, y es que los músculos grandes pueden interferir en la medición de las señales EMG de los músculos pequeños, aun cuando los electrodos se coloquen en esos músculos. Es por eso, que las señales que se obtienen, son la suma de esos potenciales sumados de forma individual y que forman la fibra del músculo. Si existiese un problema por el cual no se pudiera registrar de forma adecuada esa medición, entonces habría que insertar directamente la aguja sobre el músculo. Es por eso que, cuando se realiza una electromiografía intramuscular, es para estudiar patologías de las unidades motoras o la fisiología de la misma. En cambio, cuando se quiere medir diferentes patrones de fatiga muscular o de actividad temporal, entonces se realizan estudios sobre la superficie del músculo (Merletti et al., 2004)

Los factores como la distancia entre electrodos, el área de actividad muscular, las propiedades de la piel, el procesamiento de la señal y el contacto que se produce entre la piel y el electrodo, son fundamentales para el análisis de las señales EMG, ya que, de todo el conjunto mencionado, dependerá directamente la amplitud y sus propiedades.

De forma habitual, las señales que se obtienen a partir de los electrodos deben procesarse y amplificarse. De acuerdo al tipo de electrodos que se usen, al tamaño y al espacio que haya entre los mismos, se pueden ver modificada el ancho de banda, el tiempo y la amplitud de la señal (Gowitzke y Milner, 1999).

La variación del ancho de banda de una señal EMG, puede variar entre 10Hz y 10 KHz, y su amplitud entre 10  $\mu$ V a 5 mV, según sea el electrodo. El electrodo que comercialmente más se utiliza es el de plata-cloruro de plata (Ag-Ag Cl), que es capaz de reducir el ruido y tener una mayor estabilidad. Para tratar la bioseñal generada, es necesario implementar un sistema serie de fases, (véase la Figura 2.5).



**Figura 2.5.** Fases para el Tratamiento de las señales EMG

Para hacer cualquier tipo de medición de una bioseñal EMG, es necesario el uso de al menos tres electrodos, donde, uno se utiliza para la puesta tierra, y los otros dos para medir la entrada diferencial. A veces, se agregan electrodos activos o activados para crear un circuito de control de retroalimentación entre el sensor y el cuerpo (Webster, 2009), eliminando así la necesidad, de usar gel conductor entre los electrodos y la piel (su uso disminuye la impedancia eléctrica).

Tradicionalmente, la tecnología EMG se ha empleado para el diagnóstico médico (Luca, 1997), control protésico (Hefftner et al., 1988) y tecnologías de rehabilitación médica (Merletti y Farina, 2016). Existen estudios centrados en el uso de señales EMG, que ayudan a personas con discapacidad física, a poder interactuar con las interfaces gráficas de usuario (GUI), sustituyendo el ratón de forma continua o reconociendo gestos de forma estática (Ahsan et al., 2009; Barreto et al., 1999; Coleman, 2001; Guerreiro y Jorge, 2006). Otros ejemplos de aplicaciones EMG para la interacción hombre-máquina, incluyen el control robótico (Crawford et al., 2005; Joshi et al., 2013; Kiguchi y Hayashi, 2012), reconocimiento de voz sin voz (Manabe et al., 2003), reconocimiento de estado emocional (Benedek et al., 2005; Healey y Picard, 1998), interfaces de expresión musical (Dubost y Tanaka, 2002; Knapp y Lusted, 1990; Tanaka y Knapp, 2002), y reconocimiento de gestos genéricos, utilizándolos para controlar un reproductor de música (Fistre y Tanaka, 2002), como un puntero GUI (Rosenberg, 1998), o como un teclado numérico (Dhillon et al., 2009; Wheeler y Jorgensen, 2003).

En cuanto a los productos existentes en el mercado, debido al reciente desarrollo de las comunicaciones inalámbricas y la computación integrada, se pueden encontrar dispositivos portátiles muy extendidos para la obtención de datos EMG mediante pulseras (Saponas et al., 2009; Tabor et al., 2016). Estos sensores se ubican de forma radial y sujetos por una banda flexible alrededor de la muñeca de la mano. Un ejemplo destacado podría ser el brazalete Myo producido por la empresa Thalmic Labs, que se ha utilizado en muchas aplicaciones y experimentos (Abreu et al., 2016; Benalcázar, Jaramillo, et al., 2017; Benalcázar, Motoche, et al., 2017; Boyali et al., 2015; Sathiyarayanan y Rajan, 2016). Por último, destacar a (Singh y Chaturvedi, 2023) donde con su investigación logró reconocer 24 letras del lenguaje de signos americano con una exactitud del 99,91%.

### 2.2.3 Ultrasonido

Se pueden distinguir dos técnicas para la captura de gestos. Uno de ellos utiliza imágenes ultrasónicas (sonomiografía), y el otro aprovecha el efecto *Doppler*, usando la radiación ultrasónica en cualquier estancia. La capacidad de mostrar los músculos del interior del cuerpo en tiempo real, es una de las técnicas de uso de imágenes ultrasonido.

Los tejidos por su propia naturaleza, poseen diferentes impedancias acústicas haciendo que el paso de una onda sonora de un tejido hacia otro, desprendan distintas cantidades de energía. Gracias a este proceso, se puede formar una imagen de ultrasonido. Es poco probable que, proporcione una mayor precisión que otros sistemas que miden directamente las partes móviles del cuerpo, como cámaras o guantes de datos, pero en cambio, como en EMG, se pueden obtener beneficios potenciales en cuanto a pérdidas de información debido a la oclusión, ya que en este caso cualquier parte del cuerpo no quedará oculta a los demás. Destacar que entre la EMG y la sonomiografía, existe una diferencia y es que esta última tiene menos problemas a la hora de diferenciar entre diferentes tipos de músculos. En cambio el EMG, es bastante limitado para el reconocimiento del movimiento amplio de la mano (Sikdar et al., 2014)

## 2. Estado del Arte

---

Diversos estudios han surgido al respecto: desde 2003, midiendo la contracción muscular de varios músculos (Hodges et al., 2003), incluyendo los siguientes estudios posteriores (X. Chen et al., 2010; Guo et al., 2008; Zheng et al., 2006). En 2013 un estudio planteó una banda de transductores ultrasónicos alrededor del antebrazo y los dedos (Mujibiya et al., 2013), otros estudios publicados detectaron seis gestos en 2015 (Hettiarachchi et al., 2015). (McIntosh et al., 2017) desarrolló una sonda, que era capaz de monitorizar y detectar los gestos de una mano. Esta sonda se colocaba en el antebrazo.

Otra forma de capturar gestos, es usando el efecto *Doppler*, con señales de frecuencia ultrasónicas. Esta técnica consiste básicamente en un transmisor que emite tonos ultrasónicos continuos, que rebotan en objetos (brazos y manos en reconocimiento de gestos), dentro del campo de detección. El sensor es capaz de capturar las señales que se reflejan, para luego ser analizadas y procesadas. La señal tiene la misma frecuencia cuando se está usando sobre un elemento estático, cambiando solamente la fase y la amplitud. El cambio de frecuencia se produce cuando un objeto es dinámico, produciendo lo que se denomina el desplazamiento Doppler. Esta frecuencia tratada debidamente, es capaz de proporcionar la velocidad del objeto movimiento.

Esa reflexión que se produce cuando un objeto está en movimiento, se describe con la fórmula  $fd = 2vf / cv$ , donde la desviación de la frecuencia es la  $fd$ , la frecuencia emitida es la  $f$ ,  $c$  representa a la velocidad del sonido y por último  $v$  es la velocidad del objeto en movimiento (Tahmoush y Silvius, 2009).

Siguiendo con los mismos argumentos y parámetros, cuando existen varios objetos en movimiento, la señal que se recibe tendrá muchas frecuencias, una para cada objeto.

Un ejemplo claro, pueden ser los radares de controles utilizados por la policía, que, aunque emplea ultrasonidos, son capaces de averiguar la velocidad de los vehículos. En cambio, estos radares sólo pueden capturar la velocidad de un solo objeto (Tivive et al., 2010).

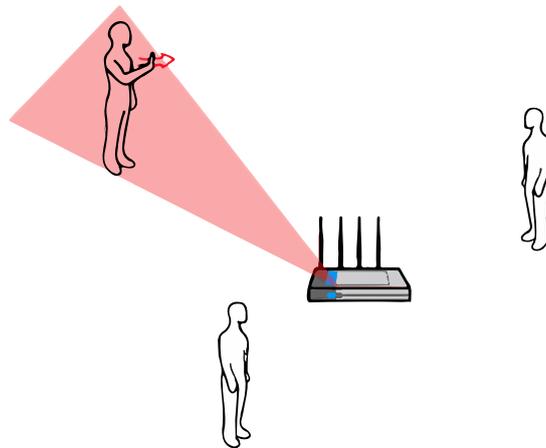
Aunque los rangos ultrasónicos se han explotado considerablemente en el mercado (Biber et al., 1980), la investigación sobre esta técnica aplicada para el reconocimiento de gestos por ultrasonido, no ha sido extensamente documentada.

Los primeros informes muestran el desarrollo de un sistema simple de reconocimiento de gestos por ultrasonidos para el control de la música (Fléty, 2000). Posteriormente, a partir del 2009, existen investigaciones desarrolladas para el reconocimiento de manos y se consideró un tema importante según (Kalgaonkar y Raj, 2009), basadas en la potencia de la señal que se refleja, usando receptores en 3D. En el año 2011, se investigó sobre el reconocimiento de gestos utilizando el telémetro ultrasónico convencional, en un robot móvil (Kreczmer, 2011). Posteriormente, tanto la Universidad de Washington como la empresa *Microsoft Research*, crearon un sistema para equipos portátiles, capaz de reconocer gestos a corta distancia (Gupta et al., 2012). Otro grupo de trabajo liderado por la Universidad de Berkeley, desarrolló un chip que contenía una gran cantidad de transductores ultrasónicos en forma de matriz, que usaban para el reconocimiento de gestos (Przybyla et al., 2012, 2014). Esta línea de investigación continuó desarrollándose,

buscando un hardware más comprimido y con sensores menos espaciados, como los trabajos de (Booij y Welle, 2014), y (Saad et al., 2018; Sang et al., 2018).

#### 2.2.4 WiFi

En la actualidad, existen muchas infraestructuras WiFi. Con el objetivo de reutilizar las WiFi existentes, lo que permitiría economizar el empleo de hardware y facilitar el despliegue de sistemas a gran escala, la búsqueda del reconocimiento de gestos a través de redes WiFi empezó hace unos 12 años (véase la Figura 2.6). Una ventaja adicional de esta técnica, consiste en su capacidad para proporcionar reconocimiento de gestos, incluso en escenarios sin visión directa. Esto se detallará más adelante.



**Figura 2.6.** Reconocimiento de diferentes posiciones de manos usando WiFi

Estos sistemas, para obtener la información necesaria y así reconocer un gesto, usan distintos indicadores. Por ejemplo, algunos trabajos de investigación se basan en el indicador de intensidad de la señal recibida (Kellogg et al., 2014), en el indicador de tiempo de vuelo de la señal que corresponden con sus siglas en inglés *ToF* (Adib et al., 2014) o en los desplazamientos *Doppler* (Pu et al., 2013). Sin embargo, estos sistemas requieren dispositivos especializados (Adib et al., 2014; Adib y Katabi, 2013), o necesitan modificar dispositivos comerciales ya existentes (Pu et al., 2013), y algunos de ellos son excesivamente susceptibles a las interferencias (Kellogg et al., 2014). En consecuencia, las investigaciones actuales se han centrado más en la observación de la información del estado del canal. Como resultado de una división de frecuencia ortogonal, la información que aporta indica posiciones más precisas, ya que vienen de una única señal. Esto ha hecho que el indicador más veraz, incluso con interferencias en interiores, sea la información del estado del canal (K. Wu et al., 2012).

Basado en el indicador de fuerza de la señal recibida, se creó un sistema de reconocimiento de gestos usando WiFi (*WiGest*) (Abdelnasser et al., 2015). Con unas tasas de eficiencia que varían entre el 96% y el 87,5%, el sistema es capaz de identificar y mapear varios gestos de la mano, usando tres puntos de acceso aéreo y uno simple. Otros sistemas como pueden ser el *WiSee*, son capaces de reconocer hasta nueve gestos

## 2. Estado del Arte

---

corporales, que interactúan con electrodomésticos conectados a WiFi (Pu et al., 2013). Esto se debe, a que las señales inalámbricas cuando rebotan sobre un objeto, producen el efecto *Doppler*.

En la misma línea introdujo el *WiVi* (Adib y Katabi, 2013), el *Witrack* (Adib et al., 2014) y *Witrack0* (Adib et al., 2015). Un hecho importante a destacar de estos sistemas, es que son capaces de detectar los movimientos humanos, aunque existan paredes en medio. Con respecto a la información del estado del canal (Nandakumar et al., 2014), presentó un sistema de reconocimiento de cuatro gestos con un 91% de éxito en escenarios de visión directa y con un 89% en escenarios con nula visión. El *WiFinger*, es un sistema creado a partir de otros trabajos de investigación (H. Li et al., 2016), capaces de capturar los gestos de los dedos, obteniendo una eficiencia de aproximadamente el 93%.

### 2.2.5 RFID Identificación por Radio Frecuencia

Estos sistemas se basan, en señales de ultra alta frecuencia (UHF). Los lectores comerciales que se usan, son capaces de detectar a poca distancia, etiquetas dentro de un determinado rango. Estas etiquetas son portadas o no por una persona, y funcionan en algunos casos con batería. Se contemplan dos opciones, una activa usando baterías y una pasiva sin ellas. Se usan de forma habitual en las cadenas de suministro y para la detección automática de objetos (Angeles, 2005; Borriello et al., 2004). Además, se han utilizado para aumentar el entorno de información digital (Want et al., 1999), para monitorear actividades humanas en interiores (Buettner et al., 2009) o para detectar interacciones humanas con objetos etiquetados con RFID (Fishkin et al., 2004). Partiendo de las señales UHF, se pueden adquirir la información necesaria para el reconocimiento de gestos, en los que se producen cambios de fase, en el RSSI (Yang et al., 2020).

Una ventaja de utilizar la tecnología RFID pasiva es su bajo coste, pero lógicamente al no contar con suministro propio de energía, los rangos de detección se reducen a centímetros. Debido a este contratiempo inherente, las investigaciones relacionadas con la detección de gestos están muy enfocadas hacia la detección de interacciones entre humanos y objetos o interfaces táctiles (Asadzadeh et al., 2012; Bouchard et al., 2014; Jayatilaka y Ranasinghe, 2017; Zou et al., 2017). Para aplicaciones de trazabilidad como identificación y localización, se suele usar las etiquetas activas debido a que tienen un alcance más largo.

### 2.2.6 Cámaras RGB y RGB-d

Actualmente, las cámaras RGB y RGB-d, son los dispositivos que más se usan para capturar gestos (Premaratne et al., 2010). Estos dispositivos captan la luz en los tres canales que tienen, rojo, verde y azul, en inglés "*red*", "*green*" y "*blue*" o RGB. Basándose en una matriz de píxeles, estos dispositivos organizan la imagen resultante, dando como resultado un algoritmo de visión artificial, que, basándose en el orden y la intensidad del color, proporcionan un significado.

Existen diferentes investigaciones que usan este tipo de dispositivos, por ejemplo, está (Chung et al., 2022), que usando el método de clasificación bidireccional con redes neuronales de memoria a corto plazo (Bi-LSTM), reconoce 100 signos, con una precisión del 98,75%. Por otro lado se encuentra (Abdelrazik et al.,

2023), que usa LSTM junto con una unidad recurrente cerrada (GRU). Previamente, utilizó el *framework MediaPipe* de Google para el proceso visual de los vídeos y sus fotogramas. La precisión alcanzada para 5 palabras del lenguaje de signos egipcio es del 94,95%. Existe otra investigación de (Kothadiya et al., 2022) donde usando LSTM y GRU, reconoció 11 signos del lenguaje de signos indio con una exactitud del 97%. (Obi et al., 2023) utiliza CNN para el reconocimiento 26 letras del lenguaje de signos americano con una precisión del 96,3%. Otra investigación dirigida por (Xie et al., 2023) utiliza una red de fusión de similitud local-temporal de múltiples escalas o (mLTSF-Net), para el reconocimiento de gestos dinámicos, usando como base de datos *RWTH-PHOENIX-Weather-2014* generada por la televisión pública alemana, la cual contiene 6841 frases del lenguaje de signos alemán. En la investigación de (Das, Biswas, et al., 2023) utilizan un clasificador CNN-LSTM, donde reconocen 8 palabras de emergencia del lenguaje de signos indio con un acierto del 94,92%. Existe también sobre el lenguaje de signos árabe estudios como (Aldhahri et al., 2023), donde usa MobileNet como sistema preentrenado basado en CNN, siendo éste un modelo de visión artificial optimizado para dispositivos móviles. Su precisión para 32 signos del alfabeto árabe es del 94,96%. Con respecto al lenguaje de signos bengalí, se encuentra la investigación de (Das, Imtiaz, et al., 2023), que trabaja con un clasificador CNN con sistema aleatorio y un entrenamiento basado en el descenso de gradiente estocástico (SGD), logrando un acierto del 91,67% en el reconocimiento de caracteres y del 97,33% de los dígitos. (K y B.J, 2023) utilizó el método de manos de *MediaPipe* para estimar los puntos de referencia de las manos, obteniendo como salida un fichero de extensión CSV y como clasificador el SVM. Esta investigación obtuvo para 30 gestos del lenguaje de signos indio una precisión del 95%. (Dabwan et al., 2023) utiliza CNN para reconocer 32 signos del lenguaje de signos árabe con una exactitud del 99%.

Las cámaras RGB-d, son una variante de las cámaras tradicionales, donde a través de su cuarto canal son capaces de medir y mostrar la profundidad. De esta forma, determinan la distancia entre el objeto capturado y el foco. Si se quieren generar esas imágenes 3D, es necesario usar alguna de estas técnicas:

### 2.2.6.1 Visión Estereoscópica

En este caso si se cree conseguir dos vistas de una escena, es necesario colocar dos cámaras pareadas. Su procedimiento es muy parecido al que utiliza el ojo humano de forma binocular. Además, hay que realizar una serie de cálculos con ambas imágenes para así obtener esa profundidad. Es necesario además, conocer cómo está calibrado el sistema, teniendo en cuenta los parámetros de distancia focal y separación entre los dos centros de cámaras (Trucco y Verri, 1998).

### 2.2.6.2 Luz Estructurada

Variando espacialmente las coordenadas X e Y, se proyecta un patrón que ilumina de forma activa la escena. Esto hace que se sustituya la segunda cámara usada en la visión estereoscópica, por una fuente de luz (Manteca Fernández, 2018). Por lo tanto, esta técnica no deja de ser una variante del anterior.

### 2.2.6.3 Cámaras TOF o de tiempo de vuelo

Este tipo de cámaras poseen la cualidad estimar de forma directa una estructura 3D, sin necesidad de usar ningún algoritmo de visión artificial. Usando una luz del espectro infrarrojo y modelando su fuente,

## 2. Estado del Arte

---

es capaz de obtener la distancia de la escena, midiendo el tiempo de vuelo de la señal reflejada (Park et al., 2011).

Para diseñar un sistema capaz de reconocer gestos para el control de aplicaciones, suele plantearse las siguientes etapas, (véase la Figura 2.7); recopilación y procesamiento de datos (captura, procesamiento y segmentación de la cámara), modelado del objeto capturado (selección y gestión de las características del objeto), definición de gestos (estáticos y dinámicos), detección de los mismos (separaciones de posturas y detección de movimiento) y por último control de transiciones (máquinas de estados).

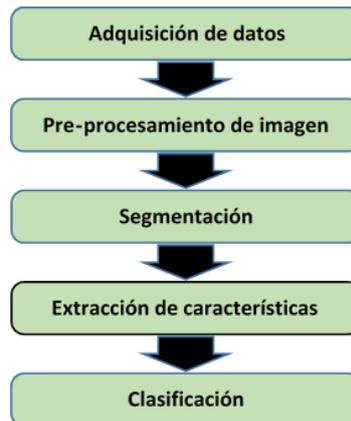


Figura 2.7. Etapas de los Sistemas Basados en la Visión

Existen muchos trabajos de investigación, que han estudiado cada una de las partes de la figura anterior. Algunos de estos trabajos son (X. Chen et al., 2010; Gavrilu y Davis, 1995)

Otros ejemplos destacados son (I. Cohen y Li, 2003), quienes emplearon el algoritmo de máquina de soporte de vectores (SVM) para clasificar las posturas corporales, utilizando un casco visual 3D construido a partir de un conjunto de datos de entrada, devolviendo imágenes en miniatura, que representan determinadas posiciones del cuerpo humano. Un ejemplo de este caso es (Mo et al., 2009), capaz de reconocer determinadas posturas como sentarse, caminar o agacharse. Además, (Corradini et al., 1999) podía analizar y reconocer posiciones mediante el uso de redes neuronales híbridas. Otro sistema que reconoce los movimientos de baile, de al menos seis articulaciones usando una cámara *Kinect*, es (Raptis et al., 2011). Este trabajo de investigación obtuvo una eficiencia del 96,9%.

Estas cámaras se han usado en muchos proyectos de investigación para el reconocimiento de gestos. El problema principal de estos dispositivos, ha sido su coste (Pisharady y Saerbeck, 2015). Una vez apareció en el mercado comercial las cámaras *Kinect* de Microsoft, los costes se hundieron, haciendo que estas cámaras que en un principio fueron de uso comercial y de ocio, se usaran para diferentes trabajos de investigación (Raheja et al., 2015). Una de las características de las *Kinect*, es que son capaces de devolver imágenes en miniatura (*Kinect para Windows - Windows apps*, s. f.).

### 2.2.7 Leap Motion

Uno de los dispositivos más usados en los últimos tiempos es el *Leap Motion*. Este dispositivo económico asequible y al mismo tiempo compacto, tiene la capacidad de capturar el movimiento de los brazos, las manos y los dedos en tiempo real (*Tracking | Leap Motion Controller | Ultraleap, s. f.*), (véase la Figura 2.8).



Figura 2.8. Modo de Funcionamiento del Leap Motion

Existen diferentes estudios, que han comprobado la precisión y robustez de este dispositivo (Guna et al., 2014; Marin et al., 2014; Smeragliuolo et al., 2016; Weichert et al., 2013). Se ha evidenciado que obtiene mejores resultados que otros dispositivos más modernos como la *Kinect* o *Myo*.

De este dispositivo, se han hecho diferentes mediciones, que analizan de forma sistemática su funcionamiento estático y dinámico. De esta forma, no sólo se consigue acotar el espacio de trabajo sino su confiabilidad y precisión. Una de las características estáticas, es que cuando se colocan objetos en ambos extremos, aparece una desviación estándar. En cuanto a los experimentos dinámicos, se observó una caída significativa en la precisión de las muestras, cuando se tomaron a más de 250 milímetros por encima del controlador.

Debido al diseño de *Leap Motion*, orientado principalmente a aplicaciones como la realidad virtual (VR), los juegos o el uso de interfaces más naturales, se pueden encontrar numerosas publicaciones al respecto (Agarwal et al., 2015; Iosa et al., 2015; Noor y Aras, 2015; Zhang y Meruvia-Pastor, 2017).

Todas las características relativas a este dispositivo, ayudan a que el mismo sea apto para investigaciones, alejadas de los juegos y realidad virtual. Este hecho se evidencia por la gran cantidad de publicaciones relacionadas con el reconocimiento de gestos, tanto estáticos como dinámicos que utilizan *Leap Motion*. Con respecto al lenguaje de signos árabe, (Mohandes, Aliyu, et al., 2014) reconoció 28 letras del alfabeto. Ellos usaron redes neuronales multicapa, añadiendo un clasificador Naves Bayes, alcanzando una eficiencia del 99%. La máquina de soporte de vectores fue usada por (Y. Chen et al., 2015), para reconocer 36 gestos entre números y alfabeto, logrando una exactitud del 80%. Ese mismo año, (McCartney et al., 2015), logró una tasa de precisión del 92,4% en una base de datos de más de 100 participantes,

## 2. Estado del Arte

---

utilizando redes neuronales convolucionales y modelos ocultos de Markov. Posteriormente usando redes neuronales y un clasificador *hidden conditional neural fields* o *HCNF*, (Lu et al., 2016), logró capturar diferentes tipos de gestos dinámicos, alcanzando una exactitud del 89,5%. Otro ejemplo de investigación para 30 signos del lenguaje de signos americano, se encuentra en (Saad et al., 2022), donde usando los clasificadores Bi-GRU y Bi-LSTM, obtuvieron una precisión del 96,97% y 97,98% respectivamente. (B. Wu et al., 2022) usando LSTM modificado, fue capaz de reconocer 20 palabras dinámicas del lenguaje de signos chino con una exactitud del 99,55%. Por otro lado está (Abdullahi y Chamnongthai, 2022) que usando un fusor capturador de vectores (FFV) combinado con Bi-LSTM obtuvo un 91% de exactitud, reconociendo 57 palabras del lenguaje de signos americano. En cambio (Katılmış y Karakuzu, 2023), usa el *extreme leaning machine* (ELM) a través de un modelo de meta-aprendizaje de ELM, llamado Meta-ELM, obteniendo para 26 palabras dinámicas del lenguaje de signos turco un acierto del 93%. Otra investigación dirigida por (Nasir et al., 2022), usa dos métodos diferentes de clasificación para el reconocimiento de 26 letras del alfabeto del lenguaje de signos americano, siendo para K-NN del 72,78% y para SVM del 79,83%. Por último está el trabajo de (Rishan et al., 2022), que usando técnicas de coincidencia de plantillas geométricas y procesamiento de lenguaje natural (NLP), logra un resultado del 77% para 3 palabras y 3 letras del lenguaje de signos de Sri Lanka.

Lógicamente, al igual que otras tecnologías explicadas anteriormente, el uso de *Leap Motion* se ha extendido a otras áreas como la robótica (Bassily et al., 2014), uso de reconocimiento de signos para el control de robots militares (Sesli, 2023), la rehabilitación médica (W.-J. Li et al., 2017), la domótica (Gándara y Bauza, 2015), la identificación y autenticación (Chan et al., 2015), la música (Hantrakul y Kaczmarek, 2014) o educación (Ling y Rui, 2016).

### 2.2.8 VR o Realidad Virtual

Hoy en día, los avances en resolución, que se incluyen en los últimos dispositivos basados en VR, ayudan a crear espacios interactivos fascinantes en diversos campos. Además de la alta resolución, hay que sumar el *finger tracking*, las sensaciones o los entornos personalizados. Estos avances benefician a aplicaciones que van dirigidas a la rehabilitación médica, educación, formación profesional y prácticas virtuales de tipo arquitectónico, bancos de pruebas virtuales para diseño e ingeniería industrial, teleoperación de robots (Cichon y Roßmann, 2017), etc. También en el campo del reconocimiento de signos cabe destacar el trabajo de (Leth, 2023), donde usando el NLP es capaz de reconocer 25 signos del lenguaje de señas danés, logrando 44,4% para signos dinámicos y un 41,3% para signos estáticos.

Los esfuerzos se han centrado en crear tanto una conciencia espacial lo más realista y precisa posible, proporcionando a los usuarios una experiencia completa. Por este motivo, existen usuarios que quieren conseguir el *head-mounted display*, este dispositivo es parecido a un casco que permite reproducir imágenes creadas por ordenador sobre una pantalla. Cuando el *head-mounted display* consigue el objetivo de crear esa conciencia espacial, se entiende que se mejora el rendimiento de los movimientos. En (Vosinakis y

Koutsabasis, 2018), muestra la compatibilidad de *Oculus Rift*, (modelo de gafas de realidad virtual), con *Leap Motion*, haciendo que el rendimiento del usuario sea significativamente mejor en entornos de realidad virtual. La dificultad que tienen estos modelos duales, en la complejidad a la hora de adaptar tanto el hardware del sensor como el procesamiento software. Debido a que estas tecnologías son muy diferentes, es necesario la utilización de otros elementos como las cámaras de profundidad, la triangulación de radio, la luz polarizada, seguimiento basado en marcadores de cámara infrarroja u otros. Esto hace, que el entorno de realidad virtual sirva como una meta-capa y de esta forma simula e integra los diferentes sistemas de seguimiento (Jetter et al., 2020). La infinidad de posibilidades que ofrece la realidad virtual ha llevado a empresas como *HTC*, *Oculus*, *Valve*, *HP*, *Microsoft* y *Sony* a fabricar sus propias gafas. Ejemplos sorprendentes son la alianza de *HP*, *Valve* y *Microsoft* para lanzar *HP Reverb G2* (*HP Reverb G2 VR Headset*, s. f.) y de esta forma alcanzar una resolución de 90 Hz, que es actualmente la misma que *Oculus Quest* (*Oculus Quest 2*, s. f.) y *HTC Vive* (*Virtual Reality | HTC España*, s. f.). Además, hay que destacar las gafas *Valve Index* (*Mejora tu experiencia.*, s. f.) con su modo experimental de 144 Hz, aumenta el campo de visión 20° más que *HTC*.

### 2.2.9 UltraLeap

Este dispositivo es relativamente nuevo, salió al mercado en el año 2019. Éste tiene la capacidad de crear la sensación de que estás tocando algo en el aire, mientras manipulas el dispositivo. A través de ondas, el dispositivo genera la sensación táctil de interactuar con un objeto. El principal problema de este dispositivo es su precio, que asciende a unos 4000 euros. Esta dificultad ha hecho que no existan muchas investigaciones que lo usen de forma habitual.

Los estudios con este dispositivo están centrados en varios campos. Los primeros que se realizaron fueron relativos a las señales hápticas o táctiles como en (Fuchs et al., 2022; Georgiou, Martinez, et al., 2022; Rakkolainen et al., 2021; Singhal et al., 2021).

Se han realizado estudios que exploran el uso de la robótica en combinación con gafas de realidad virtual como complemento (Mulot et al., 2021), para crear interfaces en automoción (Brown et al., 2020, 2022; Young et al., 2020). También se ha estudiado su uso para el manejo y automatización del hogar (Van den Bogaert et al., 2021). Otro enfoque publicado, ha sido la creación de una sensación táctil palpando objetos mientras se usan gafas de realidad virtual en (Georgiou, Martinez, et al., 2022).

Por último el uso más social del cual existen estudios, es en el reconocimiento del lenguaje braille para personas ciegas (Paneva et al., 2020).

### 2.2.10 Kinect y Azure Kinect

#### 2.2.10.1 Kinect

La cámara *Kinect* es un dispositivo desarrollado por *Microsoft* para uso comercial y de ocio en consolas de videojuegos. Se empezó a comercializar en el año 2010 y se dejó de fabricar en el año 2017.

## 2. Estado del Arte

---

Gracias a las características técnicas y a su precio, era idóneo para el reconocimiento de movimientos corporales y gestuales.

Existen muchas investigaciones relacionadas con su uso para el reconocimiento de gestos y de signos de diferentes lenguas. Dentro del reconocimiento de gestos forma extensa y sin tener en cuenta la lengua de signos, están (Y. Li, 2012; Patsadu et al., 2012; Ren et al., 2013, 2013; C. Wang et al., 2015; Y. Wang et al., 2012).

En referencia al lenguaje de signos Americano, está (Sun et al., 2015) que trabaja con 2000 frases y reconoce 86% de las mismas usando un máquina de soporte de vectores latente. En (Zafrulla et al., 2011) trabaja con 1000 signos y como clasificador usa modelos ocultos de Markov (HMM), llegando a una precisión del 76,12%. En la investigación de (Yang et al., 2020), trabaja con 24 signos y usa la técnica del campo aleatorio condicional jerárquico, con una precisión del 93,2%. Por otra parte (D. Li et al., 2020) trabaja con 2000 palabras comparando 2 modelos. Uno de estos enfoques es el basado en 2D con poses humanas y el segundo trata de dar un enfoque holístico, logrando una exactitud de hasta el 62,63%.

Con respecto al lenguaje de signos árabe existen varios trabajos relacionados con la *Kinect*. En primer lugar está (Abdel-Samie et al., 2018) que con 30 palabras consigue una precisión entre el 95,25% y el 97,58%. En (Sarhan et al., 2015) trabaja con 16 palabras usando HMM llegando al 80,47%.

En el lenguaje de signos chino existen investigaciones como (Geng et al., 2014), que trabaja con 20 gestos y usa la máquina de soporte de vectores (SVM), con una precisión del 61,46%. Por otro lado está (Chai et al., 2013), que para su investigación usa 239 palabras creando una galería de trayectorias 3D y comparando las mismas para su reconocimiento, con un porcentaje de precisión que oscila entre el 83,51% y el 96,32%. Por último como estudio relevante se encuentra (Liu et al., 2016) que con 500 palabras y usando redes neuronales de memoria a corto plazo (LSTM), logra una precisión del 63,3%.

En otros lenguajes de signos se encuentran estudios como (Lang et al., 2012), que trabaja con el lenguaje de signos alemán con una exactitud del 97%. Por otra parte está (Yauri Vidalón y De Martino, 2016), que han realizado una investigación sobre 107 palabras del lenguaje de signos brasileño relativos a términos médicos. Este último, usa como método de clasificación la deformación dinámica de tiempo (DTW) y redes neuronales con vecino más cercano (KNN), con un porcentaje de acierto entre el 90% y el 98,69% respectivamente. También hay investigaciones como la del lenguaje de signos italiano (Pigou et al., 2015), que trabaja con 20 signos usando redes neuronales convolucionales (CNN) obteniendo un acierto del 91,7%. Dentro del lenguaje de signos indio está (Mehrotra et al., 2015), que usa 37 signos con un acierto del 86,16%. Por último está (G. C. Lee et al., 2016) que usa 25 gestos del lenguaje de signos taiwanés, trabajando modelos ocultos de Markov, combinado con máquina de soporte de vectores con una precisión del 85,14%. Por último comentar el trabajo de (Bora et al., 2023), que usando el *framework* de *MediaPipe* para la extracción de imágenes y un clasificador de redes neuronales, obtiene un acierto del 99% para 9 letras del lenguaje de signos indio.

### 2.2.10.2 Azure Kinect

Este dispositivo salió al mercado en el año 2019, con un enfoque diferente al de la antigua *Kinect*. Su propósito principal es el uso en aplicaciones profesionales, respaldado por la integración de inteligencia artificial para la gestión de sistemas en entornos empresariales (*Azure Kinect DK*, s. f.).

Esta cámara tiene unas características excelentes para la gestión de entornos, pero no tiene la definición necesaria para captar de forma nítida la posición y movimiento de las manos, y mucho menos de los gestos. Existen estudios de investigación como (Brock y Gomez, 2022) que trabajan con 18 gestos usando CNN, reconociéndolos con una precisión del 41%. Al mismo tiempo, esta investigación trabajó con clasificadores de conjuntos apilados (ENS) llegando a un acierto del 70,7%. El resto de estudios publicados tienen resultados de aciertos por debajo del 75%, debido principalmente, a que el dispositivo es incapaz de reconocer los movimientos de los dedos.

## 2.3 Conclusiones

La evolución de los diferentes dispositivos y su aplicación para el reconocimiento de gestos es un tema activo desde los años 60, y sigue siendo una constante hasta la actualidad.

En la actualidad, donde las áreas de conocimiento tienen un carácter divulgativo universal, es normal que aparezcan muchos estudios relacionados con el reconocimiento de gestos, habiendo muchas investigaciones que trabajan de forma exhaustiva esta área.

Desde el *Sketchpad* en los años sesenta, hasta nuestros días, se han realizado importantes investigaciones y grandes esfuerzos para producir dispositivos que sean capaces de reconocer gestos. A partir de 1980, hubo una revolución muy importante en todo lo concerniente a las tecnologías semiconductoras. El desarrollo de sensores que eran capaces de reconocer diferentes elementos, revolucionó todo el campo de dispositivos de reconocimiento. Un ejemplo, ya comentado con anterioridad, son los guantes, *AcceleGlove*, *PowerGlove*, *CyberGlove*, *SuperGlove* y *DataGlove*. Estos dispositivos, comenzaron siendo invasivos y poco versátiles, si bien es cierto que posteriormente han ido mejorando su calidad y precisión a lo largo de los años. No sólo la mejora en el campo de los sensores ha sido fundamental para el desarrollo de dispositivos de reconocimiento, sino también todo el avance en tecnología del software que procesa dicha información.

Por ejemplo, una de las técnicas que han mejorado de forma exponencial a partir de esa década, es el procesamiento de vídeo, que tiene uno de sus principales escollos cuando trabaja en dos dimensiones, considerando que una captura gestos necesita sistemas 3D.

En el año 2010, *Microsoft* sacó al mercado la cámara *Kinect*. Dicha Cámara revolucionó el mundo del ocio, y ayudó a que muchas investigaciones pudiesen trabajar con dicha cámara. En la cámara que tenía un costo asequible y que además pertenecía al conjunto de cámaras ToF, era capaz de captar los movimientos de una persona en tiempo real y visualizar su contorno en la pantalla. El problema de este dispositivo, es que

## 2. Estado del Arte

---

no fue diseñado con el fin de reconocer los movimientos de las manos. Al mismo tiempo tenía un problema añadido, y es que su volumen era un inconveniente para su portabilidad.

En el año 2012 aparece al mercado el Leap Motion. Este dispositivo que está el conjunto de componentes ToF, tiene la capacidad de reconocer el movimiento de las manos, siendo su objetivo principal interpretar gestos manuales. Las investigaciones que han utilizado el *Leap Motion*, probaron la capacidad de este dispositivo en la detección de diferentes tipos de gestos estáticos, con porcentajes de hasta el 99% de precisión. En los trabajos de investigación donde se ha usado este dispositivo, han tenido logros de eficiencia bastante elevados. Este sensor volumétrico, tiene muchas ventajas sobre otros dispositivos. En primer lugar, esta tecnología es muy económica, su uso y transporte también es sencillo, reconoce claramente tus manos y puedes trabajar con él en múltiples plataformas. Una de las ventajas más interesantes que tiene este dispositivo, es que se puede integrar de forma muy sencilla en aparatos electrónicos, por ejemplo, en portátiles. El fabricante también proporciona todo el software necesario, siendo además su alcance, más que suficiente para reconocer, por ejemplo, la lengua de signos.

Las tecnologías comentadas anteriormente, por ejemplo, las de tipo sonora como el WiFi, el ultrasonido y el RFID, tienen problemas a la hora de detectar de forma precisa, los gestos de las manos. Muchas de estas tecnologías se basan principalmente, en cómo procesar el efecto *Doppler*.

Tabla 2.1. Dispositivos tratados en este capítulo

Dispositivo	Autores	Año	Aplicación	Enfoque	Resultados
<i>Sketchpad</i>	Sutherland	1963	Reconocimiento táctil	Primer dispositivo de interacción táctil	Abre un nuevo campo de comunicación entre hombre-máquina
<i>Videoplace</i>	Krueger et al.	1969	Realidad Virtual	Buscar una alternativa a la interacción tradicional hombre-máquina	Combina la imagen de vídeo en vivo de un participante con un mundo gráfico realizado por ordenador
<i>DataGlove</i>	Zimmerman	1983 (Actualmente en venta)	Reconocimiento de gestos	La primera invención de un guante con sensor de flexión óptica	Las señales de estos sensores se pueden procesar para su aplicación en kinesiología, fisioterapia, animación por ordenador, control remoto e interfaz hombre a máquina.
<i>PowerGlove</i>	Mattel	1989 (No se fabrica actualmente)	Realidad Virtual Reconocimiento de gestos	Dispositivo para juegos de Nintendo	Se creó una alternativa de bajo costo para investigadores en realidad virtual y reconocimiento de posturas de manos y gestos
<i>SuperGlove</i>	Nissho Electronics	1997 (No se fabrica actualmente)	Reconocimiento de gestos	Uso pionero en sensores resistivos especiales	Mide la flexión de las articulaciones metacarpofalángicas e interfalángicas proximales de los cuatro dedos y el pulgar
<i>CyberGlove</i>	CyberGlove Systems	1990 (Actualmente en venta)	Realidad Virtual Control robótico rehabilitación médica Reconocimiento de gestos	Guantes que utilizan sensores magnéticos y acelerómetros	Guantes programables con mayor impacto comercial
<i>AcceleGlove</i>	AnthroTronix	1999 (Actualmente en venta)	Realidad virtual Control robótico rehabilitación médica Reconocimiento de gestos	Guantes que utilizan sensores magnéticos y acelerómetros	Guantes programables con mayor impacto comercial
<i>Myo</i>	Thalmic Labs	2014	Realidad virtual	Pulsera que lee y recoge la actividad eléctrica de los	Pulsera mioeléctrica con mayor impacto comercial

## 2. Estado del Arte

Dispositivo	Autores	Año	Aplicación	Enfoque	Resultados
		(Actualmente en venta)	Reconocimiento de gestos Automatización del hogar	músculos para controlar, de forma inalámbrica, otros dispositivos con gestos y movimientos de brazos	
<i>Kinect</i>	Microsoft	2010-2017 (No se fabrica actualmente)	Juego Reconocimiento de gestos Control robótico rehabilitación médica Educación Identificación Autenticación	Permite a los usuarios controlar e interactuar con la consola, sin tener contacto físico con un controlador de videojuegos tradicional	Más de 10 millones de dispositivos vendidos
<i>Leap Motion</i>	UltraLeap Inc.	2012 (Actualmente en venta)	Realidad virtual Reconocimiento de gestos Control robótico rehabilitación médica Educación Identificación Autenticación	Dispositivo de bajo costo para interactuar con el ordenador a través de gestos	Startup cuya valoración alcanzó los 306 millones en 2013 y fue adquirida por 30 millones de dólares
<i>Azure Kinect</i>	Microsoft	2019 (Actualmente en venta)	En sistemas profesionales, (gestión de almacenes, de personas, ...)	Dispositivo de alto coste, con un gran desarrollo de Inteligencia Artificial para la detección de objetos y personas	Tiene muy buenos resultados en el tratamiento de imágenes y desarrollo en Inteligencia Artificial
<i>UltraLeap</i>	UltraLeap Inc.	2020 (Actualmente en venta)	En sistemas de detección de la posición de la mano de forma háptica	Dispositivo excesivamente caro y poco software desarrollado para su uso	Cada vez se vende menos debido a su precio y las limitadas posibilidades que posee

Tabla 2.2. Resumen de todas las investigaciones comentadas en este capítulo.

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(Kaushik y Jain, 2014; LaViola, 1999; Myers, 1998; Premaratne et al., 2010; Sharma y Verma, 2015)		Descripción general de la interacción basada en gestos	No aplicado	No aplicado	No aplicado	Ventajas y desventajas de varias técnicas que han surgido a lo largo del tiempo y en progreso.
(Sturman y Zeltzer, 1994)	Guantes de datos	Estudio de entrada de datos	No aplicado	No aplicado	No aplicado	Alcanza el objetivo de ser una interfaz de usuario "sin contacto" requiere de un progreso en muchas áreas, incluida la mejora de la velocidad y la precisión de los dispositivos de seguimiento y la reducción de los costos de fabricación.
(Santos et al., 2015)	Guantes de datos	Guantes para uso médico	No aplicado	No aplicado	No aplicado	El uso de guantes quirúrgicos sensorizados en cirugía laparoscópica asistida por la mano, proporciona al cirujano información adicional sobre tejidos y órganos
(Moreno et al., 2013)	Guantes de datos	Diseño de un guante de control háptico para manipulador robótico teleoperado	No aplicado	No aplicado	No aplicado	Eficiencia superior al 90% en los movimientos de apertura y cierre
(Healey y Picard, 1998)	Electromiografía EMG	Control cibernético de ordenadores portátiles	No aplicado	Señales de conductividad de la piel	Sensor de conductancia de la piel con un filtro combinado	El algoritmo de detección de sobresaltos es robusto en varios usuarios y permite que el dispositivo portátil responda automáticamente a eventos de interés potencial
(Dubost y Tanaka, 2002; Fistre y Tanaka, 2002; Knapp y Lusted,	Electromiografía EMG	Una interfaz basada en electrodos EMG activos	No aplicado	Filtrado de cálculos de potencia eficaz	No aplicado	Control de sintetizadores, secuenciadores, cajas de ritmos o cualquier otro dispositivo MIDI (Interfaz Digital de Instrumentos

## 2. Estado del Arte

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
1990; Tanaka y Knapp, 2002)						Musicales).(Dubost y Tanaka, 2002; Fistre y Tanaka, 2002; Knapp y Lusted, 1990). Ampliado a una interacción multimodal. (Tanaka y Knapp, 2002).
(Rosenberg, 1998)	Electromiografía EMG	Puntero gráfico de retroalimentación biológica (electromiogramas)	No aplicado	No aplicado	Redes neuronales	Detecta los electromiogramas de cuatro de los músculos utilizados para mover la muñeca, con un rendimiento similar al de un ratón estándar
(Dhillon et al., 2009; Wheeler y Jorgensen, 2003)	Electromiografía EMG	Captura de señales electromiográficas usando como entrada joysticks y teclados virtuales	No aplicado	No aplicado	No aplicado	Interfaz cerebro-computadora de bajo costo para personas discapacitadas y personas mayores
(Saponas et al., 2009)	Electromiografía EMG	Interpretación de electromiografía de antebrazo y clasificación de gestos de los dedos	31 segmentos temporales/milisegundo	Coherencia de fase. Media cuadrática (RMS). Transformada de Fourier (FFT).	Máquina de vectores de soporte (SVM)	Precisión de los gestos de los dedos y las posturas de los brazos por encima del 80 %
(Tabor et al., 2016)	Electromiografía EMG	Prótesis de control mioeléctrico y juegos terapéuticos con <i>Myo Armband</i>	No aplicado	No aplicado	No aplicado	Ayuda a mejorar la baja tasa de adopción de prótesis mioeléctricas
(Boyalı et al., 2015)	Electromiografía EMG	Postura de la mano y reconocimiento de gestos con <i>Myo Armband</i>	Agrupamiento subespacial ordenado (OSC)	Transformada de Fourier (FFT)	Clasificación de representación colaborativa (CRC)	Precisión superior al 97% en seis gestos con las manos y varias posturas
(Sathiyarayanan y Rajan, 2016b)	Electromiografía EMG	Proporcionar pautas sobre el uso de <i>Myo Armband</i> para el análisis de fisioterapia por parte	No aplicado	No aplicado	No aplicado	El dispositivo MYO, tiene la capacidad de poder ser utilizado para comprender los movimientos del brazo durante la etapa de fisioterapia

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(Abreu et al., 2016)	Electromiografía EMG	de médicos y pacientes. Uso de datos (EMG) proporcionados por el brazalet Myo para identificar letras del lenguaje de signos brasileño	0.25 muestras/segundo	Cálculo del valor absoluto y medio	Máquina de vectores de soporte (SVM)	Es muy difícil percibir en tiempo real gestos finos con los dedos solo con datos EMG
(Benalcázar, Jaramillo, et al., 2017; Benalcázar, Motoche, et al., 2017)	Electromiografía EMG	Reconocimiento de gestos de mano en tiempo real con pulsera Myo	200 muestras/fotograma	Rectificación de señal (ABS) y aplicación de un filtro digital de paso bajo Butterworth	Deformación dinámica del tiempo (DTW) y redes neuronales vecino más cercano (K-NN)	En (Benalcázar, Jaramillo, et al., 2017) el modelo funciona mejor (86% de precisión) que el sistema Myo (83%) para 5 tipos de gestos. En (Benalcázar, Motoche, et al., 2017) se incluye un detector de actividad muscular que acelera el tiempo de procesamiento y mejora la precisión del reconocimiento en un 89,5%
(Singh y Chaturvedi, 2023)	Electromiografía EMG	Reconocimiento de gestos de mano en tiempo real con pulsera Myo	Frecuencia a 200Hz durante 5 segundos/por signo	ANOVA, Chi-square, <i>Mutual Info</i> , ReliefF	Algoritmo catboost	24 letras del alfabeto del lenguaje de signos americano con una precisión del 99,91%
(X. Chen et al., 2010; Sikdar et al., 2014; Zheng et al., 2006)	Ultrasonido	Imágenes de ultrasonido basadas en el control para prótesis de brazo superior	Secuencias de imágenes de ultrasonido	Algoritmo de procesamiento de imágenes. Realiza un cálculo sobre los coeficientes de correlación cruzada entre varias secuencias. RMS (Media cuadrática)	Redes neuronales vecino más cercano (K-NN) (Sikdar et al., 2014b) No aplicado (X. Chen et al., 2010a; Zheng et al., 2006)	La señal de sonomiografía tiene potencial para ser un método alternativo para el control protésico

## 2. Estado del Arte

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(Mujibiya et al., 2013)	Ultrasonido	Detección de gestos basada en la propagación de ultrasonidos transdérmicos de baja frecuencia.	35, 40, 45, y 50 kHz de ondas sinusoidales	Amplitud promedio, desviación estándar, promedio lineal y logarítmico de puntos de la transformada rápida de Fourier	Máquina de vectores de soporte (SVM)	Clasificación sólida de gestos táctiles realizados en diferentes lugares del antebrazo
(Hettiarachchi et al., 2015)	Ultrasonido	Sistema de detección de actividad muscular radial con transductores de ultrasonido	1100 imágenes/gesto	Rectificación y aplicación de la transformada de Hilbert	Redes neuronales (NN), aplicando la correlación combinada cruzada	72% de éxito en el reconocimiento de cinco gestos de flexión/extensión
(McIntosh et al., 2017)	Ultrasonido	Comparación del rendimiento de diferentes posiciones de montaje de un dispositivo de ultrasonido portátil	Dividir el vídeo en flexiones o extensiones desde una posición neutra	Algoritmo de Farnebäck	Flujo óptico. Regresor de percepción multicapa.	Clasificación de flexión y extensión de 10 gestos manuales discretos con una precisión superior al 98%
(Tivive et al., 2010)	Ultrasonido	Un método de clasificación de la marcha humana basado en espectrogramas de radar <i>Doppler</i>	20 muestras/segundo	STFT (Transformada de Fourier de tiempo reducido)	Máquina de vectores de soporte (SVM)	Con siete filtros direccionales, el rendimiento de clasificación se incrementa por encima del 98%
(Biber et al., 1980)	Ultrasonido	Ondas de sonido ultrasónicas para medir la distancia a un sujeto	No aplicado	No aplicado	No aplicado	Convertir la información en rotación de la lente para obtener el enfoque correcto
(Kalgaonkar y Raj, 2009)	Ultrasonido	Dispositivo ultrasónico para reconocer gestos manuales simples	Muestras de 32 ms	Transformada rápida de Fourier y transformada discreta del coseno	Modelo de mezcla gaussiana (GMM)	88,42% de acierto en el reconocimiento de 8 gestos

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(Fléty, 2000)	Ultrasonido	Síntesis de sonido controlada por gestos con método de triangulación mediante ultrasonidos	No aplicado	Filtro paso-bajo	No aplicado	Éxito en la detección y seguimiento de movimiento en espacios de 1 m <sup>3</sup>
(Kreczmer, 2011)	Ultrasonido	Reconocimiento de gestos basado en la aplicación de telémetros ultrasónicos simples en un robot móvil	No aplicado	La velocidad y la dirección de la aleación de la mano definen una función de una relación de velocidad	No aplicado	Sensores fácilmente disponibles pero el conjunto de gestos reconocidos es limitado
(Booij y Welle, 2014; Gupta et al., 2012)	Ultrasonido	Técnica que aprovecha el altavoz y el micrófono ya integrados en el dispositivo móvil para detectar gestos en el aire	No aplicado	Velocidad, dirección, proximidad, tamaño del objetivo y variación de tiempo	No aplicado	Robustez superior al 90% en diferentes dispositivos, usuarios y entornos (Gupta et al., 2012b)
(Przybyla et al., 2012, 2014)	Ultrasonido	Sensor de profundidad ultrasónico 2D que mide el alcance y la dirección de los objetivos en el espacio (Przybyla et al., 2012b). Para aplicaciones 3D (Przybyla et al., 2014b)	No aplicado	Relación señal-ruido y ángulo de incidencia	No aplicado	Las soluciones ultrasónicas son ideales para el reconocimiento de gestos simples para teléfonos inteligentes y otros dispositivos con limitación de energía

## 2. Estado del Arte

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(Saad et al., 2018)	Ultrasonido	Reconocimiento ultrasónico de gestos manuales para dispositivos móviles	No aplicado	Redes neuronales RTD (Tiempo de ida y vuelta) y RSS (Intensidad de la señal reflejada)	Máquina de vectores de soporte (SVM)	96% de éxito para siete tipos de gestos
(Sang et al., 2018)	Ultrasonido	Sistema y métodos de reconocimiento de micro gestos manuales que utilizan detección activa ultrasónica	256 puntos de la Transformada rápida de Fourier	Sistema <i>Doppler</i> de procesamiento de señales de radar pulsado, desplazamiento y velocidad	Modelos ocultos de Markov (HMM) Redes neuronales de extremo a extremo ( <i>end-to-end</i> )	96,32% de éxito en 11 gestos manuales
(Kellogg et al., 2014)	WiFi	Sistema de reconocimiento de gestos orientado a la operación en una gama de dispositivos informáticos de bajo consumo a través de señales inalámbricas existentes	No aplicado	Detector de envolvente para eliminar la frecuencia portadora y extraer información de amplitud	Deformación dinámica del tiempo (DTW)	Logra clasificar con un 97% de acierto un conjunto de ocho gestos
(Pu et al., 2013)	WiFi	Reconocimiento de gestos extrayendo cambios <i>Doppler</i> de señales inalámbricas	Combinación de segmentos con desplazamiento <i>Doppler</i> positivo y negativo	Cálculo de la distribución de energía frecuencia-tiempo de la señal de banda estrecha	Codificación de gestos. Efecto <i>Doppler</i>	Para una persona en una habitación es capaz de reconocer 9 gestos alcanzado un 94% de eficiencia. Con más gente baja al 60%
(Adib y Katabi, 2013)	WiFi	Capturar objetos y gestos en movimiento tras una pared	Decodificación de gestos	Funciones de distribución acumulativa (CDF) de Relación señal-ruido (SNR)	No aplicado	87,5% de éxito en la detección de tres gestos en paredes de hormigón
(Abdelnasser et al., 2015)	WiFi	Gesto basado en WFI sistema de	Detecta cambios en el	Transformada discreta de Wavelet (DWT)	No aplicado	96% de precisión usando tres puntos de acceso

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
		reconocimiento que identifica diferentes primitivas de cambio de señal	indicador de intensidad de la señal recibida			
(Nandakumar et al., 2014)	WiFi	Reconocimiento de gestos WFI en dispositivos existentes	Indicador de intensidad de la señal recibida y cambio en el estado de la información	Filtros de paso-bajo. Cambios en el indicador de intensidad de la señal recibida y en el estado de la información	Información de altura y tiempo para realizar la clasificación	91% de precisión al clasificar cuatro gestos entre seis participantes
(He et al., 2015)	WiFi	Reconocimiento de gestos WFI en dispositivos existentes	180 grupos de datos de cambio en el estado de la información de cada paquete	Valor medio, desviación estándar, desviación mediana, desviación absoluta y valor máximo de los patrones de anomalía.	Máquina de vectores de soporte (SVM)	Precisión promedio de un 92% de reconocimiento en el escenario con línea de visión y de un 88% sin línea de visión
(H. Li et al., 2016)	WiFi	Reconocimiento de gestos con los dedos como entrada de texto continuo en dispositivos WFI estándar	Cambio en el estado de la información	Transformada discreta de Wavelet (DWT)	Deformación dinámica del tiempo (DTW)	Precisión promedio de clasificación de un 90.4% para reconocer gestos de dedos de 9 dígitos del lenguaje de signos americano
(Zou et al., 2017)	RFID (Identificación por radiofrecuencia)	Sistema de reconocimiento de gestos sin dispositivo basado en la salida de información de fase por dispositivos RFID comerciales	Método Varri modificado	Filtro Savitzky-Golay	Deformación dinámica del tiempo (DTW)	Logra una precisión de promedio de reconocimiento entre el 96.5% y 92.8% en el escenario de posiciones idénticas y posiciones diversas respectivamente
(Raptis et al., 2011b)	Visión	Clasificación en tiempo real de gestos de baile de	Fotogramas de vídeo	Algoritmo de seguimiento esquelético (STA)	Clasificador en cascada	Precisión promedio del 96,9% para grabaciones de movimiento

## 2. Estado del Arte

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
		animación de esqueleto usando <i>Kinect</i>				esquelético de aproximadamente 4 segundos
(Pisharady y Saerbeck, 2015)	Visión	Una revisión de los algoritmos de reconocimiento de gestos manuales basados en la visión	No aplicado	No aplicado	No aplicado	Los métodos que utilizan cámaras RGB y RGB-D se revisan con comparaciones de algoritmos cuantitativos y cualitativos.
(Raheja et al., 2015)	Visión	Reconocimiento de gestos robusto con <i>Kinect</i> a través de una comparación entre DTW (Deformación dinámica del tiempo) y HMM (Modelos ocultos de Markov)	Fotogramas de vídeo	No aplicado	Deformación dinámica del tiempo (DTW) y Modelos ocultos de Markov (HMM)	Deformación dinámica del tiempo y Modelos ocultos de Markov con ambos sistemas dan una tasa de clasificación de alrededor del 90%
(Chung et al., 2022)	Visión	Reconocimiento del lenguaje de signos chino	1200 muestras de vídeo por signo	No aplicado	Método de clasificación bidireccional con redes neuronales de memoria a corto plazo (Bi-LSTM)	Reconocimiento de 100 signos con una precisión del 98,75%
(Abdelrazik et al., 2023)	Visión	Reconocimiento del lenguaje de signos egipcio	No aplicado	No aplicado	Memoria de corto plazo (LSTM). Unidad recurrente cerrada (GRU)	Reconocimiento de 10 palabras con 5 personas. Con LSTM 90,97% y con GRU 91,23%
(Kothadiya et al., 2022)	Visión	Reconocimiento del lenguaje de signos indio	Las entradas fueron 2048 secuencias	No aplicado	Combina memoria de corto plazo	Reconocimiento de 11 signos con un acierto del 97%

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
			dimensionales de 40 veces por paso		(LSTM) con la unidad recurrente cerrada (GRU)	
(Obi et al., 2023)	Visión	Reconocimiento del lenguaje de signos americano	30526 imágenes que se dividen en 27 clases para datos de entrenamiento y 8958 imágenes que se dividen en 27 clases para datos de prueba	No aplicado	Redes convolucionales neuronales (CNN)	26 letras del lenguaje de signos americano con una tasa de éxito del 96,3%
(Xie et al., 2023)	Visión	Reconocimiento del lenguaje de signos alemán	Utilización del <i>RWTH-PHOENIX-Weather-2014</i> Todos los videos han sido preprocesados a una resolución de 25 cuadros por segundo (FPS)	No aplicado	Red de fusión de similitud local-temporal de múltiples escalas (mLTSF-Net)	Creación de un modelo de tratamiento de vídeo para el reconocimiento de gestos
(Das, Biswas, et al., 2023)	Visión	Reconocimiento de palabras de emergencia del lenguaje de signos indio	El conjunto de datos contiene un total de 824 videoclips, con 52 ejemplos en cada clase excepto "Lose"	No aplicado	Utiliza una combinación de redes convolucionales neuronales (RNN) con memoria de	Los 8 signos dieron una exactitud de 94,42%

## 2. Estado del Arte

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
			y "Thief", cada uno de los cuales contiene 50 muestras.		corto plazo (LSTM)	
(Aldhahri et al., 2023)	Visión	Reconocimiento del alfabeto del lenguaje de signos árabe	54.049 imágenes divididas en 32 clases	No aplicado	<i>MobileNet</i> es un modelo preentrenado basado en redes neuronales convolucionales (CNN)	32 signos del alfabeto del lenguaje de signos árabe con una exactitud del 94,46%
(Das, Imtiaz, et al., 2023)	Visión	Reconocimiento del lenguaje de signos bengalí	1800 imágenes de signos de caracteres.	No aplicado	Entrenamiento con descenso de gradiente estocástico (SGD) y clasificador con redes neuronales convolucionales (CNN) combinado con bosque aleatorio (RF)	La precisión con 26 caracteres es de 91,67%. La precisión de con 10 dígitos es del 97,33%
(K y B.J., 2023)	Visión	Reconocimiento del lenguaje de signos indio	100 muestras por cada signo	Se usa <i>MediaPipe</i> para generar fichero de extensión CSV	Máquina de soporte de vectores (SVM)	Reconocimiento de 30 signos con una precisión del 95%
(Dabwan et al., 2023)	Visión	Reconocimiento del lenguaje de signos árabe	506 imágenes por cada signo	Modelo con escala <i>EfficientnetB1</i> cargado y pre-entrenados en el modo <i>ImageNet</i>	Redes neuronales convolucionales (CNN)	Reconocimiento de 32 signos con una precisión del 99%

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(Agarwal et al., 2015)	<i>Leap Motion</i>	Segmentación y reconocimiento de texto escrito en 3D mediante <i>Leap Motion Interface</i>	Diferenciación parcial de la señal dentro de una ventana predefinida	Se aplica la heurística para determinar los límites de las palabras.	Modelos ocultos de Markov (HMM)	El algoritmo de segmentación de palabras basado en heurística propuesto funciona con una precisión de hasta el 80,3% y las palabras basadas en HMM (modelos ocultos de Markov) han registrado una precisión del 77,6%.
(Mohandes, Deriche, et al., 2014)	<i>Leap Motion</i>	Reconocimiento del lenguaje de signos árabe usando el <i>Leap Motion Controller</i>	Muestras de 10 fotogramas/segundo	Valor medio	Clasificador Nave Bayes (NB) y and Redes neuronales de perceptrón multicapa (NLP)	Precisión del 98% con el clasificador Naive Bayes y más del 99% con el MLP (Redes neuronales de perceptrón multicapa)
(Y. Chen et al., 2015; Lu et al., 2016; McCartney et al., 2015)	<i>Leap Motion</i>	Reconocimiento de gestos dinámicos con las manos usando <i>Leap Motion</i>	Muestra de datos	Dirección de la palma, palma normal, posiciones de las yemas de los dedos y posición del centro de la palma. Filtrado de ruido de alta frecuencia.	Máquina de vectores de soporte (SVM) y Modelos ocultos de Markov (HMM) (Y. Chen et al., 2015b). Redes neuronales convolucionales (CNN) (McCartney et al., 2015). Campos neuronales condicionales ocultos (HCNF) (Lu et al., 2016b)	Con el 80% de los datos de la muestra, la eficiencia que se logra está por encima del 93,3% (Y. Chen et al., 2015b). Es incapaz de reconocer gestos en tiempo real aunque su eficiencia es alta (McCartney et al., 2015). 95,0% de reconocimiento para el conjunto de datos <i>Handicraft-Gesture</i> y 89,5% para <i>LeapMotion-Gesture3D</i> (Lu et al., 2016b).

## 2. Estado del Arte

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(W.-J. Li et al., 2017)	<i>Leap Motion</i>	Reconocimiento de gestos con las manos para la rehabilitación posterior a un ictus mediante el movimiento de salto	Muestra de datos	Se extraen 17 características, incluidas las distancias euclidianas entre las yemas de los dedos, el ángulo de cabeceo de la palma, el ángulo de giro de la palma, el ángulo de balanceo de la palma y los ángulos entre los dedos	Máquina de vectores de soporte multiclase (SVM) Redes neuronales aplicando el vecino más cercano (k-NN)	Siete gestos para la rehabilitación residencial de pacientes post accidente cardiovascular con una precisión del 97,71%
(Gándara y Bauza, 2015)	<i>Leap Motion</i>	<i>Leap Motion</i> Controller para la autenticación a través de la geometría de la mano y los gestos	Muestra de datos	Tasa igual de error (ERR), Falsa aceptación (FAR) y Tasa falsa de rechazo (FRR)	WEKA (Entorno de Waikato para el análisis del conocimiento)	TI muestra que <i>Leap Motion</i> puede usarse con éxito tanto para autenticar a los usuarios al iniciar sesión como al realizar actividades continuas. La precisión de la autenticación alcanza el 98%
(Marin et al., 2014)	<i>Leap Motion</i> y <i>Kinect</i>	Reconocimiento de gestos de mano con dispositivos <i>Kinect</i> y de movimiento de salto	Muestra de datos	Extracto de las características de los dispositivos	Máquina de vectores de soporte (SVM)	81% de precisión con <i>Leap Motion</i> , 65% de precisión con <i>Kinect</i> y, cuando se combinan, aumenta al 91,3% con diez gestos estáticos
(Saad et al., 2022)	<i>Leap Motion</i>	Reconocimiento del lenguaje de signos americano	No aplicado	No aplicado	Usa 2 clasificadores diferentes bidireccional con unidad recurrente cerrada (Bi-GRU) y bidireccional de memoria de corto plazo (Bi-LSTM)	El sistema trabaja con 30 signos tanto estáticos como dinámicos. Para (Bi-GRU) la precisión es del 96,97% y para (Bi-LSTM) del 97,28%

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(B. Wu et al., 2022)	<i>Leap Motion</i>	Reconocimiento del lenguaje de signos chino	No aplicado	No aplicado	Memoria de corto plazo modificado (LSTM)	Reconocimiento de 20 palabras dinámicas con una exactitud del 99,55%
(Abdullahi y Chamnongthai, 2022)	<i>Leap Motion</i>	Reconocimiento del lenguaje de signos americano	10 muestras por cada signo	No aplicado	fusor capturador de vectores (FFV) combinado con bidireccional memoria de corto plazo (Bi-LSTM)	Reconocimiento de 57 palabras con una precisión del 91%
(Katılmış y Karakuzu, 2023)	<i>Leap Motion</i>	Reconocimiento del lenguaje de señas turco	6 participantes con 15 repeticiones cada palabra	No aplicado	Máquina extrema de aprendizaje (ELM)	Reconoce 26 palabras dinámicas con una exactitud del 93%
(Sesli, 2023)	<i>Leap Motion</i>	Reconocimiento de gestos para robots militares	No aplicado	No aplicado	Función de distribución acumulativa (CDF) basada en redes neuronales profundas (DNN)	La precisión de 8 gestos fue del 88,44%
(Nasir et al., 2022)	<i>Leap Motion</i>	Reconocimiento de 26 letras del lenguaje de signos americano	Las muestras fueron realizadas por 2 personas distintas un de ellas sorda	No aplicado	Vecino más cercano (K-NN). Máquina de soporte de vectores (SVM)	La investigación usa validación cruzada para ambos métodos de clasificación, obteniendo una precisión del 72,78% con K-NN y un 79,83% con SVM

## 2. Estado del Arte

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(Rishan et al., 2022)	<i>Leap Motion</i> con procesamiento del lenguaje natural (NLP)	Reconocimiento del lenguaje de signos de Sri Lanka	Se realizaron muestras de 3 palabras y de 3 letras	No aplicado	Coincidencia de plantillas geométricas y procesamiento de lenguaje natural (NLP)	Para 3 palabras tuvo un acierto del 80% y un 77% para 3 letras
(Mehrabian, 2017)	<i>Leap Motion</i>	Grado de impacto del gesto en una conversación	No aplicado	No aplicado	No aplicado	55% de impacto en conversaciones emocionales
(Ahsan et al., 2009; Barreto et al., 1999; Benedek et al., 2005; Coleman, 2001; Crawford et al., 2005; Farina et al., 2016; Guerreiro y Jorge, 2006; Hefftner et al., 1988; Joshi et al., 2013; Kiguchi y Hayashi, 2012; Luca, 1997; Manabe et al., 2003; Merletti et al., 2004; Merletti y Farina, 2016; Reaz et al., 2006; Valentinuzzi, 2007)	Datos de electromiograma (EMG)	Técnicas de análisis de señales EMG	Descomposición de EMG utilizando la optimización cuadrática media (LMS) no lineal más baja de acumuladores de orden superior	Transformada de Wavelet (DWT), Modelo autoregresivo (AR) Inteligencia artificial Modelos lineales y estocásticos; Modelado no lineal	Distancia euclidiana entre las formas de onda potenciales de la acción de la unidad motora (Luca, 1997). Estimación del número de unidades motoras. Modelos de hardware (Guerreiro y Jorge, 2006; Kiguchi y Hayashi, 2012)	Uso de técnicas basadas en señales de EMG. La mayoría de las investigaciones se encuentran en el ámbito de la biomedicina, implementación hardware, usuario final y diagnóstico clínico.
(Gowitzke y Milner, 1999)	Datos de electromiograma (EMG)	Base científica del cuerpo y sus movimientos.	No aplicado	No aplicado	No aplicado	Organización global de los elementos del sistema neuromuscular, neurorreceptores e instrumentación

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(Webster, 2009)	Datos de electromiograma (EMG)	Conceptos de instrumentación médica	No aplicado	No aplicado	No aplicado	Una vista de los principales sensores médicos
(Guo et al., 2008; Hodges et al., 2003; Zheng et al., 2006)	Datos de electromiograma (EMG)	Relaciones entre las variables medidas con ecografía y EMG	No aplicado	La relación entre cada uno de los parámetros de ultrasonido y EMG se describió con regresión no lineal (Hodges et al., 2003). Y con raíz cuadrática media en (Guo et al., 2008)	No aplicado	La sonomiografía tiene un gran potencial por ser un método alternativo para evaluar la función muscular
(Tahmoush y Silvius, 2009)	WiFi	Identificación de múltiples sujetos humanos al aire libre con señales micro- <i>Doppler</i>	No aplicado	Filtrado <i>Doppler</i> para eliminar el desorden, extracción del torso del espectrograma, filtrado del torso para reducir el ruido y extracción del período pico usando una transformada de Fourier.	Trata de comparar clasificadores para futuros trabajos de investigación.	Para el reconocimiento de personas, es capaz de alcanzar una eficiencia del 80%, tanto en largo alcance como en vista frontal. Si existen variaciones angulares, baja en el peor de los casos, hasta el 40%.
(Abdelnasser et al., 2015; Adib et al., 2014, 2015; Adib y Katabi, 2013; Pu et al., 2013; K. Wu et al., 2012)	WiFi	Sistema que rastrea el movimiento 3D de un usuario a partir de las señales de radio reflejadas en su cuerpo	Espectrograma de distancia de ida y vuelta de cada antena receptora	Filtro Kalman	No aplicado	Su precisión supera la ubicación actual de los sistemas de RF, que requieren que el usuario tenga un transceptor (Adib et al., 2014). La segunda versión (Adib et al., 2015) puede localizar hasta cinco usuarios simultáneamente, incluso si están estáticos. Precisión del 95% o superior
(K. Wu et al., 2012)	Datos de electromiograma (EMG)	Sistema de posicionamiento en interiores basado en WFI que aprovecha la	No aplicado	Procesamiento del estado del canal	No aplicado	Mejora significativamente la precisión de la ubicación en comparación con el uso del indicador de intensidad de la señal recibida

## 2. Estado del Arte

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
		información del estado del canal				
(Angeles, 2005; Borriello et al., 2004; Buettner et al., 2009; Fishkin et al., 2004; Want et al., 1999; Yang et al., 2020)	RFID	Aplicaciones de la cadena de suministro RFID	No aplicado	No aplicado	No aplicado	Directrices de gestión para implementar proactivamente aplicaciones RFID
(Bouchard et al., 2014)	RFID	Contenedor de fluidos con RFID sin batería para reconocer instancias individuales de consumo de líquidos	La serie temporal se segmenta como una sola lectura de etiqueta	Media, mediana, moda, desviación estándar y rango con respecto al indicador de intensidad y fase de la señal recibida	Naïve Bayes (NB). Máquina de vectores de soporte (SVM). Campos aleatorios condicionales lineales LCRF. Bosque aleatorio (RF).	87% de éxito en reconocer episodios de consumo de alcohol en diez voluntarios
(Trucco y Verri, 1998)	Visión	Técnicas introductorias a la visión artificial 3D	No aplicado	No aplicado	No aplicado	Conjunto de técnicas computacionales para imágenes 3D
(Manteca Fernández, 2018)	Visión	Generación de modelos 3D por luz estructurada	No aplicado	No aplicado	No aplicado	Creación de un modelo de escáner 3D basado en luz estructurada
(Park et al., 2011)	Visión	Muestreo de alta calidad en mapas de profundidad capturados desde una cámara 3D-ToF (Tiempo de vuelo) junto con una	No aplicado	No aplicado	No aplicado	El nuevo método supera los enfoques existentes para el muestreo en 3D-ToF

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
		cámara RGB de alta resolución				
(C. Wu y Aghajan, 2007)	WiFi	Estimación de la postura humana basada en modelos para el análisis de gestos en una red de cámaras inteligentes de fusión oportunista	Fotogramas de vídeo	Maximización de expectativas (EM)	No aplicado	Se describe la estimación de la postura humana incorporando el concepto de un marco de fusión oportunista
(I. Cohen y Li, 2003; Corradini et al., 1999; Gavrila y Davis, 1995; Mo et al., 2009)	Cámaras TOF	Modelo de rasgos humanos 2D (Corradini et al., 1999; Mo et al., 2009b) Seguimiento basado en modelos 3D y reconocimiento del movimiento humano a partir de imágenes reales (M. W. Cohen et al., 2018; Gavrila y Davis, 1995)	Fotogramas de vídeo	Modelo 3 DOF (libertad de movimientos). Descriptores de forma (Gavrila y Davis, 1995). Esqueleto de estrella en (I. Cohen y Li, 2003). Procesamiento del color de la piel	Deformación dinámica del tiempo (DTW) en (Gavrila y Davis, 1995) Máquina de vectores de soporte (SVM) en (I. Cohen y Li, 2003; Mo et al., 2009). Clasificadores neuronales (Corradini et al., 1999).	El experimento logra obtener un resultado de recuperación de poses 3D y clasificación de movimientos en (I. Cohen y Li, 2003; Gavrila y Davis, 1995). Resultados con modelos 2D en (Corradini et al., 1999; Mo et al., 2009b)
(Guna et al., 2014; Montalvo Martínez, 2010; Weichert et al., 2013)	Leap Motion	Una evaluación del rendimiento del <i>Leap Motion Controller</i> (Guna et al., 2014). El análisis se realizó de acuerdo con la norma ISO 9283	No aplicado	No aplicado	No aplicado	En el escenario estático se demostró que la desviación estándar fue menor a 0,5 mm en todo momento. Sin embargo, hay una caída significativa en la precisión de las muestras tomadas a más de 250 mm por encima del controlador en (Guna et al., 2014).

## 2. Estado del Arte

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
		(Weichert et al., 2013)				Con un robot industrial con una pluma de referencia se logra una precisión de 0,2 mm en (Weichert et al., 2013).
(Montalvo Martínez, 2010)	Visión	Para el cálculo de la estructura en 3D de la escena, se usan técnicas de visión estereoscópica.	No aplicado	No aplicado	No aplicado	Estudio sobre la eficacia de una serie de métodos de correspondencia estereoscópica
(Smeragliuolo et al., 2016)	<i>Leap Motion</i>	Validación del <i>Leap Motion Controller</i> utilizando tecnología de captura de movimiento de marcadores	No aplicado	No aplicado	No aplicado	El <i>Leap Motion</i> no puede proporcionar datos que sean clínicamente significativos para la flexión/extensión de la muñeca y quizás para la desviación de la muñeca.
(Noor y Aras, 2015)	Visión	Potencial de interacción multimodal y multiusuario con holografía virtual	No aplicado	No aplicado	No aplicado	Se describen las interacciones de múltiples usuarios, que admiten ordenadores locales y distribuidas, utilizando una variedad de pantallas.
(Zhang y Meruvia-Pastor, 2017)	<i>Leap Motion</i>	Operar paneles virtuales con gestos manuales en juegos de realidad virtual inmersivos	No aplicado	No aplicado	No aplicado	Interacción basada en gestos para juegos de realidad virtual utilizando el motor de juegos <i>Unity</i> , el sensor <i>Leap Motion</i> , teléfono inteligente y auriculares de realidad virtual
(Iosa et al., 2015)	<i>Leap Motion</i>	Exploración de la viabilidad de adaptar el <i>Leap Motion</i> , desarrollado para videojuegos, a la neuro-rehabilitación de ancianos con ictus subagudo	Muestra de datos	No aplicado	No aplicado	La rehabilitación con <i>Leap Motion</i> se realizó con un alto nivel de participación activa, sin efectos adversos, y contribuyó a incrementar la recuperación de las habilidades manuales

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(Bassily et al., 2014)	<i>Leap Motion</i>	Manipulación de brazos robóticos con <i>Leap Motion Controller</i>	Muestra de datos	Algoritmo de mapeo	No aplicado	Se propone esquema de manipulación de brazo robótico para permitir la incorporación de sistemas robóticos en el entorno doméstico
(Gándara y Bauza, 2015)	<i>Leap Motion</i>	El sistema que controla de forma remota diferentes electrodomésticos de un hogar u oficina a través de interfaces de usuario naturales en la interacción hombre-máquina	No aplicado	No aplicado	No aplicado	Se construye un entorno de bajo costo y en tiempo real capaz de ayudar a las personas con discapacidad
(Hantrakul y Kaczmarek, 2014)	<i>Leap Motion</i>	Implementaciones del dispositivo <i>Leap Motion</i> en síntesis de sonido y actuaciones interactivas en vivo	No aplicado	No aplicado	No aplicado	Busca empoderar a pacientes discapacitados con expresión musical utilizando tecnología de seguimiento de movimiento y sensores portátiles
(Cichon y Roßmann, 2017; Jetter et al., 2020; Ling y Rui, 2016; Vosinakis y Koutsabasis, 2018)	<i>Leap Motion</i> y gafas de Realidad Virtual	Gafas de realidad virtual y tendencias del <i>Leap Motion</i> en educación	No aplicado	No aplicado	No aplicado	La tecnología sigue actualizando y fomentando la innovación en la educación. La realidad virtual y los sensores como <i>Leap Motion</i> pretenden tener un gran impacto en ella
(Leth, 2023)	Realidad Virtual	Gafas de Realidad Virtual. Reconocimiento de 25 signos de lenguaje de signos danés	No aplicado	No aplicado	Procesamiento de lenguaje natural (NLP)	Se ha obtenido un acierto del 44,4% en signos dinámicos y un 41,3% en signos estáticos
(Fuchs et al., 2022; Georgiou, Frier, et al., 2022;	<i>UltraLeap</i>	Estudio de las señales hápticas	No aplicado	No aplicado	No aplicado	Las conclusiones que se extraen derivan del posible uso de este tipo de dispositivos en varios campos

## 2. Estado del Arte

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
Rakkolainen et al., 2021; Singh et al., 2021)						
(Mulot et al., 2021)	<i>UltraLeap</i>	Uso para el manejo de robots usando gafas de realidad virtual	No aplicado	No aplicado	No aplicado	Las conclusiones que se extraen es que al usuario le resulta más real el manejo de cualquier robot gracias a la sensación táctil
(Brown et al., 2020; Young et al., 2020)	<i>UltraLeap</i>	Uso para el manejo de elementos de un vehículo	No aplicado	No aplicado	No aplicado	Por ejemplo, en el manejo de los elementos interiores de un coche esa sensación táctil ayuda a que los conductores se centren más en la conducción
(Van den Bogaert et al., 2021)	<i>UltraLeap</i>	Uso para el manejo y automatización del hogar	No aplicado	No aplicado	No aplicado	La sensación que produce el sensor ayuda a que parezca real el manejo de los diferentes elementos del hogar
(Georgiou, Martinez, et al., 2022)	<i>UltraLeap</i>	Uso para el manejo de diferentes tipos de objetos usando gafas de realidad virtual	No aplicado	No aplicado	No aplicado	La sensación táctil de manejar objetos, al mismo tiempo que se visualiza por las gafas de realidad virtual, ayuda a manejar dichos objetos comparado con la sensación de vacío
(Paneva et al., 2020)	<i>UltraLeap</i>	Reconocimiento del lenguaje Braille para personas ciegas	No aplicado	No aplicado	No aplicado	Al ser un sistema táctil ayuda a la diferenciación de palabras por puntos
(Y. Li, 2012; Patsadu et al., 2012; Ren et al., 2013; C. Wang et al., 2015; Y. Wang et al., 2012)	<i>Kinect</i>	Reconocimiento de gestos	No aplicado	No aplicado	No aplicado	Para el reconocimiento utilizan diferentes tipos de técnicas y sus resultados son múltiples.

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(Sun et al., 2015)	<i>Kinect</i>	Reconocimiento del lenguaje de signos americano de 2000 frases	No aplicado	No aplicado	Máquina de vectores de soporte (SVM).	Reconoce el 86% de las frases
(Zafrulla et al., 2011)	<i>Kinect</i>	Reconocimiento del lenguaje de signos americano de 1000 palabras	No aplicado	No aplicado	Modelos ocultos de Markov (HMM).	Alcanza una precisión del 76,12%.
(Yang et al., 2020)	<i>Kinect</i>	Reconocimiento del lenguaje de signos americano de 24 palabras	No aplicado	No aplicado	Técnica del campo aleatorio condicional jerárquico usando redes neuronales (NN)	Obtiene una precisión del 93,2%
(D. Li et al., 2020)	<i>Kinect</i>	Reconocimiento del lenguaje de signos americano de 2000 palabras	No aplicado	No aplicado	Usa dos modelos; Enfoque holístico basado en la apariencia visual; poses humanas en 2D	Obtiene una precisión del 62,63%
(Abdel-Samie et al., 2018)	<i>Kinect</i>	Reconocimiento de 30 palabras del lenguaje de signos árabe.	No aplicado	No aplicado	Deformación dinámica del tiempo (DTW).	Los resultados varían según los márgenes de error entre un 95,25% y un 97,58%.
(Sarhan et al., 2015)	<i>Kinect</i>	Reconocimiento de 16 palabras del lenguaje de signos árabe.	No aplicado	No aplicado	Modelos ocultos de Markov (HMM).	Llega a una precisión del 80,47%.
(Geng et al., 2014)	<i>Kinect</i>	Reconocimiento de 20 gestos del	No aplicado	No aplicado	Máquina de vectores de soporte (SVM)	Precisión del 61,46%

## 2. Estado del Arte

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
		lenguaje de signos chino				
(Chai et al., 2013)	<i>Kinect</i>	Reconocimiento de 239 gestos del lenguaje de signos chino	No aplicado	No aplicado	Galería de trayectorias 3D comparando las generadas con los patrones	El porcentaje de precisión oscila entre 83,51% y el 96,32%.
(Liu et al., 2016)	<i>Kinect</i>	Reconocimiento de 500 gestos del lenguaje de signos chino	No aplicado	No aplicado	Uso de redes neuronales (LSTM)	Precisión del 63,3%
(Lang et al., 2012)	<i>Kinect</i>	Reconocimiento de del lenguaje de signos alemán	No aplicado	No aplicado	Modelos ocultos de Markov (HMM)	Acierto del 97%
(Yauri Vidalón y De Martino, 2016)	<i>Kinect</i>	Reconocimiento de 107 palabras del lenguaje de signos brasileño	No aplicado	No aplicado	Deformación dinámica del tiempo (DTW); Redes neuronales vecino más cercano (K-NN)	Para el DTW el porcentaje es del 90%. Para el K-NN es del 98,96%
(Pigou et al., 2015)	<i>Kinect</i>	Reconocimiento de 20 gestos del lenguaje de signos italiano	No aplicado	No aplicado	Redes neuronales convolucionales (CNN)	Acierto del 91,7%
(Mehrotra et al., 2015)	<i>Kinect</i>	Reconocimiento de 37 gestos del lenguaje de signos indio	No aplicado	No aplicado	Máquina de vectores de soporte multiclase (SVMM)	Porcentaje de acierto del 86,16%

Referencia	Tipo de dispositivo	Enfoque	Segmentación	Extracción de características	Clasificación	Conclusiones
(G. C. Lee et al., 2016)	<i>Kinect</i>	Reconocimiento de 25 gestos del lenguaje de signos taiwanés	No aplicado	No aplicado	Combina Modelos ocultos de Markov (HMM) con Máquina de vectores de soporte (SVM)	Consigue una precisión del 85,14%
(Bora et al., 2023)	<i>Kinect</i>	Reconocimiento de 9 letras del lenguaje de signos indio	No aplicado	MediaPipe	Redes neuronales (NN)	Consigue una exactitud del 99%
(Brock y Gomez, 2021)	<i>Azure Kinect</i>	Reconocimiento con 18 gestos	No aplicado	No aplicado	Redes neuronales convolucionales (CNN); Clasificadores de conjuntos aplicados (ENS)	Precisión del 41% con CNN y del 75% con ENS.



# Capítulo 3: Materiales y Métodos



## 3. Materiales y métodos

### 3.1 Introducción

En la primera parte de este capítulo, se explica el funcionamiento del controlador del *Leap Motion*. A continuación, se explica cómo se ha realizado la adquisición de datos. En el siguiente apartado se muestra cómo se ha generado la base de datos con las 176 palabras seleccionadas para el desarrollo de la investigación. Seguidamente, se explica los métodos usados para generar patrones y el proceso seguido para recoger el movimiento de las manos, el funcionamiento del comparador y la técnica de *Dinamic Time Warping* (DTW). En la última parte, se explica la implementación del comparador utilizando las diferentes bibliotecas de *Machine Learning* de Python.

### 3.2 Adquisición

#### 3.2.1 Sensor Leap Motion

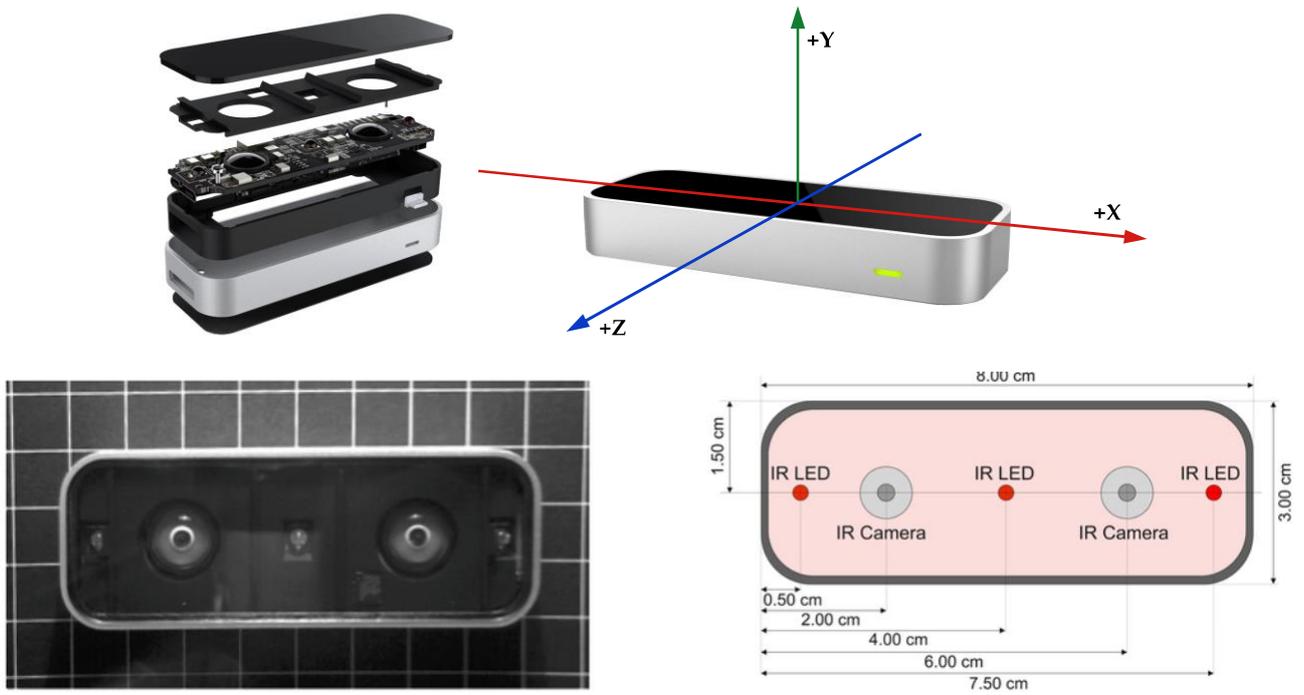
Cómo se ha comentado en el capítulo anterior, este sensor es capaz de capturar el movimiento de las manos, los dedos y los brazos. Es un dispositivo compacto, versátil y económico que contiene dos cámaras infrarrojas, (véase la Figura 3.1), con un ángulo de 120° y tres LED infrarrojos, por lo que se clasificaría como una cámara ToF, según lo descrito en el capítulo 2 del Estado del Arte. El funcionamiento de este dispositivo está en el rango del espectro no visible, con una longitud de onda de 850 nm. Éste es capaz de alcanzar hasta 200 fps o fotogramas por segundo. También tiene la capacidad, de ir adaptándose a la luz de su entorno y asegurando una resolución de imagen alta y constante.

Una de las características que se pueden ver en la Figura 3.1, son la barrera de plástico que existen entre los LEDs y las propias cámaras, logrando una iluminación uniforme en todo el espacio.

Otra característica que hay que destacar, es que la tecnología que se utiliza es del tipo CMOS, haciendo que la digitalización ocurra dentro de cada una de las celdas, dando como resultado una mayor velocidad de captura y consumo bajo de hardware (*Interacción 3D con Leap Motion*, 2014).

Para realizar la función de regular la iluminación, se utiliza el microcontrolador (Macronix MX25L3206E (*MX25L3206E DATASHEET*, s. f.)), haciendo que la información que se recoge se envía por USB.

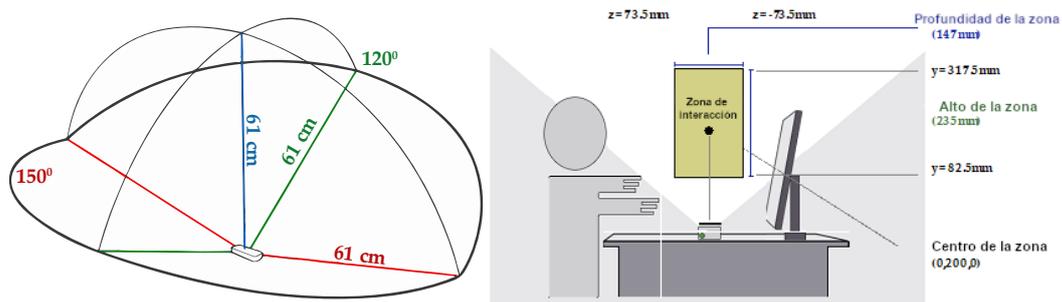
#### 4. Metodologías Experimentales



**Figura 3.1.** Diferentes componentes del sistema de hardware de Leap Motion

Por otro lado, el controlador USB ejecuta la transferencia de datos al ordenador al que esté conectado el dispositivo. Los puertos serie UART\_Rx y UART\_Tx, son los encargados de recibir los datos del dispositivo a través del puerto USB 3.0.

El área de interacción, que se puede observar en la parte izquierda de la Figura 3.2, forma un semi-hemisferio con un ángulo de visión de  $150^\circ$  y con un rango efectivo que se extiende desde los 25 milímetros hasta los 600 milímetros, definiendo lo que se denomina casa interacción. El rango de altura de interacción (véase la parte derecha de la Figura 3.2), se puede configurar entre 73,5 milímetros y 235 milímetros.



**Figura 3.2.** Área de interacción del Leap Motion

La intensidad que puede dar el USB junto con el ángulo de visión de la cámara, influyen en el área de interacción. Cabe señalar que el ángulo de visión, se identifica con la Ecuación 3.1, donde también se ve afectado por la distancia focal y el tamaño del sensor, de la siguiente manera: “ $d$ ” es la diagonal del sensor y “ $f$ ” representa la distancia focal.

$$\alpha = 2 * \tan^{-1} \left( \frac{d}{2f} \right) \quad (3.1)$$

Una de las características más interesantes que posee el Leap Motion, es que tiene conexión USB, siendo esta compatible con sistemas operativos como el Mac, Windows, Android y Linux. Los desarrolladores además tienen la posibilidad de usar una API, para que puedan trabajar en diferentes lenguajes de programación como, C++, JavaScript, Python, C#, o Objective-C. Para el acceso de los datos de seguimiento, es necesario trabajar con el lenguaje C. El *Unity* y el *Unreal Engine*, tienen la posibilidad de obtener complementos para sus motores gráficos.

En este caso, el controlador analiza cada imagen en busca de manos y dedos, y envía un conjunto de fotogramas al ordenador a través del puerto USB. Cada objeto de la captura, (véase la Figura 3.3), contiene clases de tipo Mano, con sus correspondientes subclases de Brazo (un objeto similar a un hueso que proporciona la orientación, la longitud, el ancho y los extremos de un brazo), así como las subclases de Dedo, que contienen datos relacionados con la identificación de tipos, posición de dirección, orientación, velocidad, largo y ancho, ángulo entre ellos, e incluso porcentaje de coincidencia con el modelo interno que contiene este dispositivo. A parte de todos estos parámetros, también puede reconocer ciertos gestos dinámicos predefinidos (Gesto) y herramientas (Herramienta) (*API Overview — Leap Motion JavaScript SDK v3.2 Beta documentation, 2022*).

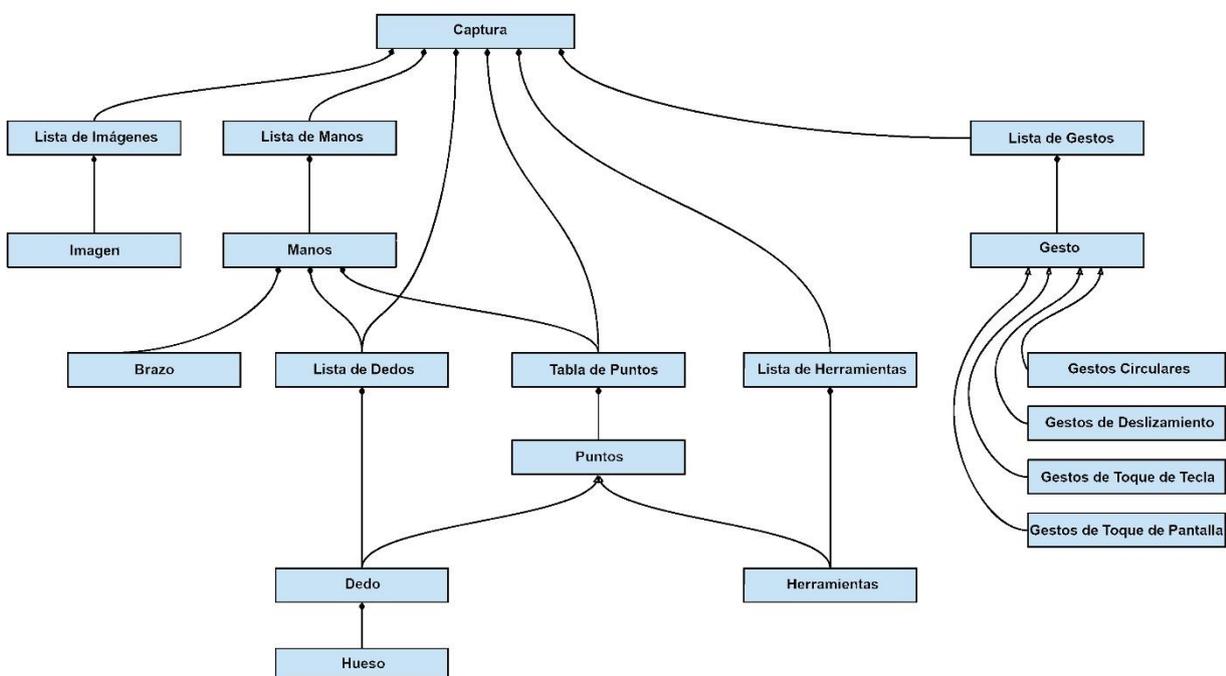


Figura 3.3. Objeto Frame del Leap Motion

Aun cuando las partes de la mano debido a la oclusión no son visibles, el sistema proporciona un seguimiento predictivo para intentar que no se pierdan en el área interactiva. Para ello el software usa las

#### 4. Metodologías Experimentales

partes que son visibles de una mano, junto con las posiciones más probables en un momento dado, para su predicción (*API Overview — Leap Motion JavaScript SDK v3.2 Beta documentation*, s. f.).

Además del modelo de seguimiento del esqueleto de la mano, (véase la parte izquierda de la Figura 3.4), *Leap Motion* también proporciona una pantalla en la que se pueden observar en tiempo real, las secuencias grabadas de una escena que sucede frente al dispositivo, (véase la parte derecha de la Figura 3.4).

Las cámaras biconvexas, son las encargadas de captar la radiación infrarroja de una escena, al mismo tiempo se genera una imagen digitalizada, usando los datos proporcionados por los sensores.

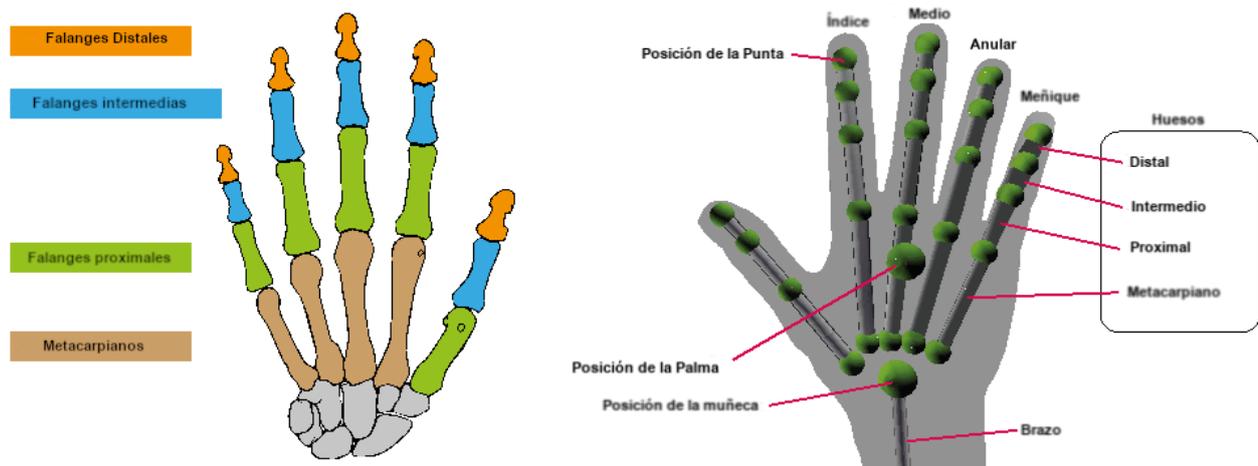


Figura 3.4. Seguimiento del modelo Esquelético de las Manos

A medida que se van obteniendo imágenes, se puede observar una distorsión que producen las lentes, (véase la Figura 3.5) sobre estas imágenes ópticas, deformándolas. El nombre para la distorsión que produce el *Leap Motion*, es distorsión compleja, siendo una mezcla entre la distorsión de corsé y de barril (Guna et al., 2014).

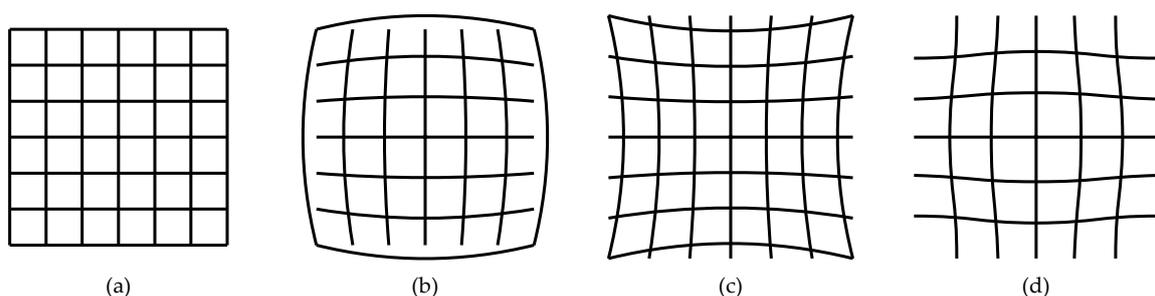


Figura 3.5. Distorsión que las lentes pueden producir. (a) Sin distorsión. (b) Distorsión de barril. (c) Distorsión de corsé. (d) Distorsión de mostacho.

Para poder calibrar la imagen, es necesario superponer una malla de puntos (véase la Figura 3.6). Esta información de distorsión, va acompañada para cada uno de los datos. También se puede corregir el brillo usando el mapa de calibración.

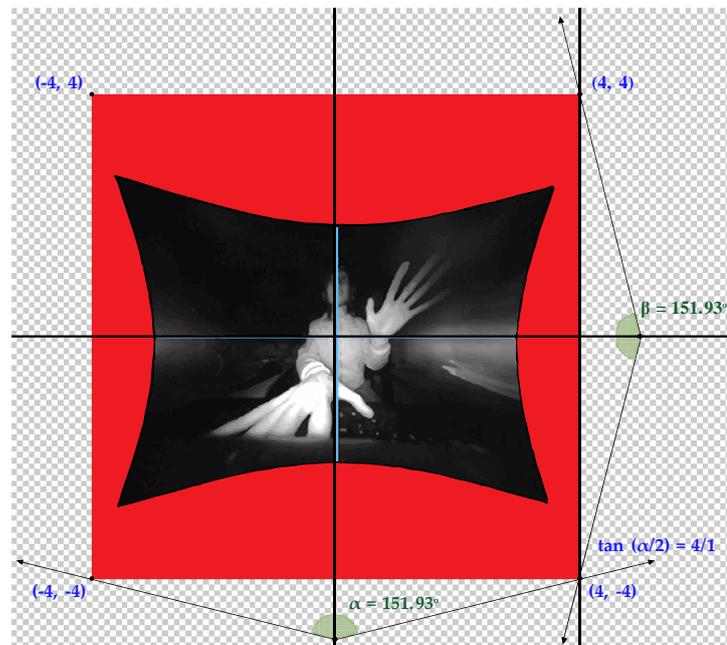


Figura 3.6. Imagen con la Distorsión Corregida antes de llegar al Microcontrolador

Para determinar la posición dentro un sistema de coordenadas, se utiliza la técnica de visión telescópica (Montalvo Martínez, 2010). Éstas se establecen, una vez que el microcontrolador haya corregido e identificado las imágenes de manos y dedos.

### 3.2.2 Proceso de Adquisición

Uno de los primeros desarrollos de esta investigación, fue la creación de un software que fuese capaz de adquirir los datos, y almacenar cada una de las muestras. De esta manera, se fue creando el diccionario gestual. También se desarrolló un soporte realizado por impresión 3D, para que el *Leap Motion* trabajara de forma frontal y así grabar los signos de una forma más simple, (véase la Figura 3.7). También se creó un protocolo de grabación, para que las muestras se grabaran de forma precisa.

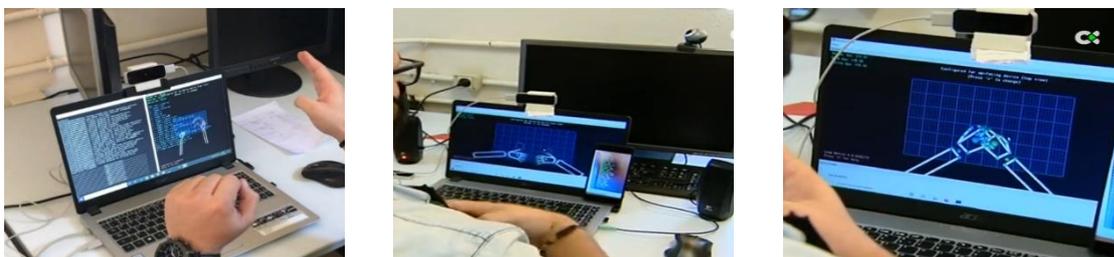
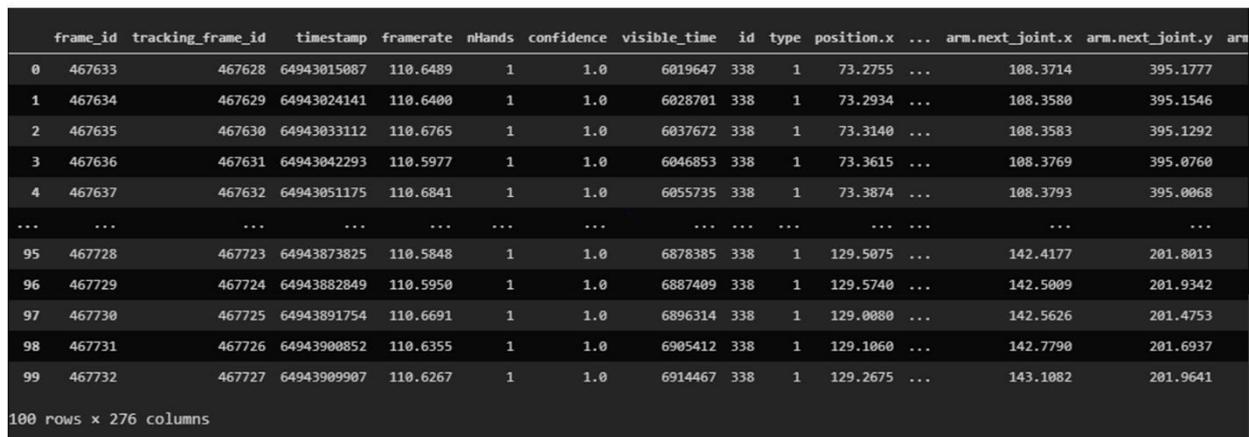


Figura 3.7. Grabación de muestras

Para cada una de las palabras, se obtiene un archivo de texto (signo dinámico), donde cada una de esas muestras tiene 276 parámetros (véase la Tabla 3.6). Cada uno de ellos hace referencia al movimiento de dedos y manos, y de la posición en un momento dado. Para la validación de cada una de las muestras, se desarrolló un software que visualizaba el movimiento de las manos y de los dedos.

## 4. Metodologías Experimentales

La base de datos total a la finalización del proceso fue de 5780 muestras, donde cada una pertenecía a un archivo con los 276 parámetros *del Leap Motion Controler*. Un detalle a tener en cuenta, es la velocidad de muestreo, la cual varía debido algunos factores. Uno de ellos es el tipo de hardware que procesa las muestras, otro corresponde al tipo de puerto y otro al software del sistema operativo, entre otros. Los ficheros tienen un tamaño que van desde los 350KB a los 2MB. Por ejemplo, en el signo de la palabra “hola”, (véase la Figura 3.8), se muestrea alrededor de 110 fotogramas por segundo (fps), para cada mano.



frame_id	tracking_frame_id	timestamp	framerate	nHands	confidence	visible_time	id	type	position.x	...	arm.next_joint.x	arm.next_joint.y	arm
0	467633	64943015087	110.6489	1	1.0	6019647	338	1	73.2755	...	108.3714	395.1777	
1	467634	64943024141	110.6400	1	1.0	6028701	338	1	73.2934	...	108.3580	395.1546	
2	467635	64943033112	110.6765	1	1.0	6037672	338	1	73.3140	...	108.3583	395.1292	
3	467636	64943042293	110.5977	1	1.0	6046853	338	1	73.3615	...	108.3769	395.0760	
4	467637	64943051175	110.6841	1	1.0	6055735	338	1	73.3874	...	108.3793	395.0068	
...	...	...	...	...	...	...	...	...	...	...	...	...	...
95	467728	64943873825	110.5848	1	1.0	6878385	338	1	129.5075	...	142.4177	201.8013	
96	467729	64943882849	110.5950	1	1.0	6887409	338	1	129.5740	...	142.5009	201.9342	
97	467730	64943891754	110.6691	1	1.0	6896314	338	1	129.0080	...	142.5626	201.4753	
98	467731	64943900852	110.6355	1	1.0	6905412	338	1	129.1060	...	142.7790	201.6937	
99	467732	64943909907	110.6267	1	1.0	6914467	338	1	129.2675	...	143.1082	201.9641	

100 rows x 276 columns

Figura 3.8. Muestra del archivo del signo hola

### 3.3 Base de Datos

De la base total de muestras, 3520 se han usado para entrenar, dejando 2260 para la verificación. Una característica importante de la lengua de signos española, es que no utiliza preposiciones, y donde una oración contiene un verbo en infinitivo, sustantivos o adjetivos. En el futuro, se prevé añadir un módulo para procesar el lenguaje natural. Uno de los primeros objetivos de este proyecto de investigación, es considerar, que la propuesta mejora con respecto a lo explicado en el capítulo dos de esta tesis, usando más signos dinámicos.

Para poder cumplir con los principios éticos sobre experimentos en personas humanas, se ha aplicado la declaración de Helsinki. Cabe destacar que el sensor sólo captura el movimiento de las manos, por lo que es prácticamente imposible reconocer a ninguno de los usuarios.

#### 3.3.1 Conjunto de Palabras Usadas

Desde un principio, para darle una practicidad a esta investigación, se consideró enmarcarla en un caso real de la vida cotidiana de cualquier persona con discapacidad auditiva. Para ello se mostró que uno de esos casos podía ser la visita a un médico. Para tal fin, se escogieron palabras relacionadas con la salud y se añadieron verbos y palabras de uso cotidiano que le diesen forma a una conversación entre paciente y médico. El conjunto total de palabras fueron 176, (véase la Tabla 3.1).

Para la fase metodológica y experimental, trabajar directamente con ese conjunto de palabras era muy complejo, por lo que se consideró trabajar en conjuntos de 50, 100 y finalmente 176 palabras, para comprobar la eficacia del sistema.

En las Tablas 3.2, 3.3 y 3.4 se pueden observar los diferentes conjuntos de palabras escogidas de forma aleatoria.

**Tabla 3.1.** Conjunto de 176 palabras usadas en la investigación

Verbos						
Caer	Caminar	Cocinar	Comer	Consultar	Correr	Curar
Desaparecer	Descansar	Duchar	Engordar	Es	Hacer	Intentar
Ir	Llevar	Medir	Operar	Pasear	Preguntar	Quemar
Reducir	Romper	Sentir	Sufrir	Tener		
Palabras médicas						
Accidente	Alergia	Alzheimer	Ambulancia	Análisis sangre	Ansiedad	Asma
Audífono	Bacterias	Bata	Cabeza	Cáncer	Circulación	Corazón
Cuello	Cuidado	Depresión	Desmayo	Diabetes	Eructo	Espalda
Estómago	Estrés	Fiebre	Flema	Glúteo	Gordo	Guantes
Hemorragia	Herida	Hígado	Hinchazón	Hombro	Hormona	Hospital
Ictus	Implante	Infarto	Inflamación	Intestino	Inyección	Lesión
Lumbago	Mandíbula	Manos	Mareo	Mascarilla	Medicamentos	Médico
Molesto	Ojos	Oreja	Órganos	Ovarios	Oxígeno	Pánico
Párpado	Pecho	Pies	Próstata	Pulmones	Puntos sutura	Salud
Sangre	Seguro	Senos	Sensación	Sordo	Suero	Tensión
Urgencia	Vacuna	Vagina	Vejiga	Virus	Vómito	
Palabras de uso cotidiano						
Abril	Además	Adiós	Agosto	Ahora	Aire	Algo
Alta	Anoche	Antes	Ayer	Bien	Boca	Buenas noches
Buenas tardes	Buenos días	Centro	Coche	Codo	De acuerdo	Delgado
De nada	Después	Día	Diciembre	Ejemplo	Ellos	Esfuerzo
Familia	Febrero	Fin	Forma	Frio	Funciona	Gracias
Hambre	Hasta	Hombre	Hora	Información	Junio	Lluvia
Mamá	Marzo	Mayor	Mediodía	Miércoles	Moto	Mucho
Nada	No	Noche	No lo sabía	No sé	Octubre	Peligro
Poco	Privado	Problema	Qué tal	Raro	Regular	Riesgo
Ruido	Sábado	Seguro	Septiembre	Sexo	Siempre	Sin
Solo	Suyo	También	Viernes	Ya		

**Tabla 3.2.** Conjunto de 100 palabras usadas en la investigación

Verbos						
Caer	Cocinar	Comer	Desaparecer	Engordar	Hacer	Intentar
Ir	Llevar	Operar	Pasear	Reducir	Sufrir	
Palabras médicas						
Alta	Ambulancia	Asma	Audífono	Bacterias	Bata	Bien
Boca	Cáncer	Circulación	Corazón	Cuidado	Desmayo	Eructo
Espalda	Estrés	Fiebre	Flema	Gordo	Hemorragia	Herida
Hinchazón	Hormona	Hospital	Implante	Infarto	Inyección	Lesión
Lumbago	Manos	Mareo	Medicamentos	Ojos	Órganos	Oxígeno
Párpado	Próstata	Puntos sutura	Salud	Sordo	Tensión	Vacuna
Vejiga	Sordo	Virus				
Palabras de uso cotidiano						
Abril	Además	Aire	Algo	Anoche	Antes	Ayer
Buenas noches	Buenas tardes	Buenos días	De acuerdo	De nada	Día	Ejemplo
Ellos	Familia	Forma	Funciona	Hasta	Hombre	Hora
Información	Mamá	Marzo	Mayor	Mediodía	Miércoles	Moto
Nada	Noche	No sé	Peligro	Poco	Privado	Que tal
Ruido	Sensación	Septiembre	Sin	También	Ya	

#### 4. Metodologías Experimentales

Tabla 3.3. Conjunto de 50 palabras usadas en la investigación

Verbos						
Cocinar	Engordar	Hacer	Intentar	Reducir	Sufrir	
Palabras médicas						
Alta	Asma	Bata	Boca	Cáncer	Corazón	Desmayo
Ejemplo	Espalda	Flema	Hemorragia	Hinchazón	Hospital	Infarto
Lumbago	Ojos	Órganos	Párpado	Próstata	Salud	Seguro
Sordo	Tensión	Vacuna	Vejiga	Virus		
Palabras de uso cotidiano						
Abril	Aire	Anoche	Ayer	Buenas tardes	Buenos días	De acuerdo
Familia	Funciona	Hora	Marzo	Moto	Peligro	Privado
Que tal	Ruido	Septiembre	Ya			

Para demostrar que el sistema era robusto e independiente de la persona que signa, se generó una base de datos de 50 palabras aleatorias, donde 15 personas anónimas signaron, forzando al mismo tiempo, que cada palabra fuera ejecutada por 3 personas diferentes. El conjunto de esas palabras utilizadas se muestra en la Tabla 3.4.

Tabla 3.4. Conjunto de 50 palabras usadas como muestras externas para la investigación

Verbos						
Cocinar	Comer	Descansar	Llevar	Preguntar	Quemar	Sentir
Palabras médicas						
Alta	Ansiedad	Asma	Boca	Cabeza	Circulación	Corazón
Inflamación	Mandíbula	Mascarilla	Molesto	Mucho	Nada	Ojos
Oxígeno	Sordo	Virus				
Palabras de uso cotidiano						
Abril	Adiós	Ahora	Aire	Anoche	Ayer	Bien
Buenas noches	Delgado	De nada	Día	Ejemplo	Fin	Funciona
Gracias	Hambre	Hombre	Lluvia	Mamá	Marzo	No sé
Peligro	Regular	Roto	También	Ya		

### 3.4 Preprocesamiento de Datos

Para el desarrollo de esta propuesta, ha sido necesario usar bibliotecas de inteligencia artificial o IA. El Python es uno de los lenguajes que más proporciona este tipo de bibliotecas, por eso se decidió trabajar con él.

El funcionamiento del sistema tiene como primer paso, cargar en la memoria del dispositivo todos los archivos que contienen los parámetros. Estos archivos, están configurados de forma que el eje Y de la matriz que forman estos parámetros, está el número de fotogramas, y en el eje X las variables. Cada una de estas variables definen un campo del LMC. De forma que, cada una de las señales contienen fotogramas, siendo éste, un número variable (véase la Figura 3.8) que depende de la duración del signo.

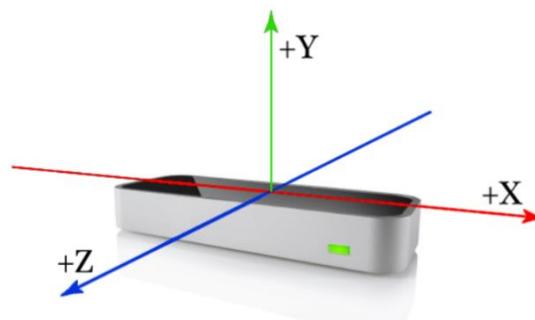
La raíz del modelo de datos parte del objeto "Captura" (véase la Figura 3.3), o "Frame", haciendo una captura en cada uno de los intervalos de actualización. A partir de ahí, se puede extraer el listado de las manos y los dedos en el instante en que fue creado. Los dedos de una mano provienen del objeto "Mano". La clase de "Dedo" o "Finger" define las características de los dedos, al mismo tiempo que el objeto "Brazo", da la posición a la que está unida su mano. Por último, una de las características que posee el objeto "Hueso" o "Bone", es la inclusión de las falanges de los dedos y los metacarpianos (véase la Figura 3.4).

Uno de los elementos que incorpora el LMC, son objetos de imagen no procesadas, junto con la cuadrícula de calibración. Todos los elementos presentados en la Tabla 3.5, se derivan del modelo descrito anteriormente que tienen como base la tabla de todos los parámetros que proporciona el sensor, enunciados en la tabla 3.6. Una de las características que nos da la información de cada una de las muestras, aparte del *timestamp* y el tiempo visible, son los datos relacionados con la orientación, dirección, velocidad, eje normal y el ancho de la palma de la mano (véase la Figura 3.9). También se incluye en esa información, la distancia y ángulos entre los dedos, anchos, rotaciones y articulaciones de cada uno de los huesos y por último el identificador demuestra o id de la mano y el dedo.



**Figura 3.9.** *Movimiento de la mano*

Cómo se comentó anteriormente, el LMC, para evitar perder las manos debido una oclusión, usa un método predictivo, para que no se pierda. Las direcciones se definen según este posicionado el dispositivo y de acuerdo a las coordenadas de la figura 3.10.



**Figura 3.10.** *Ejes del Leap Motion*

## 3.5 Generación de Patrones

### 3.5.1 Parámetros

Para mejorar el rendimiento del sistema, y mejorar su velocidad, se estudió la forma de reducir algunos de los parámetros de las muestras, sin que la detección de los movimientos supusiera la pérdida de información.

El sensor genera la información del movimiento de las manos, en un archivo de tipo texto que contiene 276 parámetros, (véase la Tabla 3.6). Para hacer el sistema más rápido, se redujo el número de parámetros. Muchos de estos parámetros no son necesarios a la hora de trabajar, debido a que contienen

#### 4. Metodologías Experimentales

información adyacente y no fundamental. Uno de estos parámetros es el número de fotograma de cada una de las muestras. Una vez que se realizaron las distintas pruebas, se decidió escoger 74 parámetros según se recoge en la tabla 3.5. Estos parámetros, son los que se consideran que contienen información fundamental a la hora de trabajar con la muestra. Estos parámetros básicos son los que incluyen la rotación de cada uno de las partes de los dedos, necesarios para interpretar como se mueve la mano. El otro parámetro elegido es la posición de cada dedo con respecto a los ejes X, Y y Z. Cada uno de los números tiene una asignación directa con la tabla 3.6.

**Tabla 3.5.** *Parámetros seleccionados*

Parámetros								
4	8	9	10	11	27	28	29	32
41	42	43	52	53	54	63	64	65
74	75	76	79	88	89	90	99	100
101	110	111	112	121	122	123	126	135
136	137	146	147	148	157	158	159	168
169	170	173	182	183	184	193	194	195
204	205	206	215	216	217	220	229	230
231	240	241	242	251	252	253	262	263
264	272							

Para reducir la complejidad del sistema, la mano derecha se separó de la izquierda durante la generación de los patrones.

El principal objetivo para la selección de parámetros, es el sostenimiento de la invariancia del tamaño rotación y traslación, para que, de esta forma, se aplique en diferentes posiciones del sensor, condiciones ambientales distintas o lugares distintos.

#### 3.5.2 Generación

Para establecer la independencia de las muestras, las grabaciones se hicieron en tres sesiones diferentes, donde en cada sesión se grababan 10 muestras. El espacio temporal entre cada una de las sesiones, era al menos de 10 días. Aparte se realizaron grabaciones con personas independientes a la persona especialista, con una sesión de 10 muestras y 50 palabras. Calculado el tiempo que se necesitaba para grabar cada una de las sesiones, y comprobando la duración del proyecto, se decidió grabar 176 palabras. El hecho de grabar todas las palabras juntas, podría significar que, si hubiera habido algún problema, no habría tiempo para rectificar. Entonces se estableció un protocolo de evaluación donde se creaban conjuntos de 50 palabras. Por lo tanto, se comprobó la evolución con 50, 100 y 176 palabras. Por último, con la sesión independiente del primer usuario, y para así comprobar la invarianza del sistema, se consideró grabar 50 palabras con al menos 15 personas diferentes, donde cada palabra fue signada por al menos tres personas.

Teniendo en cuenta las circunstancias tanto del hardware como del software usado, se comprobó que las muestras tenían tamaños diferentes, por lo que se decidió antes de generar cualquier patrón equalizar el tamaño de los archivos para que el número de fotograma fuera el mismo.

Habiendo 30 muestras de tres sesiones independientes, se decidió utilizar 20 muestras para la generación de patrones, donde las mismas fueron grabados por un especialista en la lengua de signos. Si se considera una palabra  $Q$  descompuesta por 20 observaciones de  $Q1$  a  $Q20$  y considerando  $x$  como cada uno de los parámetros de cada muestra.

Primero se buscó la serie más corta, adaptando el tamaño de cada uno de los archivos a ese valor más pequeño. Seguidamente, se creó el patrón promedio dividiendo cada una de sus parámetros por el número de muestras.

$$\frac{\sum_{i=1}^{i=20} Q_i(x_n)}{20} \quad (3.1)$$

De esta forma se consigue que el patrón contenga todas las características del signo sin que se pierda información vital para su reconocimiento, aportando fiabilidad al sistema.

Por último, hay que tener en cuenta que es necesario que el tamaño de cada uno de los archivos no difiera mucho entre ellos.

## 3.6 Comparador

### 3.6.1 Introducción al Dynamic Time Warping, DTW

La deformación dinámica del tiempo (DTW), es una técnica de clasificación (Rakthanmanon et al., 2012), que se utiliza en muchos campos de investigación. Se utiliza de forma habitual para la comparación de series temporales de distintas longitudes. Un campo donde se usa mucho, es en el reconocimiento de voz, donde cada persona puede hablar y decir algo, pero no todos hablamos a la misma velocidad diciendo lo mismo. Con DTW se resuelve este desajuste de tiempo.

En la actualidad, esta técnica se usa en múltiples proyectos de investigación. Por ejemplo, (Tuzcu y Nas, 2005) lo utiliza para el reconocimiento automático de señales de ECG; (Legrand et al., 2008), para el reconocimiento de cromosomas; (Kovacs-Vajna, 2000), en reconocimiento de huellas dactilares; mientras que (Rath y Manmatha, 2003) lo utiliza en el desarrollo del reconocimiento de documentos manuscritos. También se pueden encontrar investigaciones que utilizan DTW para el reconocimiento de firmas en (Okawa, 2019; Piyush Shanker y Rajagopalan, 2007). Dentro de las investigaciones basadas en el reconocimiento de voz se encuentran (Amin y Mahmood, 2008; Muda et al., 2010). Por ejemplo, en el área de reconocimiento gestos faciales la usa (Adwan y Arof, 2012), En cambio, (Arici et al., 2014), la utiliza en el reconocimiento de gestos corporales. Dentro del campo de reconocimiento de gestos, existen muchas investigaciones donde algunos de ellos utilizan la *Kinect* (Calin et al., 2016; Raheja et al., 2015; Reyes et al., 2011; Riofrío et al., 2017).

#### 4. Metodologías Experimentales

---

Finalmente, dentro de las investigaciones que se ocupan del reconocimiento de la lengua de signos se encuentran (Ahmed et al., 2016; Jambhale y Khaparde, 2014; Kuzmanic y Zanchi, 2007).

##### 3.6.2 Desarrollo del DTW

Una de las cuestiones que tener en cuenta cuando se está trabajando con movimientos gestuales, es que no todos movemos las manos de la misma forma y al mismo tiempo. Incluso, nosotros mismos podemos variar la velocidad de las manos según nuestra necesidad. Si se quieren comparar series temporales hay que considerar las siguientes cuestiones:

Si se parte de dos series temporales  $K$  y  $T$ , y al mismo tiempo cada una de ellas tiene una longitud  $m$  y  $n$  respectivamente, hay que considerar cada captura del sensor, como un elemento de la función obteniendo la resta para todas las capturas como:

$$K = k_1, k_2, \dots, k_i, \dots, k_n \quad (3.2)$$

$$T = t_1, t_2, \dots, t_i, \dots, t_n \quad (3.3)$$

Suponiendo que  $K = T$ , no es necesario calcular la distancia de las 2 secuencias. Si se considera  $K \neq T$ , ocurre que existe una desalineación y entonces no queda más opciones que intentar que ambas secuencias, se alineen. Es necesario entonces, generar una matriz de  $m \times n$ , asignando a cada punto de la matriz  $(i, j)$ , al punto  $K_i$  y  $T_j$  del alineamiento.

Ahora el seguimiento del mismo se define como un camino de deformación del camino regular, y se usa  $w$ , (véase la Figura 3.11), para indicar que el elemento  $K$  de  $w$  con el resultado final se obtiene de  $wK = (i,j)K$ :

Se deben cumplir las siguientes condiciones, al mismo tiempo:

1. Condición de contorno;  $w_1 = (1, 1)$  y  $wK = (m, n)$
2. Continuidad; Si  $wK-1 = (A', B')$ , entonces el siguiente punto  $w$  para el camino  $K = (A, B)$  debe cumplir  $(A-A') \leq 1$  y  $(B-B') \leq 1$
3. Monótono; Si  $wK-1 = (A', B')$ , entonces el siguiente punto  $w$  para el camino  $K = (A, B)$  debe cumplir  $0 \leq (A-A')$  y  $0 \leq (B-B')$ .

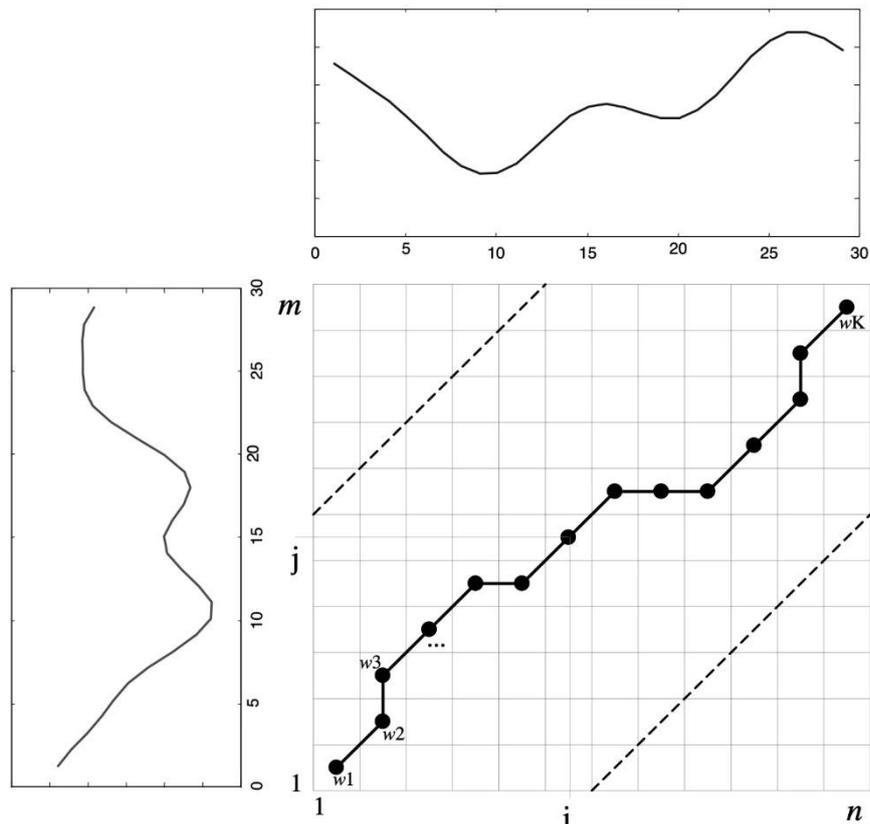


Figura 3.11. Ejemplo de una ruta de deformación en el tiempo

De esta forma, se intenta lograr el objetivo de acortar y alargar ambas secuencias temporales, estableciendo la distancia más corta y haciendo que la distancia  $\gamma(i,j)$  sea mínima.

$$W = w_1, w_2, \dots, w_k, \dots, w_K \quad \max(m,n) \leq K < m + n - 1 \quad (3.4)$$

En el ejemplo gráfico de la Figura 3.12a, se muestra un camino generado por el DTW, correspondiente a una palabra. La Figura 3.12b representa un signo de la misma clase, y se puede comprobar que ambos son similares. Por otro lado, en la Figura 3.12c, se puede observar cómo se comporta un signo diferente del patrón.

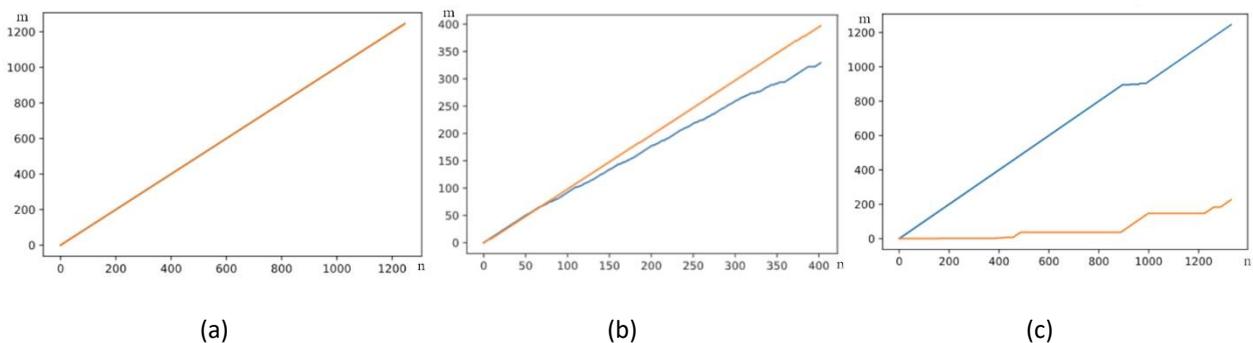


Figura 3.12. Comparador DTW. (a) Ruta del patrón de la palabra. (b) Comparación de una muestra de palabra (azul) con su propio patrón (naranja). (c) Comparación de una palabra (naranja) distinta del patrón (azul).

#### 4. Metodologías Experimentales

##### 3.6.3 Esquema del comparador

La Figura 3.13, muestra la estructura de la que está compuesta el comparador, donde 10 muestras de cada palabra es comparada con los patrones, dando como resultado la distancia más corta con respecto a la entrada.

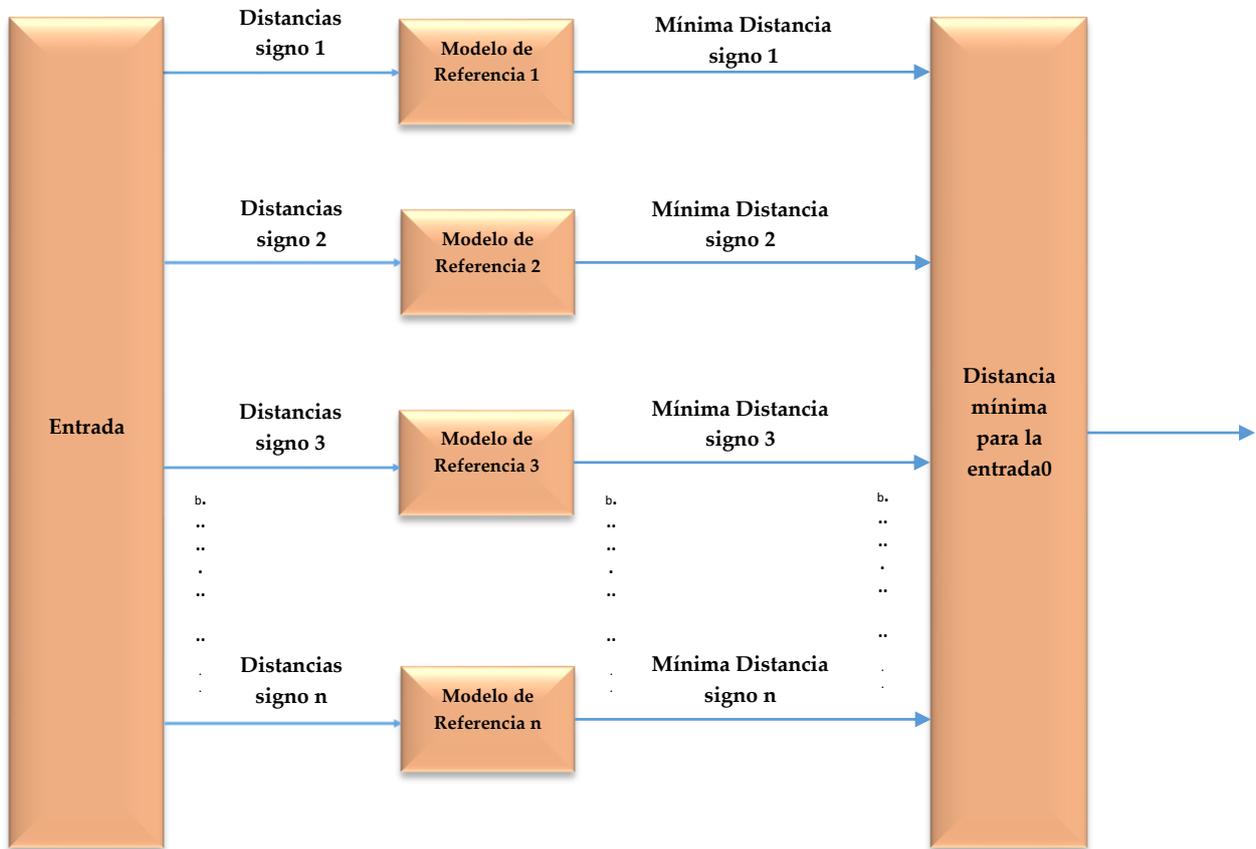


Figura 3.13. Esquema del comparador

Siguiendo el esquema de la figura anterior, la parte de el "Modelo de referencia", corresponde a los patrones. Estos patrones utilizan dos sesiones independientes de 10 muestras cada una.

Debido a que el sistema trabaja a velocidades muy altas de muestreo, se decidió realizar un diezmado de las muestras, para que el clasificador funcionase de forma más rápida.

### 3.7 Implementación

La técnica que se ha usado para implementar el sistema descrito en esta tesis, es el DTW, capaz de comparar diferentes series temporales de distintas longitudes. En la actualidad, cuando se buscan herramientas de *machine learning* en las que se puedan implementar esos algoritmos, prácticamente todas están desarrolladas en lenguaje de programación Python. A partir de ese momento se analizaron los diferentes paquetes que existen, escogiéndolo el paquete *tslearn* que ya tiene implementado este algoritmo. Este sistema, también tiene implementado el aprendizaje automático, y se decidió utilizar porque en

diferentes investigaciones estudiadas que usaban tecnologías con redes neuronales, los resultados podrían mejorarse.

Las librerías más habituales de IA en el lenguaje de programación de Python son *numpy*, *pandas*, *os*, *tslearn* y *sklearn*. Por último, comentar que el módulo *os*, es necesario para la gestión de archivos recordando que cada uno de los archivos corresponde a una muestra. Toda la parte el procesamiento automático y análisis de series temporales, están incluidos en el paquete de *tslearn*, concretamente el *tslearn.metrics.dtw*, donde se obtienen las diferentes métricas, necesarias para la evaluación del sistema. Estas métricas se encuentran en la sección 4.4 de este proyecto investigación.

#### 4. Metodologías Experimentales

**Tabla 3.6.** Conjunto de 276 parámetros completos del *Leap Motion*

	0	1	2	3	4	5	6	7	8	9
<b>0</b>	info.frame_id	tracking_frame_id	info.timestamp	framerate	nHands	confidence	visible_time	id	type	palm.position.x
<b>1</b>	palm.position.y	palm.position.z	pinch_distance	grab_angle	pinch_strength	grab_strength	palm.velocity.x	palm.velocity.y	palm.velocity.z	palm.normal.x
<b>2</b>	palm.normal.y	palm.normal.z	palm.width	palm.direction.x	palm.direction.y	palm.direction.z	palm.orientation.w	palm.orientation.x	palm.orientation.y	palm.orientation.z
<b>3</b>	0	digits[0].finger_id	digits[0].is_extended	digits[0].bones[0].width	digits[0].bones[0].next_joint.x	digits[0].bones[0].next_joint.y	digits[0].bones[0].next_joint.z	digits[0].bones[0].prev_joint.x	digits[0].bones[0].prev_joint.y	digits[0].bones[0].prev_joint.z
<b>4</b>	digits[0].bones[0].rotation.w	digits[0].bones[0].rotation.x	digits[0].bones[0].rotation.y	digits[0].bones[0].rotation.z	digits[0].bones[1].width	digits[0].bones[1].next_joint.x	digits[0].bones[1].next_joint.y	digits[0].bones[1].next_joint.z	digits[0].bones[1].prev_joint.x	digits[0].bones[1].prev_joint.y
<b>5</b>	digits[0].bones[1].prev_joint.z	digits[0].bones[1].rotation.w	digits[0].bones[1].rotation.x	digits[0].bones[1].rotation.y	digits[0].bones[1].rotation.z	digits[0].bones[2].width	digits[0].bones[2].next_joint.x	digits[0].bones[2].next_joint.y	digits[0].bones[2].next_joint.z	digits[0].bones[2].prev_joint.x
<b>6</b>	digits[0].bones[2].prev_joint.y	digits[0].bones[2].prev_joint.z	digits[0].bones[2].rotation.w	digits[0].bones[2].rotation.x	digits[0].bones[2].rotation.y	digits[0].bones[2].rotation.z	digits[0].bones[3].width	digits[0].bones[3].next_joint.x	digits[0].bones[3].next_joint.y	digits[0].bones[3].next_joint.z
<b>7</b>	digits[0].bones[3].prev_joint.x	digits[0].bones[3].prev_joint.y	digits[0].bones[3].prev_joint.z	digits[0].bones[3].rotation.w	digits[0].bones[3].rotation.x	digits[0].bones[3].rotation.y	digits[0].bones[3].rotation.z	1	digits[1].finger_id	digits[1].is_extended
<b>8</b>	digits[1].bones[0].width	digits[1].bones[0].next_joint.x	digits[1].bones[0].next_joint.y	digits[1].bones[0].next_joint.z	digits[1].bones[0].prev_joint.x	digits[1].bones[0].prev_joint.y	digits[1].bones[0].prev_joint.z	digits[1].bones[0].rotation.w	digits[1].bones[0].rotation.x	digits[1].bones[0].rotation.y
<b>9</b>	digits[1].bones[0].rotation.z	digits[1].bones[1].width	digits[1].bones[1].next_joint.x	digits[1].bones[1].next_joint.y	digits[1].bones[1].next_joint.z	digits[1].bones[1].prev_joint.x	digits[1].bones[1].prev_joint.y	digits[1].bones[1].prev_joint.z	digits[1].bones[1].rotation.w	digits[1].bones[1].rotation.x
<b>10</b>	digits[1].bones[1].rotation.y	digits[1].bones[1].rotation.z	digits[1].bones[2].width	digits[1].bones[2].next_joint.x	digits[1].bones[2].next_joint.y	digits[1].bones[2].next_joint.z	digits[1].bones[2].prev_joint.x	digits[1].bones[2].prev_joint.y	digits[1].bones[2].prev_joint.z	digits[1].bones[2].rotation.w
<b>11</b>	digits[1].bones[2].rotation.x	digits[1].bones[2].rotation.y	digits[1].bones[2].rotation.z	digits[1].bones[3].width	digits[1].bones[3].next_joint.x	digits[1].bones[3].next_joint.y	digits[1].bones[3].next_joint.z	digits[1].bones[3].prev_joint.x	digits[1].bones[3].prev_joint.y	digits[1].bones[3].prev_joint.z
<b>12</b>	digits[1].bones[3].rotation.w	digits[1].bones[3].rotation.x	digits[1].bones[3].rotation.y	digits[1].bones[3].rotation.z	2	digits[2].finger_id	digits[2].is_extended	digits[2].bones[0].width	digits[2].bones[0].next_joint.x	digits[2].bones[0].next_joint.y
<b>13</b>	digits[2].bones[0].next_joint.z	digits[2].bones[0].prev_joint.x	digits[2].bones[0].prev_joint.y	digits[2].bones[0].prev_joint.z	digits[2].bones[0].rotation.w	digits[2].bones[0].rotation.x	digits[2].bones[0].rotation.y	digits[2].bones[0].rotation.z	digits[2].bones[1].width	digits[2].bones[1].next_joint.x
<b>14</b>	digits[2].bones[1].next_joint.y	digits[2].bones[1].next_joint.z	digits[2].bones[1].prev_joint.x	digits[2].bones[1].prev_joint.y	digits[2].bones[1].prev_joint.z	digits[2].bones[1].rotation.w	digits[2].bones[1].rotation.x	digits[2].bones[1].rotation.y	digits[2].bones[1].rotation.z	digits[2].bones[2].width
<b>15</b>	digits[2].bones[2].next_joint.x	digits[2].bones[2].next_joint.y	digits[2].bones[2].next_joint.z	digits[2].bones[2].prev_joint.x	digits[2].bones[2].prev_joint.y	digits[2].bones[2].prev_joint.z	digits[2].bones[2].rotation.w	digits[2].bones[2].rotation.x	digits[2].bones[2].rotation.y	digits[2].bones[2].rotation.z
<b>16</b>	digits[2].bones[3].width	digits[2].bones[3].next_joint.x	digits[2].bones[3].next_joint.y	digits[2].bones[3].next_joint.z	digits[2].bones[3].prev_joint.x	digits[2].bones[3].prev_joint.y	digits[2].bones[3].prev_joint.z	digits[2].bones[3].rotation.w	digits[2].bones[3].rotation.x	digits[2].bones[3].rotation.y
<b>17</b>	digits[2].bones[3].rotation.z	3	digits[3].finger_id	digits[3].is_extended	digits[3].bones[0].width	digits[3].bones[0].next_joint.x	digits[3].bones[0].next_joint.y	digits[3].bones[0].next_joint.z	digits[3].bones[0].prev_joint.x	digits[3].bones[0].prev_joint.y
<b>18</b>	digits[3].bones[0].prev_joint.z	digits[3].bones[0].rotation.w	digits[3].bones[0].rotation.x	digits[3].bones[0].rotation.y	digits[3].bones[0].rotation.z	digits[3].bones[1].width	digits[3].bones[1].next_joint.x	digits[3].bones[1].next_joint.y	digits[3].bones[1].next_joint.z	digits[3].bones[1].prev_joint.x
<b>19</b>	digits[3].bones[1].prev_joint.y	digits[3].bones[1].prev_joint.z	digits[3].bones[1].rotation.w	digits[3].bones[1].rotation.x	digits[3].bones[1].rotation.y	digits[3].bones[1].rotation.z	digits[3].bones[2].width	digits[3].bones[2].next_joint.x	digits[3].bones[2].next_joint.y	digits[3].bones[2].next_joint.z
<b>20</b>	digits[3].bones[2].prev_joint.x	digits[3].bones[2].prev_joint.y	digits[3].bones[2].prev_joint.z	digits[3].bones[2].rotation.w	digits[3].bones[2].rotation.x	digits[3].bones[2].rotation.y	digits[3].bones[2].rotation.z	digits[3].bones[3].width	digits[3].bones[3].next_joint.x	digits[3].bones[3].next_joint.y
<b>21</b>	digits[3].bones[3].next_joint.z	digits[3].bones[3].prev_joint.x	digits[3].bones[3].prev_joint.y	digits[3].bones[3].prev_joint.z	digits[3].bones[3].rotation.w	digits[3].bones[3].rotation.x	digits[3].bones[3].rotation.y	digits[3].bones[3].rotation.z	4	digits[4].finger_id

	0	1	2	3	4	5	6	7	8	9
<b>22</b>	digits[4].is_extended	digits[4].bones[0].width	digits[4].bones[0].next	digits[4].bones[0].next	digits[4].bones[0].next	digits[4].bones[0].previous	digits[4].bones[0].previous	digits[4].bones[0].previous	digits[4].bones[0].rotation.w	digits[4].bones[0].rotation.x
<b>23</b>	digits[4].bones[0].rotation.y	digits[4].bones[0].rotation.z	digits[4].bones[1].width	digits[4].bones[1].next	digits[4].bones[1].next	digits[4].bones[1].next	digits[4].bones[1].previous	digits[4].bones[1].previous	digits[4].bones[1].previous	digits[4].bones[1].rotation.w
<b>24</b>	digits[4].bones[1].rotation.x	digits[4].bones[1].rotation.y	digits[4].bones[1].rotation.z	digits[4].bones[2].width	digits[4].bones[2].next	digits[4].bones[2].next	digits[4].bones[2].next	digits[4].bones[2].previous	digits[4].bones[2].previous	digits[4].bones[2].previous
<b>25</b>	digits[4].bones[2].rotation.w	digits[4].bones[2].rotation.x	digits[4].bones[2].rotation.y	digits[4].bones[2].rotation.z	digits[4].bones[3].width	digits[4].bones[3].next	digits[4].bones[3].next	digits[4].bones[3].next	digits[4].bones[3].previous	digits[4].bones[3].previous
<b>26</b>	digits[4].bones[3].previous	digits[4].bones[3].rotation.w	digits[4].bones[3].rotation.x	digits[4].bones[3].rotation.y	digits[4].bones[3].rotation.z	arm.width	arm.next_joint.x	arm.next_joint.y	arm.next_joint.z	arm.prev_joint.x
<b>27</b>	arm.prev_joint.y	arm.prev_joint.z	arm.rotation.w	arm.rotation.x	arm.rotation.y	arm.rotation.z				



# Capítulo 4:

# Metodología

# Experimental



## 4. Metodología experimental

### 4.1 Introducción

En el desarrollo de la investigación, se diseñó una metodología experimental que fuera capaz de garantizar que los resultados tuviesen una veracidad contrastada, y que fuesen dando paso a paso, datos positivos, que mostrasen que el sistema propuesto funcionase de forma correcta.

### 4.2 Validación cruzada o “Cross Validation”

En la actualidad, para la comprobación de resultados en muchas investigaciones, existen diferentes técnicas que ayudan a validar los resultados para considerarlos independientes, y ratificar que las muestras no han sido sesgadas de forma intencionada. Un ejemplo, de técnica usada es la validación cruzada.

La validación cruzada garantiza la independencia de los datos entre el entrenamiento y la prueba de test. Existen varias investigaciones dentro del reconocimiento de signos que usan esta técnica, por ejemplo, (Mapari y Kharat, 2016; Mohandes, Aliyu, et al., 2014; Simos y Nikolaidis, 2016; Vaitkevicius et al., 2019). En esta investigación, como en las referenciadas, se usan con el objetivo de predecir y estimar la precisión de los modelos y la invariancia entre los datos de los experimentos.

El sistema de validación cruzada, (véase la Figura 4.1), que se ha empleado en esta investigación, como en (Chong y Lee, 2018; Nasir et al., 2022), es la “validación cruzada dejando una fuera” o “*Leave-one-out-cross-validation*” con sus siglas en inglés LOOCV, y en ambos casos se utiliza para el reconocimiento del lenguaje de signos. Esta técnica de validación se usa en muchos campos, por ejemplo en astronomía (Welbanks et al., 2023), en el tratamiento de imágenes (Hossain et al., 2023; Jamil et al., 2023), para el tratamiento de vídeo (Singam y Wlode, 2023), en los biocombustibles (Xiao et al., 2023), en el estudio de las corrientes oceánicas (Deogharia y Sil, 2023), en el estudio y aprendizaje en el entorno universitario (Abulela et al., 2023), en el tratamiento de imágenes radiológicas (Hernández y López-Córtés, 2023), en el estudio de ADN (Robeck et al., 2023), estudios de voz para predecir el grado de depresión de un paciente (Y. Wang et al., 2023). Otro ejemplo se puede encontrar en (Zhu et al., 2023), donde utiliza la validación en diferentes procesos químicos. Dentro del campo de la medicina se pueden encontrar investigaciones como la de (Deng et al., 2023), que se basa en estudios para la detección temprana de enfermedades neuro generativas, o (Raets et al., 2023) donde trabaja con datos oncológicos y por último en medicina ginecológica (Kaliappan et al., 2023) para el estudio de fetos.

Esta validación consiste en usar una serie de datos para el entrenamiento dejando fuera los datos de test.

#### 4. Metodología Experimental

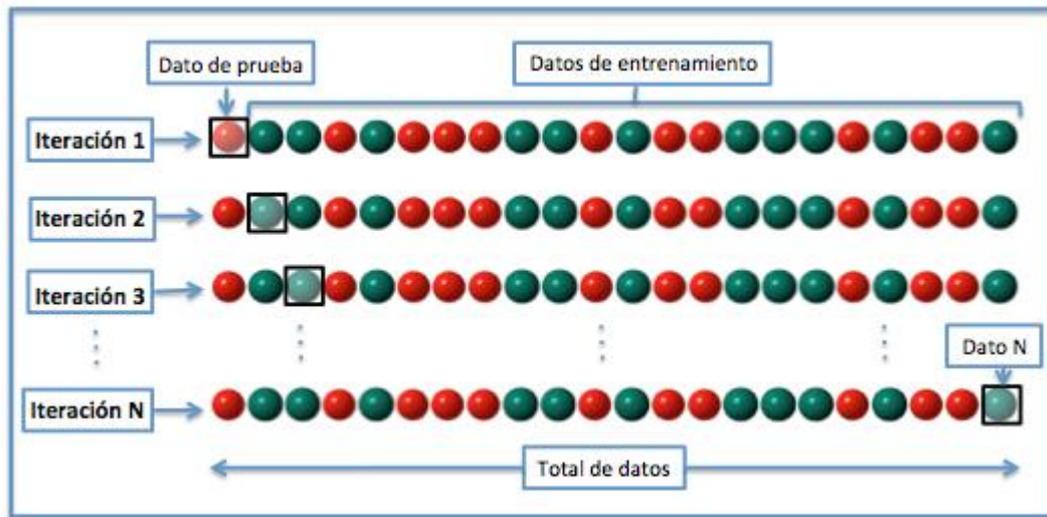
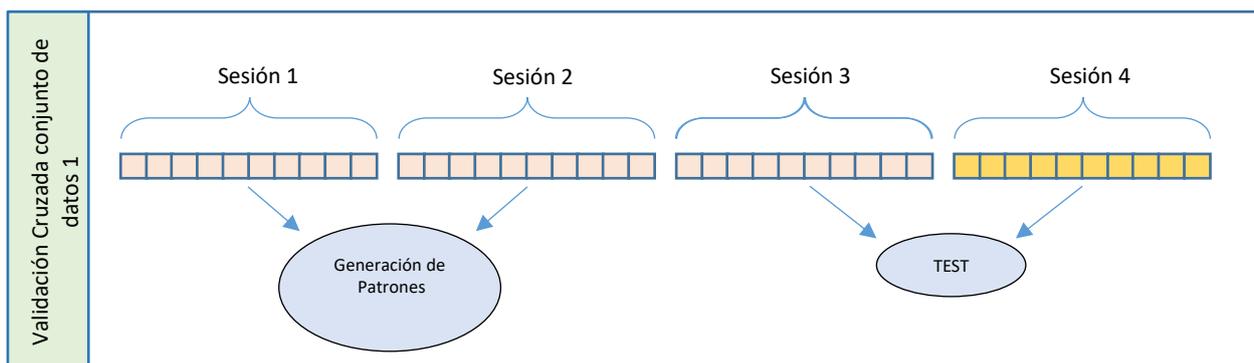


Figura 4.1. Validación cruzada dejando a uno fuera

En el caso de esta investigación, consiste en dividir la base de datos de la que se dispone en conjunto de 10 muestras. La persona especialista con discapacidad auditiva grabó las 5280 muestras. Todas las muestras grabadas se usaron para la generación de patrones y muestras comparativas, donde las mismas se dividieron en 3 sesiones independientes de 10 muestras cada una. Se utilizaron 2 sesiones para crear los patrones, (véase la Figura 4.2), y con la sesión restante y la cuarta sesión independiente, se testean los patrones. Cada conjunto de muestras grabadas por el especialista se realizó con una diferencia de al menos 10 días, para que su independencia fuera lo más amplia posible. La cuarta sesión donde intervinieron 15 personas diferentes de forma aleatoria, se grabó a posteriormente para verificar la robustez del sistema.

La cuarta sesión de 10 muestras por 50 palabras, (500 muestras en total), se compara con los patrones generados en cada una de las 3 sesiones de la validación cruzada. Estas 10 muestras fueron realizadas por 15 personas distintas y cada una de las palabras, fue signada por al menos 3 personas diferentes de forma aleatoria.

Una vez separadas las sesiones se crean el conjunto de datos. El primer conjunto de datos establece que las sesiones 1 y 2 de cada palabra se usarán para generar los patrones y la tercera sesión para la realización del test. El conjunto de datos 2, se configura usando las sesiones 2 y 3 para los patrones y la sesión 1 para el test. Por último, el tercer conjunto de datos usa, para la generación de patrones la sesiones 1 y 3 y para test la sesión 2 (véase la Figura 4.2).



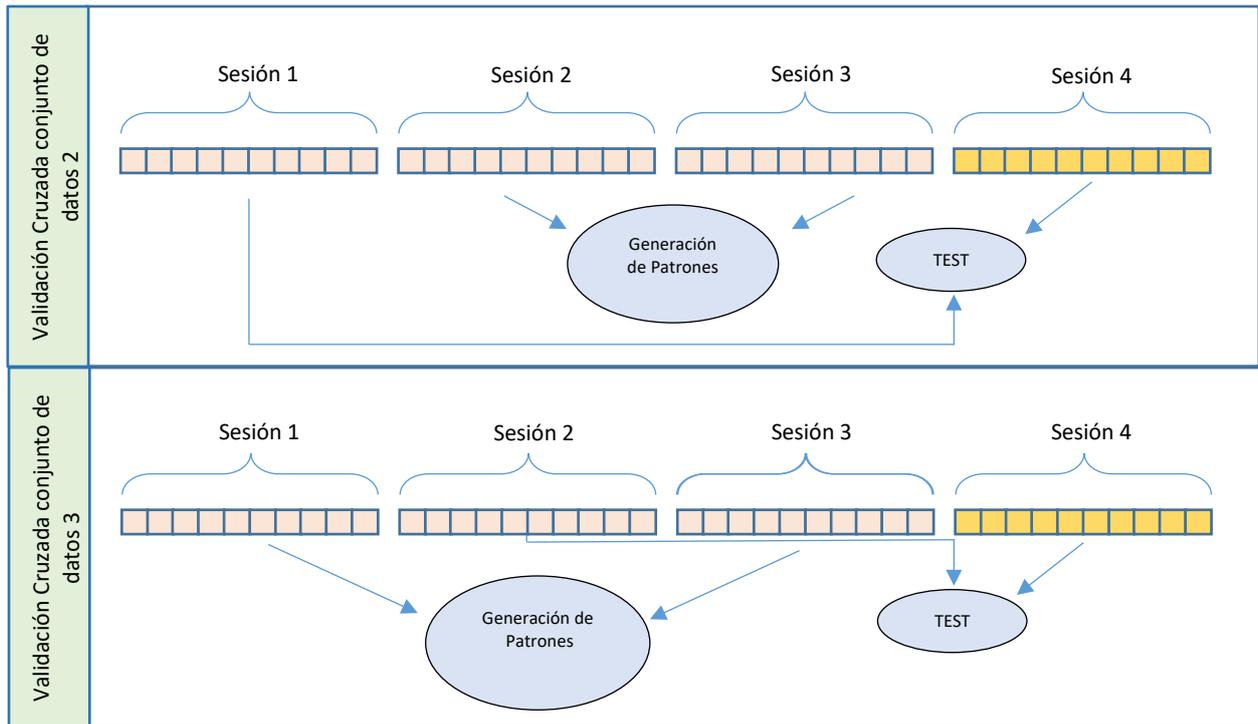


Figura 4.2. Procedimiento utilizado para el testeo de las muestras y ejecución de la validación cruzada

### 4.3 Sistema implementado

El sistema físico implementado (véase la Figura 4.3), tiene 2 partes. La primera, es el bloque de entrenamiento. Aquí se cogen 20 muestras de las 30 existentes de las grabaciones de la persona experta, y se generan los patrones. Estos patrones son una de las entradas al clasificador.

El segundo bloque, lo componen las “muestras comparables”. Dentro de las “muestras comparables” existen 2 tipos, las 10 muestras generadas por la persona experta en lenguaje de signos y las 10 muestras de la 4ª sesión generada por las 15 personas anónimas. Ambas muestras, entran de forma independiente al clasificador usando el mismo patrón, obteniendo una salida en forma de matriz de confusión.

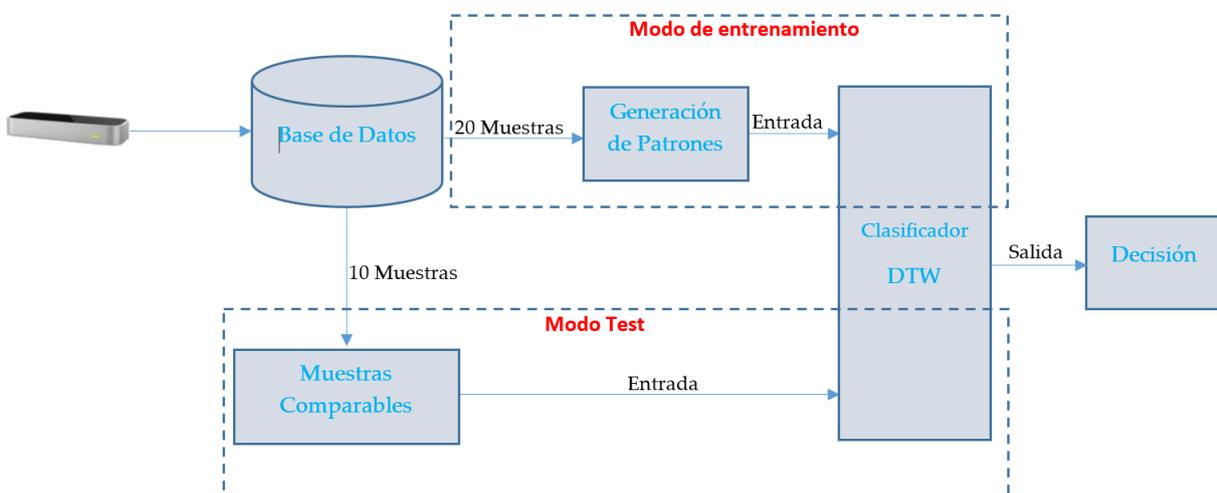


Figura 4.3. Sistema Implementado

## 4. Metodología Experimental

### 4.4 Calidad del sistema

La calidad puede definirse como la conformidad relativa con las especificaciones planteadas en la hipótesis, a lo que al grado en que esta tesis cumple las especificaciones de diseño y trazabilidad. En esta investigación, para cuantificar que los experimentos realizados con las 4 sesiones de palabras, cumplen con esos criterios, se han usado las librerías Python *Scikit-Learn*. Estas librerías extraen las coincidencias o confusiones de las muestras con sus patrones en forma de matriz de confusión.

#### 4.4.1 Matriz de confusión bi-clase

En un sistema Bi-clase partimos de una matriz de 2x2, representada en la Figura 4.4. La entrada que se quiere verificar, "Real" corresponde a las filas y lo que se ha predicho o sea la salida "Predicción" corresponde a las columnas. Cada una de las celdas corresponde a:

- Falso Positivo [FP]; Es cuando se espera una salida positiva y ésta no ocurre
- Falso Negativo [FN]; Es cuando se espera una salida negativa y ésta no ocurre
- Verdadero Positivo [VP]; Es cuando se espera una salida positiva y ésta ocurre
- Verdadero Negativo [VN]; Es cuando se espera una salida negativa y ésta ocurre

		Predicción	
		VN	FP
Real	VN		
	FN		

Figura 4.4. Modelo gráfico de los parámetros

#### 4.4.2. Precisión

La métrica de precisión muestra la proporción de verdaderos positivos frente a falsos negativos. Esta medida es muy útil para identificar la capacidad que tiene el sistema de predecir la salida, (véase la Figura 4.5).

		Predicción	
		VN	FP
Real	VN		
	FN		

Figura 4.5. Modelo gráfico del parámetro de precisión

$$precisión = \frac{VP}{VP + FP} \times 100 \quad (4.1)$$

#### 4.4.3 Recuperación o "Recall"

Esta medida muestra la proporción de verdaderos positivos frente a falsos negativos, sobre la cantidad de elementos que el modelo es capaz de identificar, (véase la Figura 4.6).

		Predicción	
		VN	FP
Real	VN		
	FN		

Figura 4.6. Modelo gráfico del parámetro de recuperación o "recall"

$$\text{recuperación} = \frac{VP}{VP + FN} \times 100 \quad (4.2)$$

#### 4.4.4 F1

Esta función combina los parámetros de precisión y recuperación. Con esta medida se compara el rendimiento de los 2 parámetros. Se calcula realizando la media armónica de precisión y recuperación.

$$F1 = 2 \cdot \frac{\text{precisión} \cdot \text{recuperación}}{\text{precisión} + \text{recuperación}} \times 100 \quad (4.3)$$

#### 4.4.5 Exactitud

Este parámetro mide el porcentaje de éxito del sistema y su fórmula corresponde con:

$$\text{exactitud} = \frac{VP + VN}{VP + VN + FP + FN} \times 100 \quad (4.4)$$

#### 4.4.2 Matriz de confusión multi-clase

Cuando existe un clasificador multiclase, la complicación aumenta porque se tendría que calcular la "precisión" y el "recall" para cada una de las palabras. Por consiguiente, la tarea se hace más ardua a la hora de encontrar un único valor numérico que refleje el buen desempeño del clasificador. En esta tesis se ha empleado el método "macro average", que consiste en calcular el promedio simple de las métricas de cada clase o palabra.

En este caso, se parte de una matriz de confusión multiclase (véase Figura 4.7), como  $MC(i,j)$  de  $N \times N$  que tiene 2 ejes, donde en cada fila está la palabra a predecir y en cada una de las columnas se encuentra el resultado predicho.

#### 4. Metodología Experimental

Real=1	VP	FN	FN	FN	FN
Real=2	FP	VN	VN	VN	VN
Real = ...	FP	VN	....	VN	VN
Real=N-1	FP	VN	VN	VN	VN
Real=N	FP	VN	VN	VN	VN

Predicho =1	Predicho =2	Predicho = ...	Predicho =N-1	Predicho =N
-------------	-------------	----------------	---------------	-------------

(a)

Real=1	VN	VN	VN	FP	VN
Real=2	VN	VN	VN	FP	VN
Real = ...	VN	VN	....	FP	VN
Real=N-1	FN	FN	FN	VP	FN
Real=N	VN	VN	VN	FP	VN

Predicho =1	Predicho =2	Predicho = ...	Predicho =N-1	Predicho =N
-------------	-------------	----------------	---------------	-------------

(b)

**Figura 4.7.** Matriz de confusión multiclase. (a) Clase 1. (b) Clase N-1

La “precisión” *macro average* se calcula usando la ecuación 4.5 y sustituyendo la “precisión” por la ecuación 4.1:

$$MacroAvgPrecisión = \frac{\sum_{j=1}^N precisión_j}{N} \quad (4.5)$$

Para calcular la *macro average* recuperación o “recall” se calcula a través de la ecuación 4.6 y sustituyendo la “recuperación” por la ecuación 4.2:

$$MacroAvgRecuperación = \frac{\sum_{i=1}^N precisión_i}{N} \quad (4.6)$$

Para calcular el *macro average* F1 se debe usar la ecuación 4.7 donde se calcula la relación existente entre la “precisión” y la “recuperación”.

$$MacroAvgF1 = 2 \cdot \frac{MacroAvgPrecisión \cdot MacroAvgRecuperación}{MacroAvgPrecisión + MacroAvgRecuperación} \quad (4.7)$$

Por último, para el cálculo de la “exactitud” se ha empleado la ecuación 4.8 como la media las exactitudes de cada palabra, sustituyendo “exactitud” por la ecuación 4.4:

$$exactitud = \frac{\sum_{i=1}^N exactitud_i}{N} \quad (4.8)$$

En el caso de esta investigación, un ejemplo de la matriz que se va usar está representada en la figura 4.8. En las filas y en las columnas se muestran cada una de las palabras, en las filas se muestra las palabras a predecir y en las columnas las palabras predichas.

En una matriz de confusión en la diagonal se muestra el éxito de las palabras. Los recuadros verdes corresponden a los aciertos entre 8 y 10, ambos inclusive, mientras que los recuadros naranjas muestran una tasa de aciertos entre 5 y 7, ambos inclusive. El rojo indica que el número de aciertos es inferior a 4. Todos los cuadros de color rojo que están fuera de la diagonal, en el cruce de una palabra con otra, no se consideran aciertos.

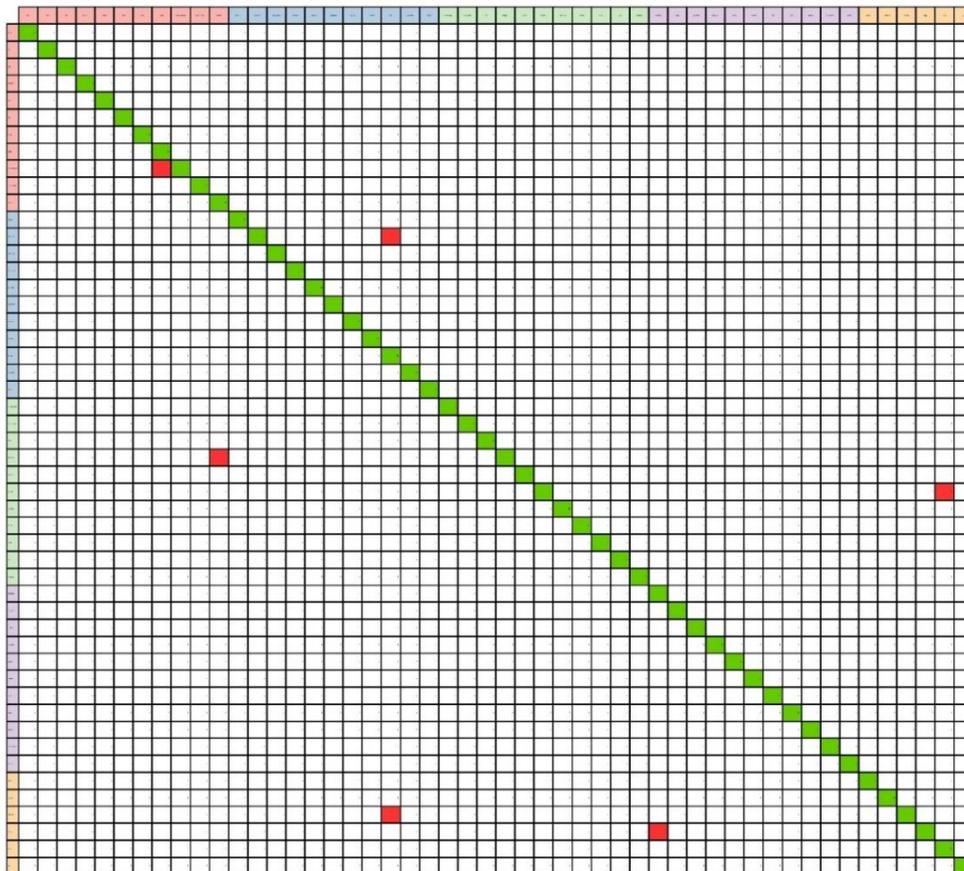


Figura 4.8. Matriz de confusión de 50 palabras

#### 4.5 Procedimiento

En primer paso que se dio, fue buscar información acerca de las diferentes investigaciones que se estaban llevando a cabo, y comprobar que se podía aportar a lo que ya existía. Se hizo un estudio exhaustivo sobre el estado del arte y a partir de ahí se empezaron a tomar decisiones.

Uno de esas decisiones, fue la de establecer un protocolo de actuación. Después del estado del arte, se consideró que el número de palabras debería ser superior a 100 para que el sistema se considerara robusto. También era imprescindible que hubiesen más de una persona signando y comparando cada una de las palabras.

El siguiente paso, fue ponerse en contacto con la asociación de personas sordas ASOR Las Palmas, para solicitar ayuda para las grabaciones de los signos, y la localización de una persona experta en lenguaje de signos. Una vez que se habían ejecutado los 2 primeros pasos, entonces, se desarrolló a partir del *firmware* del *Leap Motion*, un software para la captura de los movimientos de las manos.

#### **4. Metodología Experimental**

---

Con dicho software se empezaron a grabar los primeros signos: buenos días, buenas tardes, buenas noches y hola y se empezaron a comparar diferentes muestras sin patrones, para constatar que se estaba realizando el trabajo de forma correcta. El primer problema que apareció, fue enseñar a la persona experta como debía grabar las muestras para que fueran válidas y que ella misma pudiese comprobar si estaban realizadas de forma correcta. El segundo problema surgió con el tratamiento del fichero y el tamaño de los mismos. Este problema fue tratado con posterioridad, cuando se empezaron a extraer los parámetros del fichero y se concluyó que no todos esos parámetros eran necesarios para interpretar un signo.

A continuación, se implementó el sistema como se ha explicado en el apartado 3.7 de esta tesis. En este punto surgieron algunas cuestiones. Una de ellas, era el cómo se tenía que entrenar el sistema, es decir, cuantas muestras eran necesarias para generar un patrón y al mismo tiempo, se pudiese realizar un test del mismo. Entonces, observando otras investigaciones y calculando el tiempo que se necesitaba para obtener una muestra, se estableció que con 20 muestras era suficiente para generar un patrón robusto. También se consideró que con 10 muestras era suficiente para la comprobación de resultados, y al mismo tiempo, se observó que, con ese número de sesiones, se podrían generar los parámetros de calidad del sistema usando una validación cruzada.

Para la generación de patrones, se buscó la forma más simple y completa para que no se perdiese información del signo. Se optó por usar el método expuesto en el apartado 3.5.2 de esta tesis. A partir de ese momento, se empezó a entrenar el sistema y valorar los diferentes resultados, al mismo tiempo que se afinaba en el número de parámetros necesarios de cada muestra, para así alcanzar el objetivo marcado de 176 palabras y poder establecer una conversación médica entre paciente y doctor. Se decidieron el conjunto de palabras de la Tabla 3.1, porque encaraban bien esa conversación entre médico y paciente. Se preestableció un marco para grabar muestras de diferentes personas, y así establecer un sistema robusto e independiente del usuario. Cuando se estaban realizando las grabaciones y las comprobaciones, llegó la pandemia por el COVID-19, y el confinamiento. Esto supuso dar un giro a los objetivos marcados y supuso un cambio a la hora de grabar con distintas personas. Por lo tanto, al final se tomó la determinación de que fuese una sola persona la que grabase todas las muestras, es decir la persona experta.

Se hicieron conjuntos de 50 palabras para comprobar la evolución del sistema. Las pruebas se realizaron en paquetes de 50, 100 y finalmente, de 176 palabras.

Más adelante, se consideró la posibilidad de darle más robustez al trabajo incluyendo una cuarta sesión independiente del primer usuario. Entonces, se estableció un protocolo en el cual se realizaron grabaciones con 15 personas diferentes, y donde cada una de las palabras fueran signadas por al menos 3 distintas. De esta forma se estableció esa cuarta sesión y se perfilaron las 50 palabras de muestras externas que se incluyen en la investigación. Las palabras se seleccionaron teniendo en cuenta en primer lugar la aleatoriedad, y en segundo lugar que fuesen signos simples de gesticular, ya que las personas que decidieron participar de forma anónima en esta sesión, no eran signantes especialistas.

# Capítulo 5:

# Resultados



## 5. Resultados

### 5.1 Introducción

Los resultados de esta investigación se han obtenido siguiendo los procedimientos expuestos en capítulos anteriores, y trazando una línea de trabajo progresiva que se expone en cada uno de los apartados. En el siguiente apartado, se muestran los resultados de las validaciones entre los patrones generados por la persona especialista y las “muestras comparativas”. Además, se han añadido las muestras de la cuarta sesión generada por 15 personas diferentes. Finalmente se expone el análisis global de los resultados analizando la eficacia del sistema y la discusión con respecto a otras investigaciones.

### 5.2 Validación Cruzada

#### 5.2.1 Conjunto de palabras 1

Es necesario comprobar tanto el éxito el conjunto de palabras como su evolución a medida que se añaden nuevas palabras. La primera premisa tener en cuenta para la observación de datos, es que un cruce corresponde a una confusión de dos palabras. Se alcanzó una eficiencia del 98,80% para el conjunto de las 50 primeras palabras.

Los parámetros de calidad obtenidos a partir de la matriz, (véase la Figura 5.1), se pueden observar de forma numérica en la Tabla 5.1. Este resultado que se acerca mucho al 100 % de exactitud, es muy superior a las otras investigaciones expuestas en el capítulo 1 y enumeradas en la Tabla 5.14 de este capítulo.

*Tabla 5.1. Parámetros de calidad del conjunto de palabras 1*

Nº Palabras	Precisión	Recuperación	F1	Exactitud
50	98,93%	98,80%	98,80%	98,80%
100	97,96%	97,60%	97,59%	97,60%
176	96,32%	95,17%	95,02%	95,17%

Con 100 palabras los aciertos descendieron hasta el 97,60%. Los parámetros obtenidos se extrajeron de la matriz de confusión localizada en el anexo I, Figura AI.1.

Finalmente, se probó con el conjunto total de la base de datos de 176 palabras, bajando la exactitud hasta el 95,17%. Estos datos se extrajeron de la matriz de confusión de la Figura AI.2 del anexo I.

Como se explicó en el capítulo tres de esta tesis, se creó un conjunto de muestras externas independientes, de la persona especialista que signaba. Se grabaron cincuenta palabras con 10 muestras por cada una de ellas, y ejecutada por 15 usuarios anónimos. Los usuarios anónimos son distintos de la persona que grabó las primeras sesiones. Además, para esta cuarta sesión, se usaron los modelos generados para cada uno de los conjuntos de datos de la validación cruzada. La eficiencia del experimento este conjunto de palabras, mostrados en la Tabla 5.2, logró un 93,80%.

## 5. Resultados

Tabla 5.2. Parámetros de calidad de las muestras de usuarios externos (conjunto 1)

Nº Palabras	Precisión	Recuperación	F1	Exactitud
50	94,80%	93,80%	93,74%	93,80%

### 5.2.2 Análisis de resultados 1

De acuerdo a la descripción aportada en el capítulo 4.4.1 de esta tesis, el comportamiento con 50 palabras signadas por la misma persona que realizó las muestras, corresponde con la Figura 5.1. En esta figura se pueden comprobar los diferentes cruces que se producen en algunas palabras. Por ejemplo, la palabra “Buenas tardes” se confunde 1 vez con “boca”, la palabra “corazón” se confunde al menos 1 vez con “flema”. La confusión de las otras 3 palabras también se produce una vez. Por ejemplo, las palabras “boca” y “corazón” tienen cierta similitud a la hora de signarse, y dicha confusión puede considerarse normal.

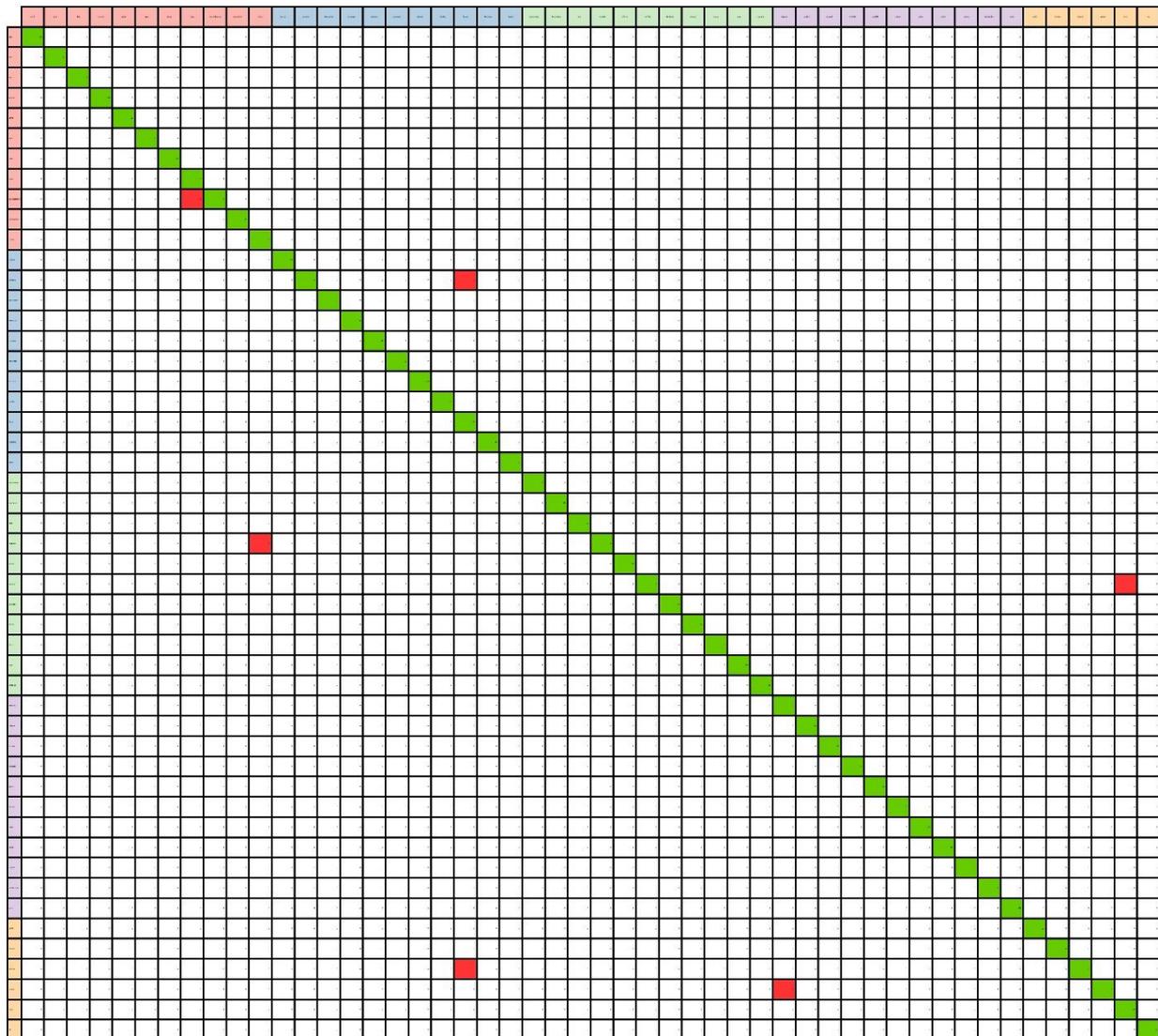


Figura 5.1. Matriz de confusión de 50 palabras conjunto 1

En la Figura 5.2 se muestra la matriz de confusión de las 50 palabras signadas en la cuarta sesión. Su eficiencia fue del 93,8% para 50 palabras, escogidas de forma aleatoria del conjunto de 176. En esta matriz los cruces de palabras fueron mayores con respecto a la matriz anterior.

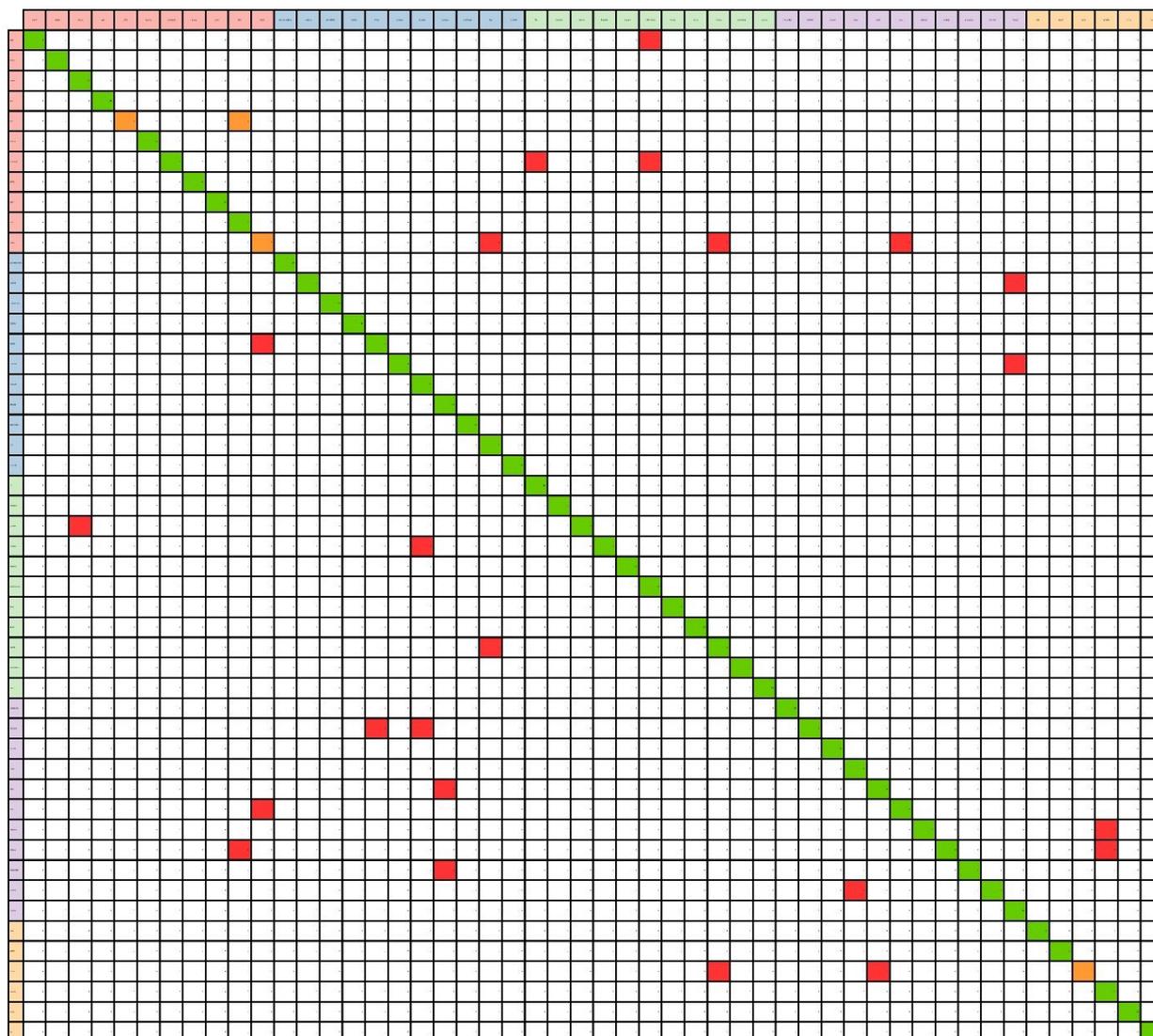


Figura 5.2. Matriz de confusión 50 palabras con muestras externas conjunto 1

Se pueden observar los cruces producidos en la Tabla 5.3, donde se comprueba como en la mayoría de los casos sólo se produce un cruce. Para estos cruces existen varias razones: Una de ellas es la similitud del signo, y otra muy importante es la propia idiosincrasia de la persona que signa, aunque el movimiento aparentemente sea el mismo, no todos movemos las manos de la misma forma y eso se traduce en una confusión de palabras por parte del sistema.

Tabla 5.3. Cruces de palabras basadas en la matriz de la Figura 5.2

Palabras y número de veces que se cruzan	
	1 2 3 4 5 6 7 8 9 10
Abril	Inflamación
Alta	Bien
Ansiedad	Fin Inflamación
Boca	Día

## 5. Resultados

	Mamá	
	Ojos	
Cabeza	Regular	
Comer	Boca	
Corazón	Regular	
Gracias	Ahora	
Hambre	Delgado	
Mama		Día
Molesto	Comer	
	Delgado	
No se		De nada
Ojos	Boca	
Oxígeno	También	
Peligro	Bien	
	También	
Preguntar	De nada	
Quemar	Nada	
Sordo	Mamá	No sé

### 5.2.3 Conjunto de palabras 2

En las primeras 50 palabras se obtuvo un acierto o exactitud del 91,2%, extrayendo los parámetros de calidad a partir de las matrices de la Figura 5.3, y mostrando dichos datos numéricos en la tabla 5.4. Con 100 palabras la exactitud descendió hasta el 85,70%, extrayendo los datos a partir de la Figura AI.3 del anexo I. El sistema bajó aproximadamente 5 puntos porcentuales. Finalmente, se probó con el conjunto total de la base de datos de 176 palabras, llegando a una exactitud del 85,39% siendo éste un valor similar si se compara con las 100 palabras. Estos datos fueron extraídos a partir de la Figura AI.4 del anexo I.

**Tabla 5.4.** Parámetros de calidad del conjunto de palabras 2

Nº Palabras	Precisión	Recuperación	F1	Exactitud
50	92,54%	91,12%	89,28%	91,12%
100	86,86%	85,70%	83,49%	85,70%
176	88,06%	85,39%	83,60%	85,39%

Con respecto a los parámetros obtenidos de la cuarta sesión, (véase la Figura 5.4), y de sus respectivos datos numéricos, (véase la Tabla 5.5), los aciertos alcanzaron el 84,40%.

**Tabla 5.5.** Parámetros de calidad de las muestras de usuarios externos (conjunto 2)

Nº Palabras	Precisión	Recuperación	F1	Exactitud
50	85,22%	84,40%	82,10%	84,40%

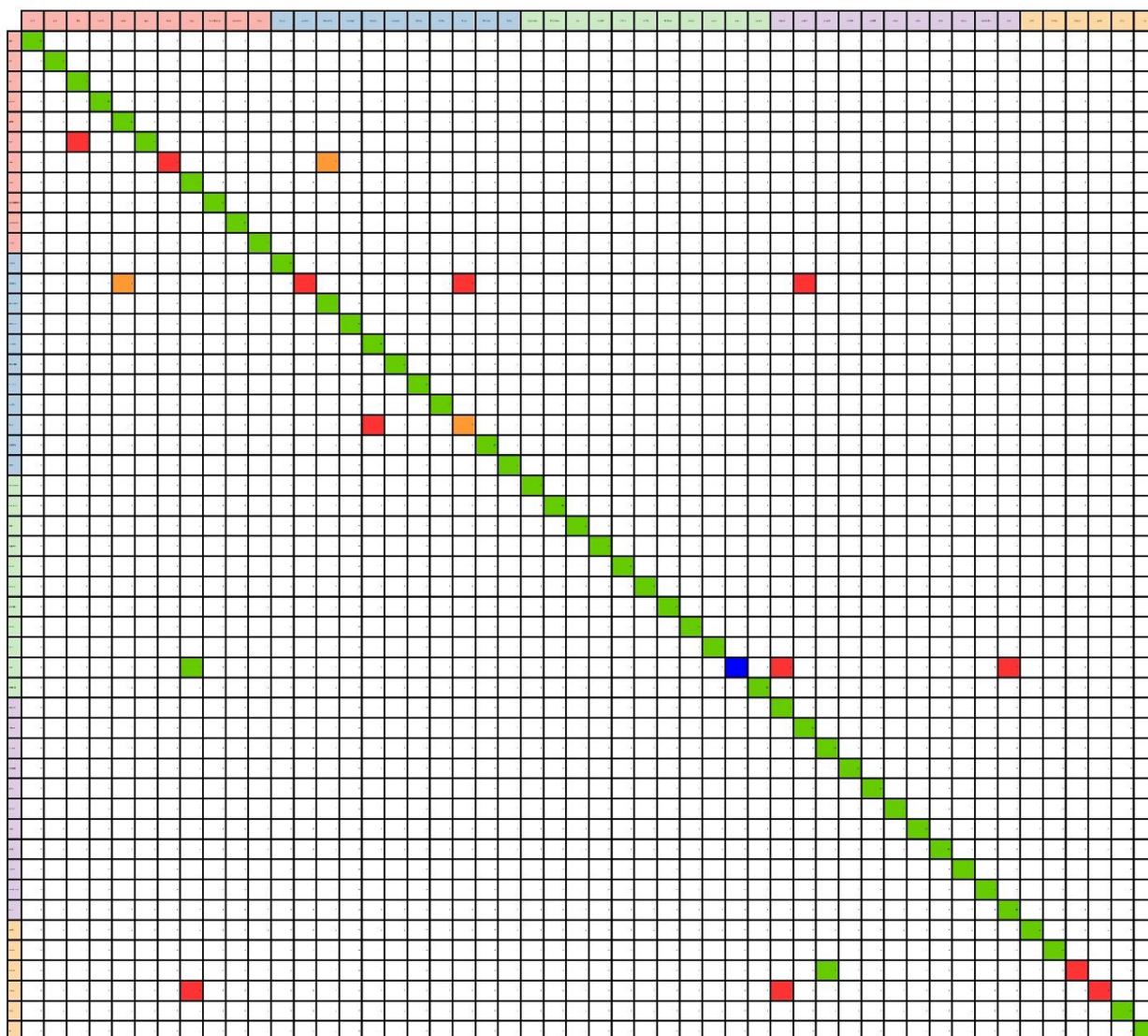


Figura 5.3. Matriz de confusión de 50 palabras conjunto 2

### 5.2.4 Análisis de resultados 2

Con respecto a los cruces generados por la matriz de la Figura 5.3, y expuestos en la Tabla 5.6, se pueden sacar varias conclusiones. Por ejemplo, la palabra “ojos” no fue reconocida ninguna vez, siendo ésta hasta 8 veces reconocida por la palabra “boca”. Dicha palabra es similar a “ojos”. Otras palabras que se asemejan son “ayer” y “alta”, “corazón” y “asma”. Como conclusión final de este análisis es que solamente se cruzan 7 palabras del conjunto de 50.

Tabla 5.6. Cruces de palabras basadas en la matriz de la Figura 5.3

		Palabras y número de veces que se cruzan									
		1	2	3	4	5	6	7	8	9	10
Ayer	Alta										
Bata	De acuerdo										

## 5. Resultados

Corazón	Peligro	Flema	Asma
Flema	Ejemplo		
Ojos	Párpado		Boca
	Sordo		
Vacuna			Privado
Vejiga	Párpado	Boca	

A la hora de analizar los resultados de la sesión cuarta, (véanse la Figura 5.4, y Tabla 5.7), los datos han sido sensiblemente más bajos, unos 6,5 puntos porcentuales, si se comparan con los anteriores. En esta matriz existen 2 palabras que nunca han sido reconocidas “abril” y “quemar”, que se muestran en la diagonal en color azul.

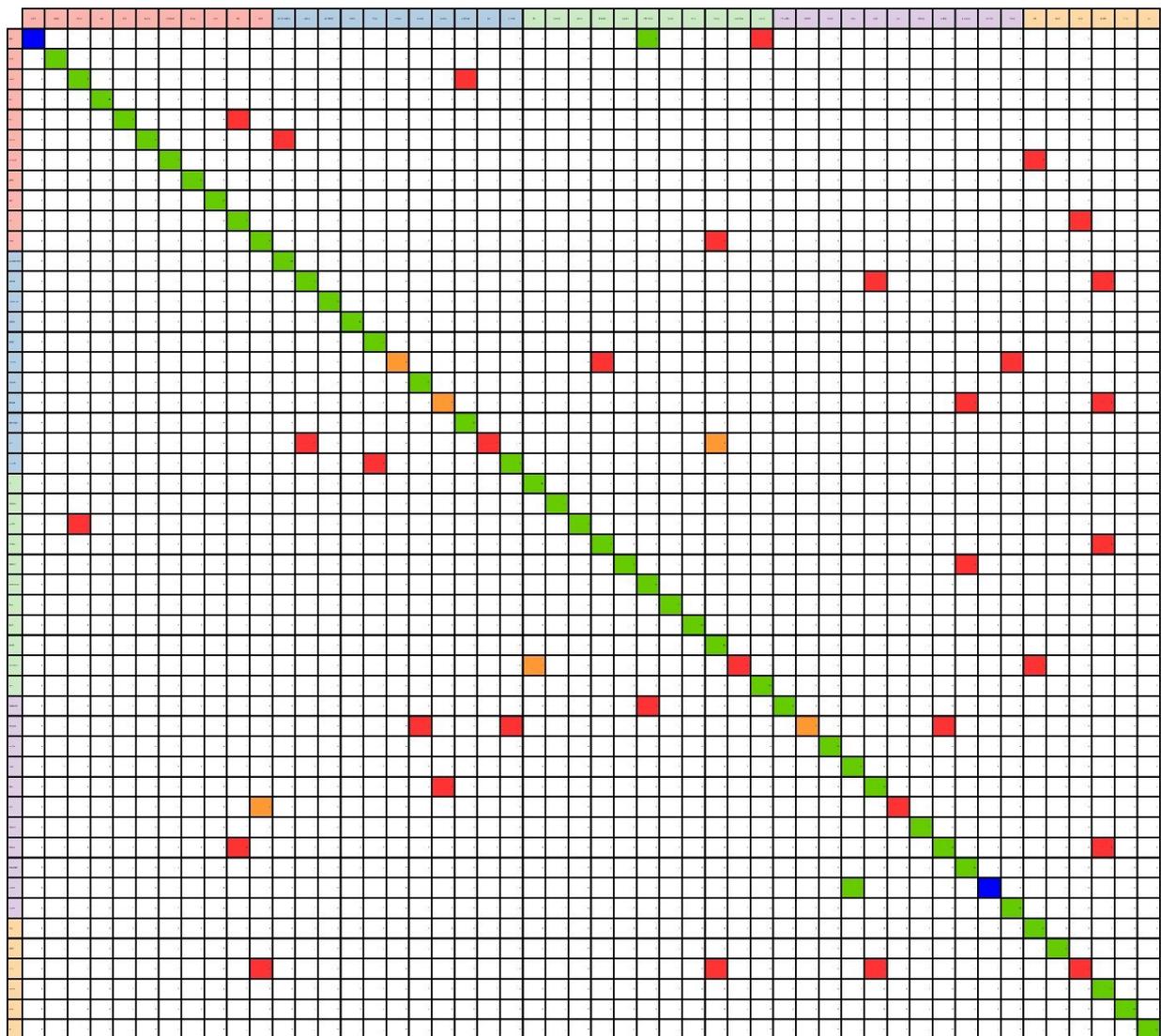


Figura 5.4. Matriz de confusión 50 palabras con muestras externas conjunto 2

Esta validación cruzada es la que arroja peores resultados con respecto a las validaciones de los otros conjuntos de palabras. Uno de los hechos que ha producido esta caída, ha sido la generación de patrones donde algunas de las palabras, tenían una o varias muestras erróneas. Esos errores vienen provocados por

una grabación deficiente, debido a varias circunstancias: inicio de la grabación con las manos mal posicionadas, error en la parada de la grabación si el usuario no mantenía las manos inmóviles o grabación errónea en disco.

Los resultados pormenorizados con respecto a los cruces se pueden visualizar en la Tabla 5.7. En ésta se pueden observar que existen bastantes más cruces que en el experimento anterior. En este caso además existen muchas palabras semejantes que se confunden por ejemplo “abril” e “inflamación”, “anoche” y “buenas noches”, “boca” y “mamá”, “cabeza” y “no sé” y por último “mascarilla” e “inflamación”.

**Tabla 5.7.** Cruces de palabras basadas en la matriz de la figura 5.4

Palabras y número de veces que se cruzan										
	1	2	3	4	5	6	7	8	9	10
Abril		Marzo						Inflamación		
Inflamación		Marzo								
Ahora	Descansar									
Alta		Bien								
Anoche	Buenas noches									
Ansiedad	Roto									
Bien	Sordo									
Boca		Mamá								
Cabeza	No sé También									
Corazón	Regular		Hambre							
De nada	También	Preguntar								
Día			Cabeza			Mamá				
Ejemplo	Comer									
Gracias	Ahora									
Hambre	También									
Hombre	Preguntar									
Mandíbula		Roto			Fin					
Mascarilla		Inflamación								
Molesto	Delgado Ejemplo Peligro									
No sé	De nada									
Ojos								Boca		
Peligro	Bien También									
Sordo	Mamá	No sé	Boca							

## 5. Resultados

---

### 5.2.5 Conjunto de palabras 3

En la última validación del conjunto de 50 palabras, se obtuvo un acierto del 94,6%. Los diferentes parámetros obtenidos se pueden visualizar en la Figura 5.5 y en la Tabla 5.8. En la validación de 100 palabras la exactitud descendió hasta el 90,90%, según los parámetros extraídos a partir de en la Figura AI.5 del anexo I, siendo unos 4 puntos porcentuales más bajo con el doble de palabras. Con el conjunto total de palabras el sistema alcanzó un acierto de 88,75%, es decir, unos 2 puntos inferior, pero con 76 palabras más. Los datos de este último experimento se obtuvieron a partir de la matriz de confusión de la Figura AI.6 del anexo I.

Tabla 5.8. Parámetros de calidad del conjunto de palabras 3

Nº Palabras	Precisión	Recuperación	F1	Exactitud
50	97,83%	94,60%	93,86%	94,60%
100	91,16%	90,90%	89,62%	90,90%
176	90,45%	88,75%	87,73%	88,75%

Siguiendo con la misma dinámica en la exposición de resultados, en la cuarta sesión se obtuvo un acierto del 87,20%, con los datos extraídos de la Figura 5.6 y enumerados en la Tabla 5.9.

Tabla 5.9. Parámetros de calidad de las muestras de usuarios externos (conjunto 3)

Nº Palabras	Precisión	Recuperación	F1	Exactitud
50	88,20%	87,20%	85,92%	87,20%

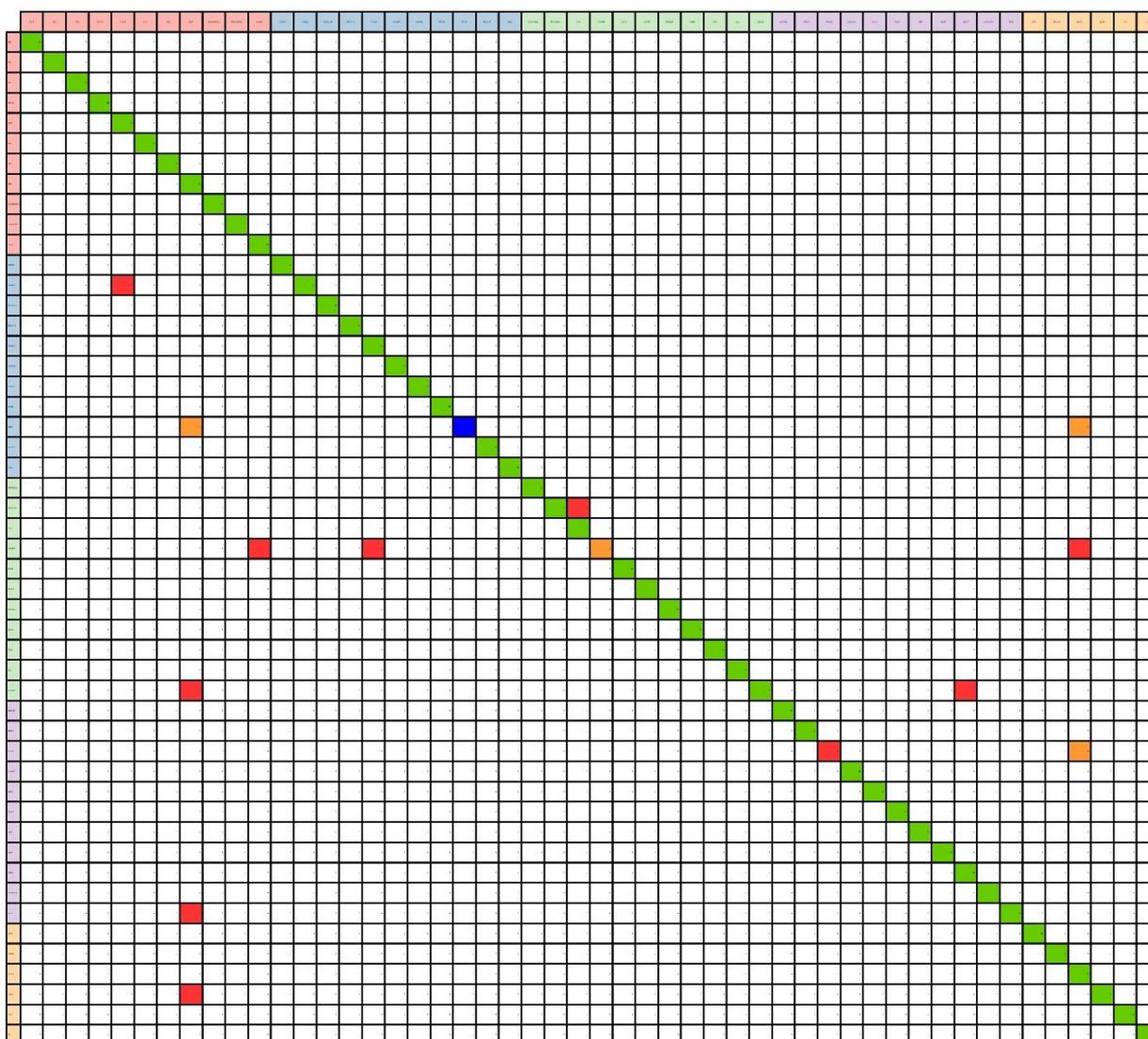


Figura 5.5. Matriz de confusión de 50 palabras conjunto 3

### 5.2.6 Análisis de resultados 3

En esta validación, los resultados son semejantes a los de la primera validación cruzada. Los datos mostrados en la Tabla 5.4, revelan las diferentes confusiones que se producen entre alguna de ellas, donde además se puede observar como la palabra “flema” no se reconoce nunca.

Con respecto a los cruces generados por la matriz de la Figura 5.5, se pueden extraer las siguientes conclusiones mostrados en la Tabla 5.10. Destacar que la palabra “flema” no fue reconocida ninguna vez, siendo hasta 5 veces reconocida por las palabras “boca” y “vacuna”. Es importante destacar como al final son sólo 8 palabras de las 50 las que se confunden con otras palabras, destacando que 4 de ellas solamente se cruzan una vez con otra palabra.

Tabla 5.10. Cruces de palabras basadas en la matriz de la figura 5.5

Palabras y número de veces que se cruzan									
1	2	3	4	5	6	7	8	9	10

## 5. Resultados

Corazón	Asma		
Flema			Boca Vacuna
Hinchazón	Hora		
Hospital	Cáncer Vacuna	Ejemplo	
Órganos	Boca Seguro		
Privado			Vacuna
Sordo	Boca		
Vejiga	Boca		

Los resultados obtenidos con la cuarta sesión, son más bajos que los datos arrojados por la persona especialista. Son 7 puntos de diferencia entre ambas, (véase la Figura 5.6). Esto sugiere la importancia no solo de crear patrones precisos, sino también de asegurar que los movimientos de las manos sean lo más explícitos posible. En esta matriz la palabra “día”, por ejemplo, no ha sido reconocida nunca.

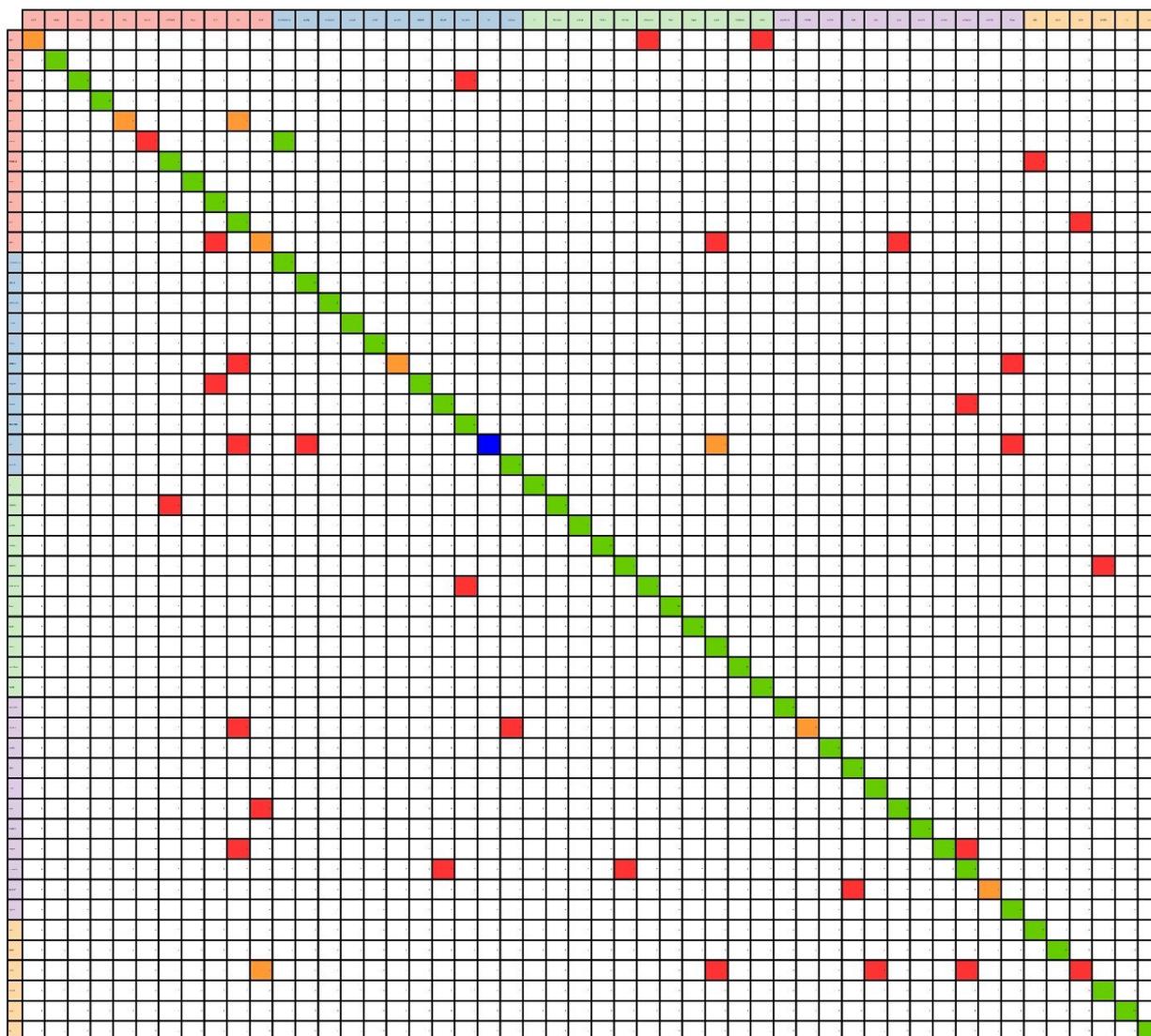


Figura 5.6. Matriz de confusión 50 palabras con muestras externas conjunto 3

Si se observa la Tabla 5.11, se pueden sacar algunas conclusiones interesantes. En primer lugar, ha habido 20 palabras que se han confundido al menos una vez con otra palabra, y donde 9 de ellas se han confundido sólo con 1 palabra. También es necesario considerar la similitud entre diferentes palabras signadas, por ejemplo, la palabra “abril” es semejante a las palabras “inflamación” y “marzo”. Otro ejemplo de similitud entre signos se puede encontrar entre las palabras “anoche” y “buenas noches”, “inflamación” y “descansar” o entre “sordo” y “no sé”.

Tabla 5.11. Cruces de palabras basadas en la matriz de la figura 5.6

Palabras y número de veces que se cruzan										
	1	2	3	4	5	6	7	8	9	10
Abril	Inflamación	Marzo								
Ahora		Descansar								
Alta					Bien					

## 5. Resultados

Anoche				Buenas noches
Ansiedad	Roto			
Bien	Sordo			
Boca	Ayer Mamá		Ojos	
Corazón	Regular	Bien		
Delgado	Ayer			
De nada		Preguntar		
Día	Bien Regular		Cabeza	Mamá
Funciona	Ansiedad			
Hombre	También			
Inflamación	Descansar			
Molesto	Ejemplo	Bien		
Ojos		Boca		
Peligro	Bien Preguntar			
Preguntar	De nada Hombre			
Quemar			Nada	
Sordo	Mamá No sé Preguntar			Boca

### 5.3 Resultado global de la validación cruzada

En el análisis final se realizó la media de los 3 experimentos, enumerados en las Tablas 5.12 y 5.13. Estas tablas muestran que la precisión promedio para 176 palabras es casi del 90% para la misma persona y aproximadamente del 88,50% para la cuarta sesión independiente. Además, en las mismas tablas se pueden contemplar las variaciones de cada una de las validaciones con respecto a la media.

**Tabla 5.12.** Media de los parámetros de calidad de la validación cruzada

Conjunto de palabras	Nº Palabras	Precisión	Variación	Recuperación	Variación	F1	Variación	Exactitud	Variación
Conjunto 1	50	96,42%	-2,51	94,84%	-3,96	93,98%	-4,82	94,84%	-3,96
Conjunto 2			+3,88		+3,72		+4,70		+3,72
Conjunto 3			-1,41		+0,24		+0,12		+0,24
Conjunto 1	100	91,99%	-5,97	91,40%	-6,20	90,23%	-7,36	91,40%	-6,20
Conjunto 2			+5,13		+5,70		+6,74		+5,70
Conjunto 3			+0,83		+0,50		+0,61		+0,50
Conjunto 1	176	91,61%	-4,71	89,77%	-5,40	88,78%	-6,24	89,77%	-5,40
Conjunto 2			+3,55		+4,38		+5,18		+4,38
Conjunto 3			+1,16		+1,02		+1,05		+1,02

Tabla 5.13. Media de los parámetros de calidad de la validación cruzada de las muestras externas

Conjunto de palabras	Nº Palabras	Precisión	Variación	Recuperación	Variación	F1	Variación	Exactitud	Variación
Conjunto 1			-5,40		-5,34		-6,49		-5,34
Conjunto 2	50	89,40%	+4,18	88,46%	+4,06	87,25%	+5,15	88,46%	+4,06
Conjunto 3			+1,20		+1,26		+1,33		+1,26

#### 5.4 Discusión

Como análisis general los resultados muestran que la eficiencia del sistema está por encima del 89,77% en todas las fases de la validación cruzada. Los experimentos que utilizan esta técnica y donde se pueden comparar de forma más directa son, (Chong y Lee, 2018; Mapari y Kharat, 2016; Mohandes, Aliyu, et al., 2014; Simos y Nikolaidis, 2016; Vaitkevičius et al., 2019). La comparativa de este sistema con el resto, se ha hecho con respecto a la exactitud de los experimentos, al número de palabras y al tamaño de las muestras.

A medida que se agregan palabras, la eficiencia del sistema disminuye, lo que puede convertirse en un problema para reconocer algunas palabras. La independencia del sistema se validó con la cuarta sesión, donde los parámetros de calidad descendieron un 6,4% en aciertos, un 7,02% en precisión y un 5,69% para F1 entre el conjunto de palabras de la persona especialista y los usuarios anónimos lo cual asegura una invariancia entre usuarios. Además, la “precisión” siempre está 1 y 2 puntos porcentuales por encima de la exactitud. Este dato es relevante, porque dicho parámetro predice de forma positiva la probabilidad de que la palabra sea encontrada de forma correcta y exista un acierto. El parámetro de “recuperación” coincide con la “exactitud” porque en una matriz multiclase donde sólo se tienen en cuenta las filas y las columnas, un verdadero negativo (VN) no existe para el cálculo de la misma.

Para qué fuese más fácil de comparar los proyectos de investigación que se han estudiado, éstos se han dividido en tres tipos, uso de signos dinámicos, uso de signos estáticos y uso de signos dinámicos-estáticos. En ella se puede comprobar la limitación en el número de signos y la propuesta de esta investigación, donde se propone un sistema basado en un generador de patrones usando el “*Dynamic Time Warping*” (DTW), para demostrar que se puede aumentar el número de palabras manteniendo las tasas de éxito.

Los resultados obtenidos en esta tesis con los de otras publicaciones comentadas en el capítulo uno apartado de *Trabajos Relativos*, se pueden contemplar en la Tabla 5.14. Destacar que existen algunas diferencias con respecto a las investigaciones mostradas en la Tabla 5.14. Por ejemplo, existen trabajos como (Simos y Nikolaidis, 2016), que está por encima del 99% pero que reconoce 24 signos estáticos que es un número bastante inferior a las 176 palabras de esta investigación. Otras investigaciones que están por encima del 95% de precisión (Alnahhas y Alkhatib, 2020; Anwar et al., 2017; Avola et al., 2019; C. K. M. Lee et al., 2021; Mohandes, Aliyu, et al., 2014; Mohandes, Deriche, et al., 2014; Naglot y Kulkarni, 2016), tienen una mezcla de signos estáticos y dinámicos, y donde la investigación que más palabras reconoce tiene 40 signos, de los cuales 18 son estáticos y 12 son dinámicos. Otro trabajo a tener en cuenta es (Hisham y Hamouda,

## 5. Resultados

2018), que usa la misma técnica que esta tesis en el reconocimiento de signos dinámicos, el DTW pero, que se diferencia de esta investigación en 2 cuestiones muy importantes. La primera que usa solamente 12 signos dinámicos y la segunda que no usa un procedimiento de validación cruzada para la verificación de muestras. Con respecto a los signos dinámicos cabe señalar la propuesta de (Elons et al., 2014) que trabaja con 50 palabras y 4 personas distintas y sin validación cruzada, y donde la precisión está sobre el 88%, es decir, por debajo de esta investigación en 50 palabras y una persona 94,86% y por debajo de 50 palabras y 15 personas diferentes, 88,29%. El trabajo de (Jenkins, 2021), tiene valores de reconocimiento de entre el 96,4% y el 99,9% usando diferentes técnicas de clasificación. Esta investigación tiene 2 cuestiones diferenciadoras a tener en cuenta. La primera es que no usa validación cruzada y la segunda que sólo trabaja con 13 palabras del lenguaje de signos americano.

Tabla 5.14. Trabajos relativos usando Leap Motion

Referencia	Gestos	Conjunto de Datos	Clasificador	Exactitud
<b>Signos Estáticos</b>				
(Funasaka et al., 2015)	24 Letras del Lenguaje de Signos Americano	No aclarado	Árbol de Decisión	82,71%
(Marin et al., 2014) *	10 Signos del Lenguaje Americano	14 personas × 10 conjuntos × 10 muestras	SVM	80,86%
(Simos y Nikolaidis, 2016)	24 Signos del Lenguaje de Griego	6 personas × 10 conjuntos	MLP (Traslación de los huesos)	99,08%
			MLP (Traslación de la palma de la mano)	98,96%
(Mapari y Kharat, 2016)	32 Signos del Lenguaje Americano	146 personas × 1 muestra /letra	MLP	90%
(Vaitkevicius et al., 2019)	24 Letras del Lenguaje de Signos Americano Uso de Frases	12 personas × 10 muestras	Análisis de Regresión Lineal	86,1%
<b>Signos Estáticos y Dinámicos</b>				
(Mohandes, Deriche, et al., 2014)	28 Signos del Lenguaje Árabe	10 muestras/letra	KNN	97,1%
			HMM	97,7%
(Mohandes, Aliyu, et al., 2014)	28 Signos del Lenguaje Árabe	10 muestras/letra	Naïve Bayes	98,3%
			MLP	99,1%
(Hisham y Hamouda, 2018)*	16 Palabras estáticas del Lenguaje de Signos Árabe 20 Palabras dinámicas	2 personas × 200 muestras	Estático	
			Neural Network	90,35%
			K-NN	95,22%
			SVM (RBF Kernel)	89,12%
			SVM (Multi-Kernel)	90,78%
			Dinámico	
DTW	96,41%			
(Naglot y Kulkarni, 2016)	26 Letras del Lenguaje de Signos Americano	4 personas × 20 muestras/letras	MLP-BP	96,15%
(Chong y Lee, 2018)	26 Letras del Lenguaje de Signos Americano 10 Números	12 personas No definida las muestras	SVM (letras)	80,30%
			DNN (letras)	93,81%
			SVM (total)	72,79%
			DNN (total)	88,79%
(C. K. M. Lee et al., 2021)	26 Letras del Lenguaje de Signos Americano	100 personas × 1 muestras/letra	LSTM	97,96%
			SVM	98,35%
			RNN	98,19%

(Tao et al., 2018)	26 Letras del Lenguaje de Signos Americano	5 conjuntos × 450 muestras/letra	CNN	80,1–99,7%
(Anwar et al., 2017)	26 Letras del Lenguaje de Signos Indonesio	5 personas × 10 muestras/letra	K-NN	95,15%
			SVM	93,85%
(Alnahhas y Alkhatib, 2020)	15 Palabras estáticas/dinámicas del Lenguaje de Signos Árabe	5 personas × 10 muestras/letra	LSTM	96%
(Avola et al., 2019)	18 Signos Estáticos/ 12 Signos dinámicos/	20 personas × 60 muestras /palabra	LSTM	96,41%
<b>Signos Dinámicos</b>				
(Elons et al., 2014)	50 Palabras dinámicas del Lenguaje de Signos Árabe	4 personas × 1 conjunto	MLP	88%
(Jenkins, 2021)	13 Palabras del Lenguaje de Signos Americano	10 muestras	Neural Network	99,9%
			Random Forest	99,7%
			SVM	99,9%
			K-NN	98,7%
			Naïve Bayes	96,4%
Esta investigación	50 Palabras del Lenguaje de Signos Español	1 persona × 30 muestras/palabra	DTW	94,86%
Esta investigación	100 Palabras del Lenguaje de Signos Español	1 persona × 30 muestras/palabra	DTW	91,40%
Esta investigación	176 Palabras del Lenguaje de Signos Español	1 persona × 30 muestras/palabra	DTW	89,77%
Esta investigación	50 Palabras del Lenguaje de Signos Español	15 personas diferentes × 10 muestras/palabra	DTW	88,29%

\* Trabajos que han usado la *Kinect* junto con el LMC.

Destacar que, para la creación de los patrones en esta investigación, se usaron 20 muestras de dos sesiones independientes, donde ha hecho que la calidad y eficiencia del patrón mejore. Una de las cuestiones, que se han sacado durante el desarrollo de esta investigación, es que, con un patrón de 50 muestras grabadas por distintas personas, haría que el sistema fuera más robusto mejorando la eficiencia del sistema.

El sistema desarrollado en esta investigación, mejora los resultados expuestos en la tabla anterior, destacando que este sistema mejora frente a otras técnicas de “*Deep Learning*” y “*Machine Learning*”. Los resultados con 50 signos y/o palabras, muestran una robustez frente a otros estudios.

### 5.5 Eficacia del sistema

En esta investigación se muestra un sistema de reconocimiento de palabras para la lengua de signos española, mejorando el número de palabras del estado del arte, que se encontraban entre 10 y 50 signos dinámicos. Al mismo tiempo, se ha aumentado el reconocimiento de señales dinámicas, aplicando técnicas de reconocimiento automático mediante el DTW, mejorando los otros sistemas que han usado diferentes técnicas de “*Deep Learning*” y “*Machine Learning*”, mostrándose esta propuesta más robusta que las técnicas utilizadas hasta el momento.

Destacar que la invarianza entre diferentes usuarios ha sido validada, gracias a la cuarta sesión externa, grabada por 15 personas diferentes. Este es un valor agregado de la propuesta, ya que mantiene una exactitud de un 88,46% con validación entre usuarios con respecto al 94,86% para la validación inter-

## **5. Resultados**

---

sesión con el mismo usuario con 50 palabras. Desde un punto de vista práctico, el modelo creado por un usuario o signante puede ser utilizado por cualquier otro usuario o signante.

La utilidad de este sistema también viene definida por el número de palabras necesarias para que funcione la vida real. Este número de palabras estaría en torno a las 1000.

### **5.5.1 Coste computacional**

Una de las características más destacables que tiene el sistema de reconocimiento de signos desarrollado en esta tesis, es su funcionamiento en tiempo real. El hardware que existe actualmente, es capaz de soportar no sólo la capacidad de cálculo del sistema sino su almacenamiento en memoria, haciendo que el tiempo de respuesta sea muy bajo. Esta característica, se puede trasladar a diferentes tipos de dispositivos portátiles para su uso por parte de los usuarios.

La facilidad para entrenar el sistema, es una de las características más destacadas del mismo. La grabación de muestras se puede realizar de una forma simple, donde cada usuario puede crear un patrón para su uso personal y para cada una de las palabras. Eso ayuda a que su reconocimiento sea más eficiente. Hay que recordar que, aunque exista un lenguaje de signos español, existen muchas palabras que cambian según la región o la zona, habiendo de facto varios dialectos.

# Capítulo 6: Conclusiones



## 6. Conclusiones

### 6.1 Valoración final

Una vez se ha completado todas las fases de desarrollo de la tesis, es necesario comprobar que se han cumplido las premisas planteadas al inicio de esta investigación. En primer lugar, se parte de la hipótesis planteada en el capítulo 1 de esta tesis:

*“Es posible crear un traductor del lenguaje de signos español usando un sensor económico, fiable y fácil de manejar, traduciendo los signos con independencia de la persona que signa.”*

Se puede considerar que esta hipótesis se ha validado de acuerdo con los resultados obtenidos y justificados con la metodología experimental.

También hay que señalar que el objetivo principal planteado, se ha cumplido, haciendo que el sistema funcione, sea fácil de manejar y al mismo tiempo económico.

Por último, destacar que, con respecto a los objetivos específicos, se ha creado una base de datos extensa de patrones que pueden ser utilizados en futuros trabajos. Con respecto al último objetivo logrado, señalar el uso de tecnologías *machine learning* para el reconocimiento de los diferentes signos.

### 6.2 Méritos logrados

Asociado a esta tesis, se han logrado una serie méritos que han aportado a valor esta investigación. El primer logro, fue ganar uno de los premios de III Edición de las Ayudas a Tecnologías Accesibles de Indra y Fundación Universia del año 2018, donde su asignación económica ha ayudado en gran medida a lograr avanzar esta investigación (véase la Figura C.1 del apéndice III).

Por otro lado, asociado a esta investigación, se han publicado 2 artículos que suman valor a esta investigación. El primer artículo se publicó en junio de 2020, (véase la Figura C.2 del apéndice III) donde versaba sobre el estado del arte. En enero de 2023 se publicó el segundo artículo (véase la Figura C.3 del apéndice III), que trataba sobre los experimentos y resultados obtenidos del sistema de reconocimiento del lenguaje de signos.

Además, hay que destacar el segundo premio del III Concurso Cátedra Telefónica de la Universidad de Las Palmas de Gran Canaria del año 2020, en la categoría de proyectos de transferencia (véase la Figura C.4 del apéndice III).

### 6.3 Futuros trabajos

A medida que se fue desarrollando la investigación, y empezaron a arrojar resultados, se consideró que había algunas cuestiones que de cara al futuro era necesario tener en cuenta.

## 6. Conclusiones

---

En primer lugar, si hay varias personas para la grabación de múltiples muestras, sería necesario la creación de un protocolo que marque de forma unívoca el trabajo, estableciendo algunas premisas básicas. La primera sería usar más de 50 muestras para la generación de patrones, haciendo que el sistema fuera más robusto e independiente. Otra premisa, es la participación de al menos 5 personas distintas en la generación de patrones. Con respecto a la mejora de la base de datos del sistema, sería conveniente la realización de una “verificación cruzada”, es decir, que cada una de las muestras del patrón, sea verificada por una persona distinta de la que participe en su generación. Muchas veces, por la aleatoriedad de las personas a la hora de grabar, debería existir un protocolo de grabación de muestras, para que el inicio y la finalización de cada signo sea la misma para todos los usuarios. Por último, habría que tener en cuenta que la muestra más pequeña no supere de los 800Kb con respecto a la más grande, y que de acuerdo con el apartado 3.2.2 de esta tesis las muestras no superen los 2 MB de tamaño ni sean inferior a los 350 KB.

Otra cuestión a señalar es que, las personas sordas no utilizan preposiciones en sus frases, al mismo tiempo que no siguen la misma sintaxis en la construcción de las mismas, como las personas parlantes. Para que la frase quede ordenada de forma coherente, y pueda ser entendida por una persona parlante, sería necesario crear un sistema de inteligencia artificial, que ordenase las palabras en el orden adecuado. Una solución externa para resolver esta cuestión puede ser el uso del *Chat GPT*, que puede construir dicha frase de forma automática.

Otro hecho que hay que valorar, es el comportamiento gestual del rostro que utilizan las personas sordas para poder comunicarse. De hecho, existen palabras que se signan exactamente igual, y sólo son capaces de diferenciarse por el rostro, un ejemplo pueden ser las palabras “gusto” y “dolor”. Ante esta disyuntiva tan particular, sería recomendable el uso de la cámara, que en este caso ya incorpora el *Leap Motion* y que, usando sistemas de inteligencia artificial pueda distinguir a cuál de los signos se hace referencia.

Otro aspecto a destacar en trabajos futuros, es el discernimiento entre palabras signadas con una mano o dos manos, lo cual hace necesario que el sistema diferencie de forma automática el uso de una o ambas manos, ayudando de esta forma, a que la clasificación de los signos sea más rápida y así simplificar la búsqueda y por lo tanto el reconocimiento del signo.

Por último, valorar el uso de otras librerías de DTW en Python. Hasta ahora se han usado librerías estándar, pero sería conveniente buscar posibles alternativas a esas librerías que ayuden a mejorar la velocidad y la eficacia del sistema, probando su uso en tiempo real.

# Bibliografía



---

## Bibliografía

- Abdelnasser, H., Youssef, M., & Harras, K. A. (2015). WiGest: A ubiquitous WiFi-based gesture recognition system. *2015 IEEE Conference on Computer Communications (INFOCOM)*, 1472-1480. <https://doi.org/10.1109/INFOCOM.2015.7218525>
- Abdelrazik, M. A., Zekry, A., & Mohamed, W. A. (2023). Efficient Deep Learning Algorithm for Egyptian Sign Language Recognition. *2023 33rd Conference of Open Innovations Association (FRUCT)*, 3-8. <https://doi.org/10.23919/FRUCT58615.2023.10142991>
- Abdel-Samie A.-G. A.-R., Elmisery F. A., M.Brisha A., & Khalil A. H. (2018). Arabic Sign Language Recognition Using Kinect Sensor. *Research Journal of Applied Sciences, Engineering and Technology*, 15(2), 57-67. <https://doi.org/10.19026/rjaset.15.5292>
- Abdullahi, S. B., & Chamnongthai, K. (2022). American Sign Language Words Recognition Using Spatio-Temporal Prosodic and Angle Features: A Sequential Learning Approach. *IEEE Access*, 10, 15911-15923. <https://doi.org/10.1109/ACCESS.2022.3148132>
- Abreu, J. G., Teixeira, J. M., Figueiredo, L. S., & Teichrieb, V. (2016). Evaluating Sign Language Recognition Using the Myo Armband. *2016 XVIII Symposium on Virtual and Augmented Reality (SVR)*, 64-70. <https://doi.org/10.1109/SVR.2016.21>
- Abulela, M. A. A., Mrutu, A. P., & Ismail, N. M. (2023). Learning and Study Strategies as Predictors of Undergraduates' Emotional Engagement: A Cross-Validation Approach. *SAGE Open*, 13(1), 21582440231155881. <https://doi.org/10.1177/21582440231155881>
- Adib, F., Kabelac, Z., & Katabi, D. (2015). *{Multi-Person} Localization via {RF} Body Reflections*. 279-292. <https://www.usenix.org/conference/nsdi15/technical-sessions/presentation/adib>
- Adib, F., Kabelac, Z., Katabi, D., & Miller, R. C. (2014). *3D Tracking via Body Radio Reflections*. 317-329. <https://www.usenix.org/conference/nsdi14/technical-sessions/presentation/adib>
- Adib, F., & Katabi, D. (2013). See through walls with WiFi! *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*, 75-86. <https://doi.org/10.1145/2486001.2486039>
- Adwan, S., & Arof, H. (2012). On improving Dynamic Time Warping for pattern matching. *Measurement*, 45(6), 1609-1620. <https://doi.org/10.1016/j.measurement.2012.02.015>

## **Bibliografia**

---

- Agarwal, C., Dogra, D. P., Saini, R., & Roy, P. P. (2015). Segmentation and recognition of text written in 3D using Leap motion interface. *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, 539-543. <https://doi.org/10.1109/ACPR.2015.7486561>
- Ahmed, W., Chanda, K., & Mitra, S. (2016). Vision based Hand Gesture Recognition using Dynamic Time Warping for Indian Sign Language. *2016 International Conference on Information Science (ICIS)*, 120-125. <https://doi.org/10.1109/INFOSCI.2016.7845312>
- Ahsan, R., Ibrahimy, M. I., & Khalifa, O. O. (2009). EMG Signal Classification for Human Computer Interaction: A Review. *European Journal of Scientific Research*, 33(3), 480-501.
- Aldahri, E., Aljuhani, R., Alfaidi, A., Alshehri, B., Alwadei, H., Aljojo, N., Alshutayri, A., & Almazroi, A. (2023). Arabic Sign Language Recognition Using Convolutional Neural Network and MobileNet. *Arabian Journal for Science and Engineering*, 48(2), 2147-2154. <https://doi.org/10.1007/s13369-022-07144-2>
- Alnahhas, A., & Alkhatib, B. (2020a). Enhancing The Recognition Of Arabic Sign Language By Using Deep Learning And Leap Motion Controller. *International Journal of Scientific & Technology Research*, 9(4), Article 4.
- Alnahhas, A., & Alkhatib, B. (2020b). Enhancing The Recognition Of Arabic Sign Language By Using Deep Learning And Leap Motion Controller. *International Journal of Scientific & Technology Research*, 9(4), 1865-1870.
- Amin, T. B., & Mahmood, I. (2008). Speech Recognition using Dynamic Time Warping. *2008 2nd International Conference on Advances in Space Technologies*, 74-79. <https://doi.org/10.1109/ICAST.2008.4747690>
- Angeles, R. (2005a). Rfid Technologies: Supply-Chain Applications and Implementation Issues. *Information Systems Management*, 22(1), Article 1. <https://doi.org/10.1201/1078/44912.22.1.20051201/85739.7>
- Angeles, R. (2005b). Rfid Technologies: Supply-Chain Applications and Implementation Issues. *Information Systems Management*, 22(1), 51-65. <https://doi.org/10.1201/1078/44912.22.1.20051201/85739.7>

- Anwar, A., Basuki, A., Sigit, R., Rahagiyanto, A., & Zikky, Moh. (2017a). Feature Extraction for Indonesian Sign Language (SIBI) Using Leap Motion Controller. *2017 21st International Computer Science and Engineering Conference (ICSEC)*, 1-5. <https://doi.org/10.1109/ICSEC.2017.8443926>
- Anwar, A., Basuki, A., Sigit, R., Rahagiyanto, A., & Zikky, Moh. (2017b). Feature Extraction for Indonesian Sign Language (SIBI) Using Leap Motion Controller. *2017 21st International Computer Science and Engineering Conference (ICSEC)*, 1-5. <https://doi.org/10.1109/ICSEC.2017.8443926>
- API Overview—Leap Motion JavaScript SDK v3.2 Beta documentation*. (s. f.). Recuperado 29 de agosto de 2022, de [https://developer-archive.leapmotion.com/documentation/javascript/devguide/Leap\\_Overview.html](https://developer-archive.leapmotion.com/documentation/javascript/devguide/Leap_Overview.html)
- API Overview—Leap Motion JavaScript SDK v3.2 Beta documentation*. (2022, julio 12). [https://developer-archive.leapmotion.com/documentation/javascript/devguide/Leap\\_Overview.html](https://developer-archive.leapmotion.com/documentation/javascript/devguide/Leap_Overview.html)
- Arici, T., Celebi, S., Aydin, A. S., & Temiz, T. T. (2014). Robust gesture recognition using feature pre-processing and weighted dynamic time warping. *Multimedia Tools and Applications*, 72(3), 3045-3062. <https://doi.org/10.1007/s11042-013-1591-9>
- Asadzadeh, P., Kulik, L., & Tanin, E. (2012). Gesture recognition using RFID technology. *Personal and Ubiquitous Computing*, 16(3), Article 3. <https://doi.org/10.1007/s00779-011-0395-z>
- Avola, D., Bernardi, M., Cinque, L., Foresti, G. L., & Massaroni, C. (2019a). Exploiting Recurrent Neural Networks and Leap Motion Controller for the Recognition of Sign Language and Semaphore Hand Gestures. *IEEE Transactions on Multimedia*, 21(1), Article 1. <https://doi.org/10.1109/TMM.2018.2856094>
- Avola, D., Bernardi, M., Cinque, L., Foresti, G. L., & Massaroni, C. (2019b). Exploiting Recurrent Neural Networks and Leap Motion Controller for the Recognition of Sign Language and Semaphore Hand Gestures. *IEEE Transactions on Multimedia*, 21(1), Article 1. <https://doi.org/10.1109/TMM.2018.2856094>
- Azure Kinect DK: Desarrollo de modelos de IA | Microsoft Azure*. (s. f.). Recuperado 7 de octubre de 2022, de <https://azure.microsoft.com/es-es/products/kinect-dk/>

## **Bibliografía**

---

- Barreto, A., Scargle, S., & Adjouadi, M. (1999). Hands-off human-computer interfaces for individuals with severe motor disabilities. *on Human-Computer Interaction: Communication, Cooperation, and Application Design-Volume 2 - Volume 2, 2*, 970-974.
- Bassily, D., Georgoulas, C., Guettler, J., Linner, T., & Bock, T. (2014). *Intuitive and adaptive robotic arm manipulation using the leap motion controller*. 1-7.
- Benalcázar, M. E., Jaramillo, A. G., Jonathan, Zea, A., Páez, A., & Andaluz, V. H. (2017). Hand gesture recognition using machine learning and the Myo armband. *2017 25th European Signal Processing Conference (EUSIPCO)*, 1040-1044. <https://doi.org/10.23919/EUSIPCO.2017.8081366>
- Benalcázar, M. E., Motoche, C., Zea, J. A., Jaramillo, A. G., Anchundia, C. E., Zambrano, P., Segura, M., Benalcázar Palacios, F., & Pérez, M. (2017). Real-time hand gesture recognition using the Myo armband and muscle activity detection. *2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM)*, 1-6. <https://doi.org/10.1109/ETCM.2017.8247458>
- Benedek, J., Hazlett, R. L., & Benedek, J. (2005). Incorporating facial emg emotion measures as feedback in the software design process. *In Proc. Human Computer Interaction Consortium*, 9.
- Biber, C., Ellin, S., Shenk, E., & Stempeck, J. (1980, octubre 1). *The Polaroid Ultrasonic Ranging System*. Audio Engineering Society Convention 67, Nueva York, NY, USA. <https://www.aes.org/e-lib/browse.cfm?elib=3680>
- Booij, W. E., & Welle, K. O. (2014). *Ultrasound detectors* (United States Patent US8792305B2). <https://patents.google.com/patent/US8792305B2/en>
- Bora, J., Dehingia, S., Boruah, A., Chetia, A. A., & Gogoi, D. (2023). Real-time Assamese Sign Language Recognition using MediaPipe and Deep Learning. *Procedia Computer Science*, 218, 1384-1393. <https://doi.org/10.1016/j.procs.2023.01.117>
- Borriello, G., Brunette, W., Hall, M., Hartung, C., & Tangney, C. (2004a). Reminding About Tagged Objects Using Passive RFIDs. En N. Davies, E. D. Mynatt, & I. Siio (Eds.), *UbiComp 2004: Ubiquitous Computing* (pp. 36-53). Springer. [https://doi.org/10.1007/978-3-540-30119-6\\_3](https://doi.org/10.1007/978-3-540-30119-6_3)

- Borriello, G., Brunette, W., Hall, M., Hartung, C., & Tangney, C. (2004b). Reminding About Tagged Objects Using Passive RFIDs. En N. Davies, E. D. Mynatt, & I. Siio (Eds.), *UbiComp 2004: Ubiquitous Computing* (pp. 36-53). Springer. [https://doi.org/10.1007/978-3-540-30119-6\\_3](https://doi.org/10.1007/978-3-540-30119-6_3)
- Bouchard, K., Bouzouane, A., & Bouchard, B. (2014). Gesture recognition in smart home using passive RFID technology. *Proceedings of the 7th International Conference on Pervasive Technologies Related to Assistive Environments*, 1-8. <https://doi.org/10.1145/2674396.2674405>
- Boyalı, A., Hashimoto, N., & Matsumoto, O. (2015). Hand posture and gesture recognition using MYO armband and spectral collaborative representation based classification. *2015 IEEE 4th Global Conference on Consumer Electronics (GCCE)*, 200-201. <https://doi.org/10.1109/GCCE.2015.7398619>
- Brock, H., & Gomez, R. (2021). Personalization of Human-Robot Gestural Communication through Voice Interaction Grounding. *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 846-853. <https://doi.org/10.1109/IROS51168.2021.9636105>
- Brock, H., & Gomez, R. (2022). Making The Unknown More Certain: A Stacked Ensemble Classifier for Open Gesture Recognition with a Social Robot. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3918-3922. <https://doi.org/10.1109/ICASSP43922.2022.9747370>
- Brown, E., Large, D. R., Limerick, H., Frier, W., & Burnett, G. (2022). Augmenting Automotive Gesture Infotainment Interfaces Through Mid-Air Haptic Icon Design. En O. Georgiou, W. Frier, E. Freeman, C. Pacchierotti, & T. Hoshi (Eds.), *Ultrasound Mid-Air Haptics for Touchless Interfaces* (pp. 119-145). Springer International Publishing. [https://doi.org/10.1007/978-3-031-04043-6\\_5](https://doi.org/10.1007/978-3-031-04043-6_5)
- Brown, E., R. Large, D., Limerick, H., & Burnett, G. (2020). Ultrahapticons: "Haptifying" Drivers' Mental Models to Transform Automotive Mid-Air Haptic Gesture Infotainment Interfaces. *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 54-57. <https://doi.org/10.1145/3409251.3411722>
- Buettner, M., Prasad, R., Philipose, M., & Wetherall, D. (2009). Recognizing daily activities with RFID-based sensors. *Proceedings of the 11th international conference on Ubiquitous computing*, 51-60. <https://doi.org/10.1145/1620545.1620553>

- Calin, A. D. (2016). Gesture Recognition on Kinect Time Series Data Using Dynamic Time Warping and Hidden Markov Models. *2016 18th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, 264-271. <https://doi.org/10.1109/SYNASC.2016.049>
- Chai, X., Li, G., Lin, Y., Xu, Z., Tang, Y., Chen, X., & Zhou, M. (2013). *Sign language recognition and translation with kinect*. 655, 4.
- Chan, A., Halevi, T., & Memon, N. (2015). Leap Motion Controller for Authentication via Hand Geometry and Gestures. En T. Tryfonas & I. Askoxylakis (Eds.), *Human Aspects of Information Security, Privacy, and Trust* (pp. 13-22). Springer International Publishing. [https://doi.org/10.1007/978-3-319-20376-8\\_2](https://doi.org/10.1007/978-3-319-20376-8_2)
- Chen, X., Zheng, Y.-P., Guo, J.-Y., & Shi, J. (2010a). Sonomyography (SMG) Control for Powered Prosthetic Hand: A Study with Normal Subjects. *Ultrasound in Medicine & Biology*, 36(7), 1076-1088. <https://doi.org/10.1016/j.ultrasmedbio.2010.04.015>
- Chen, X., Zheng, Y.-P., Guo, J.-Y., & Shi, J. (2010b). Sonomyography (SMG) Control for Powered Prosthetic Hand: A Study with Normal Subjects. *Ultrasound in Medicine & Biology*, 36(7), Article 7. <https://doi.org/10.1016/j.ultrasmedbio.2010.04.015>
- Chen, Y., Ding, Z., Chen, Y.-L., & Wu, X. (2015a). Rapid recognition of dynamic hand gestures using leap motion. *2015 IEEE International Conference on Information and Automation*, 1419-1424. <https://doi.org/10.1109/ICInfA.2015.7279509>
- Chen, Y., Ding, Z., Chen, Y.-L., & Wu, X. (2015b). Rapid recognition of dynamic hand gestures using leap motion. *2015 IEEE International Conference on Information and Automation*, 1419-1424. <https://doi.org/10.1109/ICInfA.2015.7279509>
- Chong, T.-W., & Lee, B. G. (2018a). American Sign Language Recognition Using Leap Motion Controller with Machine Learning Approach. *Sensors*, 18, 3554. <https://doi.org/10.3390/s18103554>
- Chong, T.-W., & Lee, B. G. (2018b). American Sign Language Recognition Using Leap Motion Controller with Machine Learning Approach. *Sensors*, 18, 3554. <https://doi.org/10.3390/s18103554>
- Chong, T.-W., & Lee, B. G. (2018c). American Sign Language Recognition Using Leap Motion Controller with Machine Learning Approach. *Sensors*, 18, 3554. <https://doi.org/10.3390/s18103554>

- Chung, W.-Y., Xu, H., & Lee, B. G. (2022). Chinese Sign Language Recognition with Batch Sampling ResNet-Bi-LSTM. *SN Computer Science*, 3(5), Article 5. <https://doi.org/10.1007/s42979-022-01341-4>
- Cichon, T., & Roßmann, J. (2017). *Simulation-based user interfaces for digital twins: Pre-, in-, or post-operational analysis and exploration of virtual testbeds*. 365-372.
- Cohen, I., & Li, H. (2003). Inference of human postures by classification of 3D human body shape. *2003 IEEE International SOI Conference. Proceedings (Cat. No.03CH37443)*, 74-81. <https://doi.org/10.1109/AMFG.2003.1240827>
- Cohen, M. W., Zikri, N. B., & Velkovich, A. (2018). Recognition of Continuous Sign Language Alphabet Using Leap Motion Controller. *2018 11th International Conference on Human System Interaction (HSI)*, 193-199. <https://doi.org/10.1109/HSI.2018.8430860>
- Coleman, K. (2001). Electromyography based human-computer-interface to induce movement in elderly persons with movement impairments. *Proceedings of the 2001 EC/NSF workshop on Universal accessibility of ubiquitous computing: providing for the elderly*, 75-79. <https://doi.org/10.1145/564526.564547>
- Corradini, A., Boehme, H.-J., & Gross, H.-M. (1999). *Visual-based posture recognition using hybrid neural networks*. 81-86.
- Crawford, B., Miller, K., Shenoy, P., & Rao, R. (2005). *Real-Time Classification of Electromyographic Signals for Robotic Control*. 2, 523-528.
- CyberGlove Systems LLC. (s. f.). Recuperado 25 de agosto de 2022, de <http://www.cyberglovesystems.com/>
- Dabwan, B. A., Jadhav, M. E., Ali, Y. A., & Olayah, F. A. (2023). Arabic Sign Language Recognition Using EfficientnetB1 and Transfer Learning Technique. *2023 International Conference on IT Innovation and Knowledge Discovery (ITIKD)*, 1-5. <https://doi.org/10.1109/ITIKD56332.2023.10099710>
- Das, S., Biswas, S. K., & Purkayastha, B. (2023). Automated Indian sign language recognition system by fusing deep and handcrafted feature. *Multimedia Tools and Applications*, 82(11), Article 11. <https://doi.org/10.1007/s11042-022-14084-4>

## **Bibliografía**

---

- Das, S., Imtiaz, Md. S., Neom, N. H., Siddique, N., & Wang, H. (2023). A hybrid approach for Bangla sign language recognition using deep transfer learning model with random forest classifier. *Expert Systems with Applications*, 213, 118914. <https://doi.org/10.1016/j.eswa.2022.118914>
- Deng, J., Sun, B., Kavcic, V., Liu, M., Giordani, B., & Li, T. (2023). Novel methodology for detection and prediction of mild cognitive impairment using resting-state EEG. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*. <https://doi.org/10.1002/alz.13411>
- Deogharia, R., & Sil, S. (2023). Reconstructing High-Frequency Radar Derived Ocean Surface-Current Fields Using Spatio-Temporal Kriging. *IEEE Journal of Oceanic Engineering*, 1-11. <https://doi.org/10.1109/JOE.2023.3283593>
- Dhillon, H. S., Singla, R., Rekhi, N. S., & Jha, R. (2009). EOG and EMG based virtual keyboard: A brain-computer interface. *2009 2nd IEEE International Conference on Computer Science and Information Technology*, 259-262. <https://doi.org/10.1109/ICCSIT.2009.5234951>
- Dubost, G., & Tanaka, A. (2002, septiembre). *A Wireless, Network-based Biosensor Interface for Music*. ICMC, Goteborg, Suecia. [https://www.researchgate.net/publication/255611768\\_A\\_Wireless\\_Network-based\\_Biosensor\\_Interface\\_for\\_Music](https://www.researchgate.net/publication/255611768_A_Wireless_Network-based_Biosensor_Interface_for_Music)
- Elons, A. S., Ahmed, M., Shedid, H., & Tolba, M. F. (2014a). Arabic sign language recognition using leap motion sensor. *2014 9th International Conference on Computer Engineering & Systems (ICCES)*, 368-373. <https://doi.org/10.1109/ICCES.2014.7030987>
- Elons, A. S., Ahmed, M., Shedid, H., & Tolba, M. F. (2014b). Arabic sign language recognition using leap motion sensor. *2014 9th International Conference on Computer Engineering & Systems (ICCES)*, 368-373. <https://doi.org/10.1109/ICCES.2014.7030987>
- Farina, D., Stegeman, D. F., & Merletti, R. (2016a). Biophysics of the Generation of EMG Signals. En *Surface Electromyography: Physiology, Engineering, and Applications* (pp. 1-24). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119082934.ch02>
- Farina, D., Stegeman, D. F., & Merletti, R. (2016b). Biophysics of the Generation of EMG Signals. En *Surface Electromyography: Physiology, Engineering, and Applications* (pp. 1-24). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119082934.ch02>

- Fishkin, K. P., Jiang, B., Philipose, M., & Roy, S. (2004). I Sense a Disturbance in the Force: Unobtrusive Detection of Interactions with RFID-tagged Objects. En N. Davies, E. D. Mynatt, & I. Siiio (Eds.), *UbiComp 2004: Ubiquitous Computing* (pp. 268-282). Springer. [https://doi.org/10.1007/978-3-540-30119-6\\_16](https://doi.org/10.1007/978-3-540-30119-6_16)
- Fistre, J., & Tanaka, A. (2002). Real time EMG gesture recognition for consumer electronics device control. *Sony CSL Paris Open House 10/2002*.
- Fléty, E. (2000). *3D gesture acquisition using ultrasonic sensors*. 193-207.
- Fuchs, C., Kollatsch, C., Gärtner, L., Heinz, A., Kiesling, A. von, Weinhold, S., & Klimant, F. (2022). Investigation on Tactile Perception of Ultrasonic Haptics Devices. *2022 IEEE 9th International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, 1-6. <https://doi.org/10.1109/CIVEMSA53371.2022.9853702>
- Funasaka, M., Ishikawa, Y., Takata, M., & Joe, K. (2015). *Sign Language Recognition using Leap Motion Controller*. 1, 263-269. <http://worldcomp-proceedings.com/proc/p2015/PDP7080.pdf>
- Gándara, C. V., & Bauza, C. G. (2015). IntelliHome: A framework for the development of ambient assisted living applications based in low-cost technology. *Proceedings of the Latin American Conference on Human Computer Interaction*, 1-4. <https://doi.org/10.1145/2824893.2824913>
- Gavrila, D. M., & Davis, L. S. (1995a). *Towards 3-d model-based tracking and recognition of human movement: A multi-view approach*. 3, 272-277.
- Gavrila, D. M., & Davis, L. S. (1995b). *Towards 3-d model-based tracking and recognition of human movement: A multi-view approach*. 3, 272-277.
- Geng, L., Ma, X., Xue, B., Wu, H., Gu, J., & Li, Y. (2014). Combining features for Chinese sign language recognition with Kinect. *11th IEEE International Conference on Control & Automation (ICCA)*, 1393-1398. <https://doi.org/10.1109/ICCA.2014.6871127>
- Georgiou, O., Frier, W., Freeman, E., Pacchierotti, C., & Hoshi, T. (2022). Mid-Air Haptics: Future Challenges and Opportunities. En O. Georgiou, W. Frier, E. Freeman, C. Pacchierotti, & T. Hoshi (Eds.), *Ultrasound Mid-Air Haptics for Touchless Interfaces* (pp. 385-397). Springer International Publishing. [https://doi.org/10.1007/978-3-031-04043-6\\_18](https://doi.org/10.1007/978-3-031-04043-6_18)

## **Bibliografía**

---

- Georgiou, O., Martinez, J., Abdouni, A., & Harwood, A. (2022). Mid-air Haptic Texture Exploration in VR. 2022 *IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 964-965. <https://doi.org/10.1109/VRW55335.2022.00333>
- Gowitzke, B. A., & Milner, M. (1999a). *El cuerpo y sus movimientos: Bases científicas*. Paidotribo.
- Gowitzke, B. A., & Milner, M. (1999b). *El cuerpo y sus movimientos: Bases científicas*. Paidotribo.
- Guerreiro, T., & Jorge, J. (2006). *EMG as a daily wearable interface*. 216-223.
- Guna, J., Jakus, G., Pogačnik, M., Tomažič, S., & Sodnik, J. (2014). An Analysis of the Precision and Reliability of the Leap Motion Sensor and Its Suitability for Static and Dynamic Tracking. *Sensors*, *14*(2), Article 2. <https://doi.org/10.3390/s140203702>
- Guo, J. Y., Zheng, Y., Huang, Q. H., Chen, X., & He Chen, J. F. (2008). Comparison of sonomyography and electromyography of forearm muscles in the guided wrist extension: 5th International Workshop on Wearable and Implantable Body Sensor Networks, BSN 2008, in conjunction with the 5th International Summer School and Symposium on Medical Devices and Biosensors, ISSS-MDBS 2008. *Proc. 5th Int. Workshop on Wearable and Implantable Body Sensor Networks, BSN 2008, in conjunction with the 5th Int. Summer School and Symp. on Medical Devices and Biosensors, ISSS-MDBS 2008*, 235-238. <https://doi.org/10.1109/ISSMDBS.2008.4575062>
- Gupta, S., Morris, D., Patel, S., & Tan, D. (2012a). SoundWave: Using the doppler effect to sense gestures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1911-1914. <https://doi.org/10.1145/2207676.2208331>
- Gupta, S., Morris, D., Patel, S., & Tan, D. (2012b). SoundWave: Using the doppler effect to sense gestures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1911-1914. <https://doi.org/10.1145/2207676.2208331>
- Hantrakul, L., & Kaczmarek, K. (2014). Implementations of the Leap Motion device in sound synthesis and interactive live performance. *Proceedings of the 2014 International Workshop on Movement and Computing*, 142-145. <https://doi.org/10.1145/2617995.2618020>

- 
- He, W., Wu, K., Zou, Y., & Ming, Z. (2015). WiG: WiFi-Based Gesture Recognition System. *2015 24th International Conference on Computer Communication and Networks (ICCCN)*, 1-7. <https://doi.org/10.1109/ICCCN.2015.7288485>
- Healey, J., & Picard, R. W. (1998). StartleCam: A cybernetic wearable camera. *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No.98EX215)*, 42-49. <https://doi.org/10.1109/ISWC.1998.729528>
- Hefftner, G., Zucchini, W., & Jaros, G. G. (1988). The electromyogram (EMG) as a control signal for functional neuromuscular stimulation. I. Autoregressive modeling as a means of EMG signature discrimination. *IEEE Transactions on Biomedical Engineering*, 35(4), 230-237. <https://doi.org/10.1109/10.1370>
- Hernández, S., & López-Córtés, X. (2023). Evaluating deep learning predictions for COVID-19 from X-ray images using leave-one-out predictive densities. *Neural Computing and Applications*, 35(13), 9819-9830. <https://doi.org/10.1007/s00521-023-08219-3>
- Hernandez-Rebollar, J. L., Kyriakopoulos, N., & Lindeman, R. W. (2002). The AcceleGlove: A whole-hand input device for virtual reality. *ACM SIGGRAPH 2002 conference abstracts and applications*, 259. <https://doi.org/10.1145/1242073.1242272>
- Hettiarachchi, N., Ju, Z., & Liu, H. (2015). A New Wearable Ultrasound Muscle Activity Sensing System for Dexterous Prosthetic Control. *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 1415-1420. <https://doi.org/10.1109/SMC.2015.251>
- Hisham, B., & Hamouda, A. (2018a). *Arabic Sign Language Recognition using Microsoft Kinect and Leap Motion Controller* (SSRN Scholarly Paper 3389799; Número 3389799). <https://doi.org/10.2139/ssrn.3389799>
- Hisham, B., & Hamouda, A. (2018b). *Arabic Sign Language Recognition using Microsoft Kinect and Leap Motion Controller* (SSRN Scholarly Paper 3389799; Número 3389799). <https://doi.org/10.2139/ssrn.3389799>
- Hodges, P. W., Pengel, L. H. M., Herbert, R. D., & Gandevia, S. C. (2003). Measurement of muscle contraction with ultrasound imaging. *Muscle & Nerve*, 27(6), 682-692. <https://doi.org/10.1002/mus.10375>

## **Bibliografía**

---

- Hossain, D., Ghosh, T., Haider Imtiaz, M., & Sazonov, E. (2023). Ear canal pressure sensor for food intake detection. *Frontiers in Electronics*, 4. <https://www.frontiersin.org/articles/10.3389/felec.2023.1173607>
- HP Reverb G2 VR Headset. (s. f.). Recuperado 29 de agosto de 2022, de <https://www.hp.com/us-en/vr/reverb-g2-vr-headset.html>
- Interacción 3D con Leap Motion. (2014, octubre 1). inLab FIB. <https://inlab.fib.upc.edu/es/blog/interaccion-3d-con-leap-motion>
- Iosa, M., Morone, G., Fusco, A., Castagnoli, M., Fusco, F. R., Pratesi, L., & Paolucci, S. (2015). Leap motion controlled videogame-based therapy for rehabilitation of elderly patients with subacute stroke: A feasibility pilot study. *Topics in Stroke Rehabilitation*, 22, 306-316. <https://doi.org/10.1179/1074935714Z.00000000036>
- Jambhale, S. S., & Khaparde, A. (2014). Gesture recognition using DTW & piecewise DTW. *2014 International Conference on Electronics and Communication Systems (ICECS)*, 1-5. <https://doi.org/10.1109/ECS.2014.6892646>
- Jamil, R., Dong, M., Jamil, A., & Bano, S. (2023). Classification of images using Extreme Learning Machine and Regularization Extreme learning Machine (ELM) with the leave-one-out cross validation. *2023 IEEE International Conference on Emerging Trends in Engineering, Sciences and Technology (ICES&T)*, 1-6. <https://doi.org/10.1109/ICEST56843.2023.10138854>
- Jayatilaka, A., & Ranasinghe, D. C. (2017). Real-time fluid intake gesture recognition based on batteryless UHF RFID technology. *Pervasive and Mobile Computing*, 34, 146-156. <https://doi.org/10.1016/j.pmcj.2016.04.007>
- Jenkins, J. (2021). An Innovative Method for Automatic American Sign Language Interpretation using Machine Learning and Leap Motion Controller. *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 0633-0638. <https://doi.org/10.1109/UEMCON53757.2021.9666640>
- Jetter, H.-C., Rädle, R., Feuchtner, T., Anthes, C., Friedl, J., & Klokose, C. N. (2020a). «In VR, everything is possible!»: Sketching and Simulating Spatially-Aware Interactive Spaces in Virtual Reality.

- 
- Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-16.  
<https://doi.org/10.1145/3313831.3376652>
- Jetter, H.-C., Rädle, R., Feuchtner, T., Anthes, C., Friedl, J., & Klokmose, C. N. (2020b). «In VR, everything is possible!»: Sketching and Simulating Spatially-Aware Interactive Spaces in Virtual Reality. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-16.  
<https://doi.org/10.1145/3313831.3376652>
- Joshi, C. D., Lahiri, U., & Thakor, N. V. (2013). Classification of gait phases from lower limb EMG: Application to exoskeleton orthosis. *IEEE EMBS Special Topic Conference on Point-of-Care (POC) Healthcare Technologies: Synergy Towards Better Global Healthcare, PHT 2013*, 228-231.  
<https://doi.org/10.1109/PHT.2013.6461326>
- K, P., & B.J, S. (2023). Hand Landmark Distance Based Sign Language Recognition using MediaPipe. *2023 International Conference on Emerging Smart Computing and Informatics (ESCI)*, 1-7.  
<https://doi.org/10.1109/ESCI56872.2023.10100061>
- Kalgaonkar, K., & Raj, B. (2009). One-handed gesture recognition using ultrasonic Doppler sonar. *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 1889-1892.  
<https://doi.org/10.1109/ICASSP.2009.4959977>
- Kaliappan, J., Bagepalli, A. R., Almal, S., Mishra, R., Hu, Y.-C., & Srinivasan, K. (2023). Impact of Cross-Validation on Machine Learning Models for Early Detection of Intrauterine Fetal Demise. *Diagnostics*, 13(10), Article 10. <https://doi.org/10.3390/diagnostics13101692>
- Katılmış, Z., & Karakuzu, C. (2023). Double handed dynamic Turkish Sign Language recognition using Leap Motion with meta learning approach. *Expert Systems with Applications*, 228, 120453.  
<https://doi.org/10.1016/j.eswa.2023.120453>
- Kaushik, Dr. M., & Jain, R. (2014). Gesture Based Interaction NUI: An Overview. *International Journal of Engineering Trends and Technology*, 9(12), 633-636. <https://doi.org/10.14445/22315381/IJETT-V9P319>
- Kellogg, B., Talla, V., & Gollakota, S. (2014). *Bringing Gesture Recognition to All Devices*. 303-316.  
<https://www.usenix.org/conference/nsdi14/technical-sessions/presentation/kellogg>
-

## **Bibliografía**

---

- Kiguchi, K., & Hayashi, Y. (2012). An EMG-Based Control for an Upper-Limb Power-Assist Exoskeleton Robot. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics: A Publication of the IEEE Systems, Man, and Cybernetics Society*, 42(4), 1064-1071. <https://doi.org/10.1109/TSMCB.2012.2185843>
- Kinect para Windows—Windows apps.* (s. f.). Recuperado 26 de agosto de 2022, de <https://docs.microsoft.com/es-es/windows/apps/design/devices/kinect-for-windows>
- Knapp, R. B., & Lusted, H. S. (1990). A Bioelectric Controller for Computer Music Applications. *Computer Music Journal*, 14(1), 42-47. <https://doi.org/10.2307/3680115>
- Kothadiya, D., Bhatt, C., Sapariya, K., Patel, K., Gil-González, A.-B., & Corchado, J. M. (2022). Deepsign: Sign Language Detection and Recognition Using Deep Learning. *Electronics*, 11(11), Article 11. <https://doi.org/10.3390/electronics11111780>
- Kovacs-Vajna, Z. M. (2000). A fingerprint verification system based on triangular matching and dynamic time warping. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), 1266-1276. <https://doi.org/10.1109/34.888711>
- Kreczmer, B. (2011). Gestures recognition by using ultrasonic range-finders. *2011 16th International Conference on Methods & Models in Automation & Robotics*, 363-368. <https://doi.org/10.1109/MMAR.2011.6031374>
- Krueger, M. W., Gionfriddo, T., & Hinrichsen, K. (1985). VIDEOPLACE—an artificial reality. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 35-40. <https://doi.org/10.1145/317456.317463>
- Kuzmanic, A., & Zanchi, V. (2007). Hand shape classification using DTW and LCSS as similarity measures for vision-based gesture recognition system. *EUROCON 2007 - The International Conference on «Computer as a Tool»*, 264-269. <https://doi.org/10.1109/EURCON.2007.4400350>
- Lang, S., Block, M., & Rojas, R. (2012). Sign Language Recognition Using Kinect. En L. Rutkowski, M. Korytkowski, R. Scherer, R. Tadeusiewicz, L. A. Zadeh, & J. M. Zurada (Eds.), *Artificial Intelligence and Soft Computing* (pp. 394-402). Springer. [https://doi.org/10.1007/978-3-642-29347-4\\_46](https://doi.org/10.1007/978-3-642-29347-4_46)

- LaViola, J. J. J. (1999a). *A Survey of Hand Posture and Gesture Recognition Techniques and Technology* (Technical Report no. CS-99-11; Número Technical Report no. CS-99-11). Brown Univ. Provid. RI. <https://www.semanticscholar.org/paper/A-Survey-of-Hand-Posture-and-Gesture-Recognition-LaViola/856d4bf0f1f5d4480ce3115d828f34d4b2782e1c>
- LaViola, J. J. J. (1999b). *A Survey of Hand Posture and Gesture Recognition Techniques and Technology* (Technical Report no. CS-99-11). Brown Univ. Provid. RI. <https://www.semanticscholar.org/paper/A-Survey-of-Hand-Posture-and-Gesture-Recognition-LaViola/856d4bf0f1f5d4480ce3115d828f34d4b2782e1c>
- Lee, C. K. M., Ng, K. K. H., Chen, C.-H., Lau, H. C. W., Chung, S. Y., & Tsoi, T. (2021a). American sign language recognition and training method with recurrent neural network. *Expert Systems with Applications*, *167*, 114403. <https://doi.org/10.1016/j.eswa.2020.114403>
- Lee, C. K. M., Ng, K. K. H., Chen, C.-H., Lau, H. C. W., Chung, S. Y., & Tsoi, T. (2021b). American sign language recognition and training method with recurrent neural network. *Expert Systems with Applications*, *167*, 114403. <https://doi.org/10.1016/j.eswa.2020.114403>
- Lee, G. C., Yeh, F.-H., & Hsiao, Y.-H. (2016). Kinect-based Taiwanese sign-language recognition system. *Multimedia Tools and Applications*, *75*(1), 261-279. <https://doi.org/10.1007/s11042-014-2290-x>
- Legrand, B., Chang, C. S., Ong, S. H., Neo, S.-Y., & Palanisamy, N. (2008). Chromosome classification using dynamic time warping. *Pattern Recognition Letters*, *29*(3), 215-222. <https://doi.org/10.1016/j.patrec.2007.09.017>
- Leth, P. G. (2023). Danish Sign Language Recognition in Virtual Reality Using Written Language Ensemble Learning. *Universidad de Aalborg*. [https://projekter.aau.dk/projekter/files/536477376/20184813\\_masters\\_thesis.pdf](https://projekter.aau.dk/projekter/files/536477376/20184813_masters_thesis.pdf)
- Li, D., Opazo, C. R., Yu, X., & Li, H. (2020). Word-level Deep Sign Language Recognition from Video: A New Large-scale Dataset and Methods Comparison. *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 1448-1458. <https://doi.org/10.1109/WACV45572.2020.9093512>

## **Bibliografía**

---

- Li, H., Yang, W., Wang, J., Xu, Y., & Huang, L. (2016a). WiFinger: Talk to your smart devices with finger-grained gesture. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 250-261. <https://doi.org/10.1145/2971648.2971738>
- Li, H., Yang, W., Wang, J., Xu, Y., & Huang, L. (2016b). WiFinger: Talk to your smart devices with finger-grained gesture. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 250-261. <https://doi.org/10.1145/2971648.2971738>
- Li, W.-J., Hsieh, C.-Y., Lin, L.-F., & Chu, W.-C. (2017). Hand gesture recognition for post-stroke rehabilitation using leap motion. *2017 International Conference on Applied System Innovation (ICASI)*, 386-388. <https://doi.org/10.1109/ICASI.2017.7988433>
- Li, Y. (2012). Hand gesture recognition using Kinect. *2012 IEEE International Conference on Computer Science and Automation Engineering*, 196-199. <https://doi.org/10.1109/ICSESS.2012.6269439>
- Ling, H., & Rui, L. (2016). VR glasses and leap motion trends in education. *2016 11th International Conference on Computer Science & Education (ICCSE)*, 917-920. <https://doi.org/10.1109/ICCSE.2016.7581705>
- Liu, T., Zhou, W., & Li, H. (2016). Sign language recognition with long short-term memory. *2016 IEEE International Conference on Image Processing (ICIP)*, 2871-2875. <https://doi.org/10.1109/ICIP.2016.7532884>
- Lu, W., Tong, Z., & Chu, J. (2016a). Dynamic Hand Gesture Recognition With Leap Motion Controller. *IEEE Signal Processing Letters*, 23(9), Article 9. <https://doi.org/10.1109/LSP.2016.2590470>
- Lu, W., Tong, Z., & Chu, J. (2016b). Dynamic Hand Gesture Recognition With Leap Motion Controller. *IEEE Signal Processing Letters*, 23(9), 1188-1192. <https://doi.org/10.1109/LSP.2016.2590470>
- Luca, C. J. D. (1997). The Use of Surface Electromyography in Biomechanics. *Journal of Applied Biomechanics*, 13(2), 135-163. <https://doi.org/10.1123/jab.13.2.135>
- Manabe, H., Hiraiwa, A., & Sugimura, T. (2003). Unvoiced speech recognition using EMG - mime speech recognition. *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, 794-795. <https://doi.org/10.1145/765891.765996>
- Manteca Fernández, F. (2018a). Generación de modelos 3D mediante luz estructurada. *Universidad de Cantabria*. <http://hdl.handle.net/10902/15136>

- Manteca Fernández, F. (2018b). Generación de modelos 3D mediante luz estructurada. *Universidad de Cantabria*. <http://hdl.handle.net/10902/15136>
- Mapari, R. B., & Kharat, G. (2016a). American Static Signs Recognition Using Leap Motion Sensor. *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, 1-5. <https://doi.org/10.1145/2905055.2905125>
- Mapari, R. B., & Kharat, G. (2016b). American Static Signs Recognition Using Leap Motion Sensor. *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, 1-5. <https://doi.org/10.1145/2905055.2905125>
- Mapari, R. B., & Kharat, G. (2016c). American Static Signs Recognition Using Leap Motion Sensor. *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, 1-5. <https://doi.org/10.1145/2905055.2905125>
- Marin, G., Dominio, F., & Zanuttigh, P. (2014a). Hand gesture recognition with leap motion and kinect devices. *2014 IEEE International Conference on Image Processing (ICIP)*, 1565-1569. <https://doi.org/10.1109/ICIP.2014.7025313>
- Marin, G., Dominio, F., & Zanuttigh, P. (2014b). Hand gesture recognition with leap motion and kinect devices. *2014 IEEE International Conference on Image Processing (ICIP)*, 1565-1569. <https://doi.org/10.1109/ICIP.2014.7025313>
- Marin, G., Dominio, F., & Zanuttigh, P. (2014c). Hand gesture recognition with leap motion and kinect devices. *2014 IEEE International Conference on Image Processing (ICIP)*, 1565-1569. <https://doi.org/10.1109/ICIP.2014.7025313>
- Marin, G., Dominio, F., & Zanuttigh, P. (2014d). Hand gesture recognition with leap motion and kinect devices. *2014 IEEE International Conference on Image Processing (ICIP)*, 1565-1569. <https://doi.org/10.1109/ICIP.2014.7025313>
- McCartney, R., Yuan, J., & Bischof, H.-P. (2015, julio). *Gesture recognition with the leap motion controller*. 19th International Conference on Image Processing, Computer Vision, & Pattern Recognition, Las Vegas, NV, USA.

## **Bibliografía**

---

- McIntosh, J., Marzo, A., Fraser, M., & Phillips, C. (2017a). EchoFlex: Hand Gesture Recognition using Ultrasound Imaging. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1923-1934. <https://doi.org/10.1145/3025453.3025807>
- McIntosh, J., Marzo, A., Fraser, M., & Phillips, C. (2017b). EchoFlex: Hand Gesture Recognition using Ultrasound Imaging. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1923-1934. <https://doi.org/10.1145/3025453.3025807>
- Mehrabian, A. (2017). *Nonverbal Communication*. Routledge.
- Mehrotra, K., Godbole, A., & Belhe, S. (2015). Indian Sign Language Recognition Using Kinect Sensor. En M. Kamel & A. Campilho (Eds.), *Image Analysis and Recognition* (pp. 528-535). Springer International Publishing. [https://doi.org/10.1007/978-3-319-20801-5\\_59](https://doi.org/10.1007/978-3-319-20801-5_59)
- Mejora tu experiencia.* (s. f.). Valve Corporation. Recuperado 29 de agosto de 2022, de <https://www.valvesoftware.com/es/index>
- Merletti, R., & Farina, D. (Eds.). (2016). *Surface Electromyography: Physiology, Engineering, and Applications* (Piscataway, New Jersey : IEEE Press). Wiley-IEEE Press.
- Merletti, R., Rainoldi, A., & Farina, D. (2004a). Myoelectric Manifestations of Muscle Fatigue. En *Electromyography* (pp. 233-258). John Wiley & Sons, Ltd. <https://doi.org/10.1002/0471678384.ch9>
- Merletti, R., Rainoldi, A., & Farina, D. (2004b). Myoelectric Manifestations of Muscle Fatigue. En *Electromyography* (pp. 233-258). John Wiley & Sons, Ltd. <https://doi.org/10.1002/0471678384.ch9>
- Mo, H.-C., Leou, J.-J., & Lin, C.-S. (2009a). *Human Behavior Analysis Using Multiple 2D Features and Multicategory Support Vector Machine*. 46-49.
- Mo, H.-C., Leou, J.-J., & Lin, C.-S. (2009b). *Human Behavior Analysis Using Multiple 2D Features and Multicategory Support Vector Machine*. 46-49.
- Mohandes, M., Aliyu, S., & Deriche, M. (2014a). Arabic sign language recognition using the leap motion controller. *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)*, 960-965. <https://doi.org/10.1109/ISIE.2014.6864742>

- 
- Mohandes, M., Aliyu, S., & Deriche, M. (2014b). Arabic sign language recognition using the leap motion controller. *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)*, 960-965. <https://doi.org/10.1109/ISIE.2014.6864742>
- Mohandes, M., Aliyu, S., & Deriche, M. (2014c). Arabic sign language recognition using the leap motion controller. *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)*, 960-965. <https://doi.org/10.1109/ISIE.2014.6864742>
- Mohandes, M., Deriche, M., & Liu, J. (2014a). Image-Based and Sensor-Based Approaches to Arabic Sign Language Recognition. *IEEE Transactions on Human-Machine Systems*, 44(4), Article 4. <https://doi.org/10.1109/THMS.2014.2318280>
- Mohandes, M., Deriche, M., & Liu, J. (2014b). Image-Based and Sensor-Based Approaches to Arabic Sign Language Recognition. *IEEE Transactions on Human-Machine Systems*, 44(4), Article 4. <https://doi.org/10.1109/THMS.2014.2318280>
- Montalvo Martínez, M. (2010a). *Técnicas de visión estereoscópica para determinar la estructura tridimensional de la escena* [Info:eu-repo/semantics/masterThesis]. Universidad Complutense de Madrid. <https://eprints.ucm.es/id/eprint/11350/>
- Montalvo Martínez, M. (2010b). *Técnicas de visión estereoscópica para determinar la estructura tridimensional de la escena* [Info:eu-repo/semantics/masterThesis]. Universidad Complutense de Madrid. <https://eprints.ucm.es/id/eprint/11350/>
- Moreno, R., Andres, C., & Pinzón Arenas, J. (2013a). Diseño de un guante háptico de control para manipulador robótico teleoperado. *Ingenium*, 7, 19. <https://doi.org/10.21774/ing.v7i17.313>
- Moreno, R., Andres, C., & Pinzón Arenas, J. (2013b). Diseño de un guante háptico de control para manipulador robótico teleoperado. *Ingenium*, 7, 19. <https://doi.org/10.21774/ing.v7i17.313>
- Muda, L., Begam, M., & Elamvazuthi, I. (2010). *Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques*. arXiv. <https://doi.org/10.48550/arXiv.1003.4083>
- Mujibiya, A., Cao, X., Tan, D. S., Morris, D., Patel, S. N., & Rekimoto, J. (2013). The sound of touch: On-body touch and gesture sensing based on transdermal ultrasound propagation. *Proceedings of the 2013*
-

## Bibliografía

---

- ACM international conference on Interactive tabletops and surfaces*, 189-198.  
<https://doi.org/10.1145/2512349.2512821>
- Mulot, L., Gicquel, G., Zanini, Q., Frier, W., Marchal, M., Pacchierotti, C., & Howard, T. (2021). DOLPHIN: A Framework for the Design and Perceptual Evaluation of Ultrasound Mid-Air Haptic Stimuli. *ACM Symposium on Applied Perception 2021*, 1-10. <https://doi.org/10.1145/3474451.3476232>
- MX25L3206E DATASHEET*. (s. f.). Recuperado 29 de agosto de 2002, de <https://www.mdpi.com/1424-8220/20/12/3571/htm>
- Myers, B. A. (1998). A brief history of human-computer interaction technology. *Interactions*, 5(2), 44-54.  
<https://doi.org/10.1145/274430.274436>
- Naglot, D., & Kulkarni, M. (2016a). Real time sign language recognition using the leap motion controller. *2016 International Conference on Inventive Computation Technologies (ICICT)*, 3, 1-5.  
<https://doi.org/10.1109/INVENTIVE.2016.7830097>
- Naglot, D., & Kulkarni, M. (2016b). Real time sign language recognition using the leap motion controller. *2016 International Conference on Inventive Computation Technologies (ICICT)*, 3, 1-5.  
<https://doi.org/10.1109/INVENTIVE.2016.7830097>
- Nandakumar, R., Kellogg, B., & Gollakota, S. (2014). *Wi-Fi Gesture Recognition on Existing Devices* (Human-Computer Interaction arXiv:1411.5394). arXiv. <https://doi.org/10.48550/arXiv.1411.5394>
- Nasir, M., Tengku, M., & Kurniawaty, E. (2022). *Hand Gesture Recognition using Leap Motion Controller for Recognition of Javanese Script*. 124-128.  
<https://abecindonesia.org/iabec/index.php/iabec/article/download/96/65>
- Noor, A. K., & Aras, R. (2015). Potential of multimodal and multiuser interaction with virtual holography. *Advances in Engineering Software*, 81, 1-6. <https://doi.org/10.1016/j.advengsoft.2014.10.004>
- Obi, Y., Claudio, K. S., Budiman, V. M., Achmad, S., & Kurniawan, A. (2023). Sign language recognition system for communicating to people with disabilities. *Procedia Computer Science*, 216, 13-20.  
<https://doi.org/10.1016/j.procs.2022.12.106>
- Oculus Quest 2: Nuestras gafas de VR todo en uno con la tecnología más avanzada*. (s. f.). Recuperado 29 de agosto de 2022, de <https://store.facebook.com/es/quest/products/quest-2/>

- Okawa, M. (2019). Template Matching Using Time-Series Averaging and DTW With Dependent Warping for Online Signature Verification. *IEEE Access*, 7, 81010-81019. <https://doi.org/10.1109/ACCESS.2019.2923093>
- Paneva, V., Seinfeld, S., Kraiczi, M., & Müller, J. (2020). HaptiRead: Reading Braille as Mid-Air Haptic Information. *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 13-20. <https://doi.org/10.1145/3357236.3395515>
- Park, J., Kim, H., Tai, Y.-W., Brown, M. S., & Kweon, I. (2011a). High quality depth map upsampling for 3D-TOF cameras. *2011 International Conference on Computer Vision*, 1623-1630. <https://doi.org/10.1109/ICCV.2011.6126423>
- Park, J., Kim, H., Tai, Y.-W., Brown, M. S., & Kweon, I. (2011b). High quality depth map upsampling for 3D-TOF cameras. *2011 International Conference on Computer Vision*, 1623-1630. <https://doi.org/10.1109/ICCV.2011.6126423>
- Patsadu, O., Nukoolkit, C., & Watanapa, B. (2012). Human gesture recognition using Kinect camera. *2012 Ninth International Conference on Computer Science and Software Engineering (JCSSE)*, 28-32. <https://doi.org/10.1109/JCSSE.2012.6261920>
- Pigou, L., Dieleman, S., Kindermans, P.-J., & Schrauwen, B. (2015). Sign Language Recognition Using Convolutional Neural Networks. En L. Agapito, M. M. Bronstein, & C. Rother (Eds.), *Computer Vision—ECCV 2014 Workshops* (pp. 572-578). Springer International Publishing. [https://doi.org/10.1007/978-3-319-16178-5\\_40](https://doi.org/10.1007/978-3-319-16178-5_40)
- Pisharady, P. K., & Saerbeck, M. (2015a). Recent methods and databases in vision-based hand gesture recognition: A review. *Computer Vision and Image Understanding*, 141, 152-165. <https://doi.org/10.1016/j.cviu.2015.08.004>
- Pisharady, P. K., & Saerbeck, M. (2015b). Recent methods and databases in vision-based hand gesture recognition: A review. *Computer Vision and Image Understanding*, 141, 152-165. <https://doi.org/10.1016/j.cviu.2015.08.004>
- Piyush Shanker, A., & Rajagopalan, A. N. (2007). Off-line signature verification using DTW. *Pattern Recognition Letters*, 28(12), 1407-1414. <https://doi.org/10.1016/j.patrec.2007.02.016>

- Premaratne, P., Nguyen, Q., & Premaratne, M. (2010a). Human Computer Interaction Using Hand Gestures. En D.-S. Huang, M. McGinnity, L. Heutte, & X.-P. Zhang (Eds.), *Advanced Intelligent Computing Theories and Applications* (pp. 381-386). Springer. [https://doi.org/10.1007/978-3-642-14831-6\\_51](https://doi.org/10.1007/978-3-642-14831-6_51)
- Premaratne, P., Nguyen, Q., & Premaratne, M. (2010b). Human Computer Interaction Using Hand Gestures. En D.-S. Huang, M. McGinnity, L. Heutte, & X.-P. Zhang (Eds.), *Advanced Intelligent Computing Theories and Applications* (pp. 381-386). Springer. [https://doi.org/10.1007/978-3-642-14831-6\\_51](https://doi.org/10.1007/978-3-642-14831-6_51)
- Przybyla, R. J., Shelton, S. E., Guedes, A., Krigel, R., Horsley, D. A., & Boser, B. E. (2012a). IN-AIR ULTRASONIC RANGEFINDING AND ANGLE ESTIMATION USING AN ARRAY OF ALN MICROMACHINED TRANSDUCERS. *2012 Solid-State, Actuators, and Microsystems Workshop Technical Digest*, 50-53. <https://doi.org/10.31438/trf.hh2012.14>
- Przybyla, R. J., Shelton, S. E., Guedes, A., Krigel, R., Horsley, D. A., & Boser, B. E. (2012b). IN-AIR ULTRASONIC RANGEFINDING AND ANGLE ESTIMATION USING AN ARRAY OF ALN MICROMACHINED TRANSDUCERS. *2012 Solid-State, Actuators, and Microsystems Workshop Technical Digest*, 50-53. <https://doi.org/10.31438/trf.hh2012.14>
- Przybyla, R. J., Tang, H.-Y., Shelton, S. E., Horsley, D. A., & Boser, B. E. (2014a). 12.1 3D ultrasonic gesture recognition. *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, 210-211. <https://doi.org/10.1109/ISSCC.2014.6757403>
- Przybyla, R. J., Tang, H.-Y., Shelton, S. E., Horsley, D. A., & Boser, B. E. (2014b). 12.1 3D ultrasonic gesture recognition. *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, 210-211. <https://doi.org/10.1109/ISSCC.2014.6757403>
- Pu, Q., Gupta, S., Gollakota, S., & Patel, S. (2013a). Whole-home gesture recognition using wireless signals. *Proceedings of the 19th annual international conference on Mobile computing & networking*, 27-38. <https://doi.org/10.1145/2500423.2500436>
- Pu, Q., Gupta, S., Gollakota, S., & Patel, S. (2013b). Whole-home gesture recognition using wireless signals. *Proceedings of the 19th annual international conference on Mobile computing & networking*, 27-38. <https://doi.org/10.1145/2500423.2500436>

- 
- Raets, C., Aisati, C. E., Rifi, A. L., Barbé, K., & Ridder, M. D. (2023). Predicting the Response to Chemoradiotherapy in Rectal Cancer Patients Using Bayesian Evolutionary Random Forest and Three-Dimensional Discrete Fourier Transform. *2023 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 1-5. <https://doi.org/10.1109/MeMeA57477.2023.10171859>
- Raheja, J. L., Minhas, M., Prashanth, D., Shah, T., & Chaudhary, A. (2015a). Robust gesture recognition using Kinect: A comparison between DTW and HMM. *Optik*, 126(11), Article 11. <https://doi.org/10.1016/j.ijleo.2015.02.043>
- Raheja, J. L., Minhas, M., Prashanth, D., Shah, T., & Chaudhary, A. (2015b). Robust gesture recognition using Kinect: A comparison between DTW and HMM. *Optik*, 126(11), 1098-1104. <https://doi.org/10.1016/j.ijleo.2015.02.043>
- Rakkolainen, I., Freeman, E., Sand, A., Raisamo, R., & Brewster, S. (2021). A Survey of Mid-Air Ultrasound Haptics and Its Applications. *IEEE Transactions on Haptics*, 14(1), 2-19. <https://doi.org/10.1109/TOH.2020.3018754>
- Rakthanmanon, T., Campana, B., Mueen, A., Batista, G., Westover, B., Zhu, Q., Zakaria, J., & Keogh, E. (2012). Searching and mining trillions of time series subsequences under dynamic time warping. *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 262-270. <https://doi.org/10.1145/2339530.2339576>
- Raptis, M., Kirovski, D., & Hoppe, H. (2011a). Real-time classification of dance gestures from skeleton animation. *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 147-156. <https://doi.org/10.1145/2019406.2019426>
- Raptis, M., Kirovski, D., & Hoppe, H. (2011b). Real-time classification of dance gestures from skeleton animation. *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 147-156. <https://doi.org/10.1145/2019406.2019426>
- Rath, T. M., & Manmatha, R. (2003). Word image matching using dynamic time warping. *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, 2, II-II. <https://doi.org/10.1109/CVPR.2003.1211511>
-

## **Bibliografía**

---

- Reaz, M. B. I., Hussain, M. S., & Mohd-Yasin, F. (2006a). Techniques of EMG signal analysis: Detection, processing, classification and applications. *Biological Procedures Online*, 8(1), Article 1. <https://doi.org/10.1251/bpo115>
- Reaz, M. B. I., Hussain, M. S., & Mohd-Yasin, F. (2006b). Techniques of EMG signal analysis: Detection, processing, classification and applications. *Biological Procedures Online*, 8(1), Article 1. <https://doi.org/10.1251/bpo115>
- Ren, Z., Yuan, J., Meng, J., & Zhang, Z. (2013). Robust Part-Based Hand Gesture Recognition Using Kinect Sensor. *IEEE Transactions on Multimedia*, 15(5), 1110-1120. <https://doi.org/10.1109/TMM.2013.2246148>
- Reyes, M., Domínguez, G., & Escalera, S. (2011). Featureweighting in dynamic timewarping for gesture recognition in depth data. *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, 1182-1188. <https://doi.org/10.1109/ICCVW.2011.6130384>
- Riofrío, S., Pozo, D., Rosero, J., & Vásquez, J. (2017). Gesture Recognition Using Dynamic Time Warping and Kinect: A Practical Approach. *2017 International Conference on Information Systems and Computer Science (INCISCOS)*, 302-308. <https://doi.org/10.1109/INCISCOS.2017.36>
- Rishan, R. M., Jayalal, S., & Wijayasiriwardhane, T. K. (2022). Translation of Sri Lankan Sign Language to Sinhala Text: A Leap Motion Technology-based Approach. *2022 2nd International Conference on Advanced Research in Computing (ICARC)*, 218-223. <https://doi.org/10.1109/ICARC54489.2022.9754050>
- Robeck, T. R., Haghani, A., Fei, Z., Lindemann, D. M., Russell, J., Herrick, K. E. S., Montano, G., Steinman, K. J., Katsumata, E., Zoller, J. A., & Horvath, S. (2023). Multi-tissue DNA methylation aging clocks for sea lions, walruses and seals. *Communications Biology*, 6(1), Article 1. <https://doi.org/10.1038/s42003-023-04734-0>
- Rosenberg, R. (1998). *The biofeedback pointer: EMG control of a two dimensional pointer*. 162-163.
- Saad, M., Bleakley, C. J., Nigram, V., & Kettle, P. (2018). Ultrasonic hand gesture recognition for mobile devices. *Journal on Multimodal User Interfaces*, 12(1), Article 1. <https://doi.org/10.1007/s12193-017-0257-8>

- Saad, M., Yang, T., & Zhou, H. (2022). A Comparison of Bidirectional GRU and LSTM for Hand Gesture Recognition Using Leap Motion. *2022 37th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 1427-1433. <https://doi.org/10.1109/YAC57282.2022.10023591>
- Sang, Y., Shi, L., & Liu, Y. (2018). Micro Hand Gesture Recognition System Using Ultrasonic Active Sensing. *IEEE Access*, 6, 49339-49347. <https://doi.org/10.1109/ACCESS.2018.2868268>
- Santos, L., González, J. L., Turiel, J. P., Fraile, J. C., & de la Fuente, E. (2015a). *Libro de actas*. Actas de las XXXVI Jornadas de Automática. <https://www.ehu.es/documents/3444171/4484745/48.pdf>
- Santos, L., González, J. L., Turiel, J. P., Fraile, J. C., & de la Fuente, E. (2015b). *Libro de actas*. Actas de las XXXVI Jornadas de Automática. <https://www.ehu.es/documents/3444171/4484745/48.pdf>
- Saponas, T. S., Tan, D. S., Morris, D., Balakrishnan, R., Turner, J., & Landay, J. A. (2009). Enabling always-available input with muscle-computer interfaces. *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, 167-176. <https://doi.org/10.1145/1622176.1622208>
- Sarhan, N. A., El-Sonbaty, Y., & Youssef, S. M. (2015). HMM-based Arabic Sign Language Recognition using Kinect. *2015 Tenth International Conference on Digital Information Management (ICDIM)*, 169-174. <https://doi.org/10.1109/ICDIM.2015.7381873>
- Sathyanarayanan, M., & Rajan, S. (2016a). *MYO Armband for physiotherapy healthcare: A case study using gesture recognition application*. 1-6. <https://doi.org/10.1109/comsnets.2016.7439933>
- Sathyanarayanan, M., & Rajan, S. (2016b). *MYO Armband for physiotherapy healthcare: A case study using gesture recognition application*. *2016 8th International Conference on Communication Systems and Networks (COMSNETS)*, 1-6. <https://doi.org/10.1109/COMSNETS.2016.7439933>
- Sesli, E. (2023). Human-Robot Interaction (HRI) through hand gestures for possible future war robots: A leap motion controller application. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-023-15278-0>
- Sharma, R. P., & Verma, G. K. (2015). Human Computer Interaction using Hand Gesture. *Procedia Computer Science*, 54, 721-727. <https://doi.org/10.1016/j.procs.2015.06.085>
- Sikdar, S., Rangwala, H., Eastlake, E. B., Hunt, I. A., Nelson, A. J., Devanathan, J., Shin, A., & Pancrazio, J. J. (2014a). Novel Method for Predicting Dexterous Individual Finger Movements by Imaging Muscle

- Activity Using a Wearable Ultrasonic System. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(1), Article 1. <https://doi.org/10.1109/TNSRE.2013.2274657>
- Sikdar, S., Rangwala, H., Eastlake, E. B., Hunt, I. A., Nelson, A. J., Devanathan, J., Shin, A., & Pancrazio, J. J. (2014b). Novel Method for Predicting Dexterous Individual Finger Movements by Imaging Muscle Activity Using a Wearable Ultrasonic System. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(1), 69-76. <https://doi.org/10.1109/TNSRE.2013.2274657>
- Simos, M., & Nikolaidis, N. (2016a). Greek sign language alphabet recognition using the leap motion device. *Proceedings of the 9th Hellenic Conference on Artificial Intelligence*, 1-4. <https://doi.org/10.1145/2903220.2903249>
- Simos, M., & Nikolaidis, N. (2016b). Greek sign language alphabet recognition using the leap motion device. *Proceedings of the 9th Hellenic Conference on Artificial Intelligence*, 1-4. <https://doi.org/10.1145/2903220.2903249>
- Simos, M., & Nikolaidis, N. (2016c). Greek sign language alphabet recognition using the leap motion device. *Proceedings of the 9th Hellenic Conference on Artificial Intelligence*, 1-4. <https://doi.org/10.1145/2903220.2903249>
- Singam, A. K., & Wlode, K. (2023). *Revised One, a Full Reference Video Quality Assessment Based on Statistical Based Transform Coefficient* (SSRN Scholarly Paper 4511335). <https://doi.org/10.2139/ssrn.4511335>
- Singh, S. K., & Chaturvedi, A. (2023). A reliable and efficient machine learning pipeline for american sign language gesture recognition using EMG sensors. *Multimedia Tools and Applications*, 82(15), Article 15. <https://doi.org/10.1007/s11042-022-14117-y>
- Singhal, Y., Wang, H., Gil, H., & Kim, J. R. (2021). Mid-Air Thermo-Tactile Feedback using Ultrasound Haptic Display. *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, 1-11. <https://doi.org/10.1145/3489849.3489889>
- Smeragliuolo, A. H., Hill, N. J., Disla, L., & Putrino, D. (2016). Validation of the Leap Motion Controller using marked motion capture technology. *Journal of Biomechanics*, 49(9), 1742-1750. <https://doi.org/10.1016/j.jbiomech.2016.04.006>

- Sturman, D. J., & Zeltzer, D. (1994a). A survey of glove-based input. *IEEE Computer Graphics and Applications*, 14(1), Article 1. <https://doi.org/10.1109/38.250916>
- Sturman, D. J., & Zeltzer, D. (1994b). A survey of glove-based input. *IEEE Computer Graphics and Applications*, 14(1), 30-39. <https://doi.org/10.1109/38.250916>
- Sun, C., Zhang, T., & Xu, C. (2015). Latent Support Vector Machine Modeling for Sign Language Recognition with Kinect. *ACM Transactions on Intelligent Systems and Technology*, 6, 20:1-20:20. <https://doi.org/10.1145/2629481>
- Sutherland, I. E. (1963). Sketchpad: A man-machine graphical communication system. *Universidad de Cambridge, UCAM-CL-TR-574(574)*, Article 574.
- Tabor, A., Bateman, S., & Scheme, E. (2016). Game-Based Myoelectric Training. *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, 299-306. <https://doi.org/10.1145/2968120.2987731>
- Tahmoush, D., & Silvius, J. (2009a). Radar micro-Doppler for long range front-view gait recognition. *Proceedings of the 3rd IEEE international conference on Biometrics: Theory, applications and systems*, 346-351.
- Tahmoush, D., & Silvius, J. (2009b). Radar micro-Doppler for long range front-view gait recognition. *Proceedings of the 3rd IEEE international conference on Biometrics: Theory, applications and systems*, 346-351.
- Tanaka, A., & Knapp, R. B. (2002). Multimodal Interaction in Music Using the Electromyogram and Relative Position Sensing. *Virginia Tech*. <http://hdl.handle.net/10919/80530>
- Tao, W., Lai, Z.-H., Leu, M. C., & Yin, Z. (2018a). *American Sign Language Alphabet Recognition Using Leap Motion Controller*. 599-604. <https://www.proquest.com/scholarly-journals/american-sign-language-alphabet-recognition-using/docview/2553578468/se-2>
- Tao, W., Lai, Z.-H., Leu, M. C., & Yin, Z. (2018b). *American Sign Language Alphabet Recognition Using Leap Motion Controller*. 599-604. <https://www.proquest.com/scholarly-journals/american-sign-language-alphabet-recognition-using/docview/2553578468/se-2>

## **Bibliografía**

---

- Tive, F. H. C., Bouzerdoum, A., & Amin, M. G. (2010a). A human gait classification method based on radar Doppler spectrograms. *EURASIP Journal on Advances in Signal Processing*, 2010, 10:1-10:10. <https://doi.org/10.1155/2010/389716>
- Tive, F. H. C., Bouzerdoum, A., & Amin, M. G. (2010b). A human gait classification method based on radar Doppler spectrograms. *EURASIP Journal on Advances in Signal Processing*, 2010, 10:1-10:10. <https://doi.org/10.1155/2010/389716>
- Tracking | Leap Motion Controller | Ultraleap*. (s. f.). Recuperado 29 de agosto de 2022, de <https://www.ultraleap.com/product/leap-motion-controller/>
- Trucco, E., & Verri, A. (1998a). *Introductory techniques for 3-D computer vision, 1st ed.* Prentice Hall Englewood Cliffs.
- Trucco, E., & Verri, A. (1998b). *Introductory techniques for 3-D computer vision, 1st ed.* Prentice Hall Englewood Cliffs.
- Tuzcu, V., & Nas, S. (2005). Dynamic time warping as a novel tool in pattern recognition of ECG changes in heart rhythm disturbances. *2005 IEEE International Conference on Systems, Man and Cybernetics*, 1, 182-186 Vol. 1. <https://doi.org/10.1109/ICSMC.2005.1571142>
- Utilización de la lengua de signos por sexo y edad. Población de 6 y más años con discapacidad de audición*. (s. f.). INE. Recuperado 9 de febrero de 2023, de <https://www.ine.es/jaxi/Datos.htm?tpx=51628>
- Vaitkevičius, A., Taroza, M., Blažauskas, T., Damaševičius, R., Maskeliūnas, R., & Woźniak, M. (2019a). Recognition of American Sign Language Gestures in a Virtual Reality Using Leap Motion. *Applied Sciences*, 9(3), Article 3. <https://doi.org/10.3390/app9030445>
- Vaitkevičius, A., Taroza, M., Blažauskas, T., Damaševičius, R., Maskeliūnas, R., & Woźniak, M. (2019b). Recognition of American Sign Language Gestures in a Virtual Reality Using Leap Motion. *Applied Sciences*, 9(3), Article 3. <https://doi.org/10.3390/app9030445>
- Vaitkevičius, A., Taroza, M., Blažauskas, T., Damaševičius, R., Maskeliūnas, R., & Woźniak, M. (2019c). Recognition of American Sign Language Gestures in a Virtual Reality Using Leap Motion. *Applied Sciences*, 9(3), Article 3. <https://doi.org/10.3390/app9030445>

- Valentinuzzi, M. E. (2007a). Bioelectrical signal processing in cardiac and neurological applications and electromyography: Physiology, engineering, and noninvasive applications. *BioMedical Engineering OnLine*, 6(1), Article 1. <https://doi.org/10.1186/1475-925X-6-27>
- Valentinuzzi, M. E. (2007b). Bioelectrical signal processing in cardiac and neurological applications and electromyography: Physiology, engineering, and noninvasive applications. *BioMedical Engineering OnLine*, 6(1), 27. <https://doi.org/10.1186/1475-925X-6-27>
- Van den Bogaert, L., Rutten, I., & Geerts, D. (2021). Augmenting Gestural Interactions with Mid-Air Haptic Feedback: A Case Study of Mixed-Method Longitudinal UX-Testing in the Lab. En E. Karapanos, J. Gerken, J. Kjeldskov, & M. B. Skov (Eds.), *Advances in Longitudinal HCI Research* (pp. 207-225). Springer International Publishing. [https://doi.org/10.1007/978-3-030-67322-2\\_10](https://doi.org/10.1007/978-3-030-67322-2_10)
- Virtual Reality | HTC España*. (s. f.). Recuperado 29 de agosto de 2022, de <https://www.htc.com/es/virtual-reality/>
- Vosinakis, S., & Koutsabasis, P. (2018). Evaluation of visual feedback techniques for virtual grasping with bare hands using Leap Motion and Oculus Rift. *Virtual Reality*, 22(1), 47-62. <https://doi.org/10.1007/s10055-017-0313-4>
- Wang, C., Liu, Z., & Chan, S.-C. (2015). Superpixel-Based Hand Gesture Recognition With Kinect Depth Camera. *IEEE Transactions on Multimedia*, 17(1), 29-39. <https://doi.org/10.1109/TMM.2014.2374357>
- Wang, Y., Liang, L., Zhang, Z., Xu, X., Liu, R., Fang, H., Zhang, R., Wei, Y., Liu, Z., Zhu, R., Zhang, X., & Wang, F. (2023). Fast and accurate assessment of depression based on voice acoustic features: A cross-sectional and longitudinal study. *Frontiers in Psychiatry*, 14. <https://www.frontiersin.org/articles/10.3389/fpsy.2023.1195276>
- Wang, Y., Yang, C., Wu, X., Xu, S., & Li, H. (2012). Kinect Based Dynamic Hand Gesture Recognition Algorithm Research. *2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics*, 1, 274-279. <https://doi.org/10.1109/IHMSC.2012.76>

## **Bibliografía**

---

- Want, R., Fishkin, K. P., Gujar, A., & Harrison, B. L. (1999). Bridging physical and virtual worlds with electronic tags. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 370-377. <https://doi.org/10.1145/302979.303111>
- Webster, J. G. (Ed.). (2009). *MEDICAL INSTRUMENTATION APPLICATION AND DESIGN*. John Wiley & Sons, Ltd.
- Weichert, F., Bachmann, D., Rudak, B., & Fisseler, D. (2013). Analysis of the Accuracy and Robustness of the Leap Motion Controller. *Sensors*, 13(5), Article 5. <https://doi.org/10.3390/s130506380>
- Welbanks, L., McGill, P., Line, M., & Madhusudhan, N. (2023). On the Application of Bayesian Leave-one-out Cross-validation to Exoplanet Atmospheric Analysis. *The Astronomical Journal*, 165(3), 112. <https://doi.org/10.3847/1538-3881/acab67>
- Wheeler, K. R., & Jorgensen, C. C. (2003). Gestures as input: Neuroelectric joysticks and keyboards. *IEEE Pervasive Computing*, 2(2), 56-61. <https://doi.org/10.1109/MPRV.2003.1203754>
- Wu, B., Lu, Z., & Yang, C. (2022). A Modified LSTM Model for Chinese Sign Language Recognition Using Leap Motion. *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1612-1617. <https://doi.org/10.1109/SMC53654.2022.9945287>
- Wu, C., & Aghajan, H. (2007). Model-based human posture estimation for gesture analysis in an opportunistic fusion smart camera network. *2007 IEEE Conference on Advanced Video and Signal Based Surveillance*, 453-458. <https://doi.org/10.1109/AVSS.2007.4425353>
- Wu, K., Xiao, J., Yi, Y., Gao, M., & Ni, L. M. (2012a). FILA: Fine-grained indoor localization. *2012 Proceedings IEEE INFOCOM*, 2210-2218. <https://doi.org/10.1109/INFCOM.2012.6195606>
- Wu, K., Xiao, J., Yi, Y., Gao, M., & Ni, L. M. (2012b). FILA: Fine-grained indoor localization. *2012 Proceedings IEEE INFOCOM*, 2210-2218. <https://doi.org/10.1109/INFCOM.2012.6195606>
- Xiao, H., Wang, W., Bao, H., Li, F., & Zhou, L. (2023). Biodiesel-diesel blend optimized via leave-one cross-validation based on kinematic viscosity, calorific value, and flash point. *Industrial Crops and Products*, 191, 115914. <https://doi.org/10.1016/j.indcrop.2022.115914>
- Xie, P., Cui, Z., Du, Y., Zhao, M., Cui, J., Wang, B., & Hu, X. (2023). Multi-scale local-temporal similarity fusion for continuous sign language recognition. *Pattern Recognition*, 136, 109233. <https://doi.org/10.1016/j.patcog.2022.109233>

- Yang, L., Chen, J., & Zhu, W. (2020a). Dynamic Hand Gesture Recognition Based on a Leap Motion Controller and Two-Layer Bidirectional Recurrent Neural Network. *Sensors*, 20(7), Article 7. <https://doi.org/10.3390/s20072106>
- Yang, L., Chen, J., & Zhu, W. (2020b). Dynamic Hand Gesture Recognition Based on a Leap Motion Controller and Two-Layer Bidirectional Recurrent Neural Network. *Sensors*, 20(7), Article 7. <https://doi.org/10.3390/s20072106>
- Yauri Vidalón, J. E., & De Martino, J. M. (2016). Brazilian Sign Language Recognition Using Kinect. En G. Hua & H. Jégou (Eds.), *Computer Vision – ECCV 2016 Workshops* (pp. 391-402). Springer International Publishing. [https://doi.org/10.1007/978-3-319-48881-3\\_27](https://doi.org/10.1007/978-3-319-48881-3_27)
- Young, G., Milne, H., Griffiths, D., Padfield, E., Blenkinsopp, R., & Georgiou, O. (2020). Designing Mid-Air Haptic Gesture Controlled User Interfaces for Cars. *Proceedings of the ACM on Human-Computer Interaction*, 4(EICS), 81:1-81:23. <https://doi.org/10.1145/3397869>
- Zafrulla, Z., Brashear, H., Starner, T., Hamilton, H., & Presti, P. (2011). American sign language recognition with the kinect. *Proceedings of the 13th international conference on multimodal interfaces*, 279-286. <https://doi.org/10.1145/2070481.2070532>
- Zhang, Y., & Meruvia-Pastor, O. (2017). Operating Virtual Panels with Hand Gestures in Immersive VR Games. En L. T. De Paolis, P. Bourdot, & A. Mongelli (Eds.), *Augmented Reality, Virtual Reality, and Computer Graphics* (pp. 299-308). Springer International Publishing. [https://doi.org/10.1007/978-3-319-60922-5\\_24](https://doi.org/10.1007/978-3-319-60922-5_24)
- Zheng, Y. P., Chan, M. M. F., Shi, J., Chen, X., & Huang, Q. H. (2006). Sonomyography: Monitoring morphological changes of forearm muscles in actions with the feasibility for the control of powered prosthesis. *Medical Engineering & Physics*, 28(5), 405-415. <https://doi.org/10.1016/j.medengphy.2005.07.012>
- Zhu, J., Jia, M., Zhang, Y., Deng, H., & Liu, Y. (2023). Transductive transfer broad learning for cross-domain information exploration and multigrade soft sensor application. *Chemometrics and Intelligent Laboratory Systems*, 235, 104778. <https://doi.org/10.1016/j.chemolab.2023.104778>

## ***Bibliografía***

---

Zimmerman, T. G. (1985). *Optical flex sensor* (United States Patent US4542291A).

<https://patents.google.com/patent/US4542291A/en>

Zou, Y., Xiao, J., Han, J., Wu, K., Li, Y., & Ni, L. M. (2017). GRfid: A Device-Free RFID-Based Gesture Recognition

System. *IEEE Transactions on Mobile Computing*, 16(2), Article 2.

<https://doi.org/10.1109/TMC.2016.2549518>

# Apéndice I



## Apéndice I

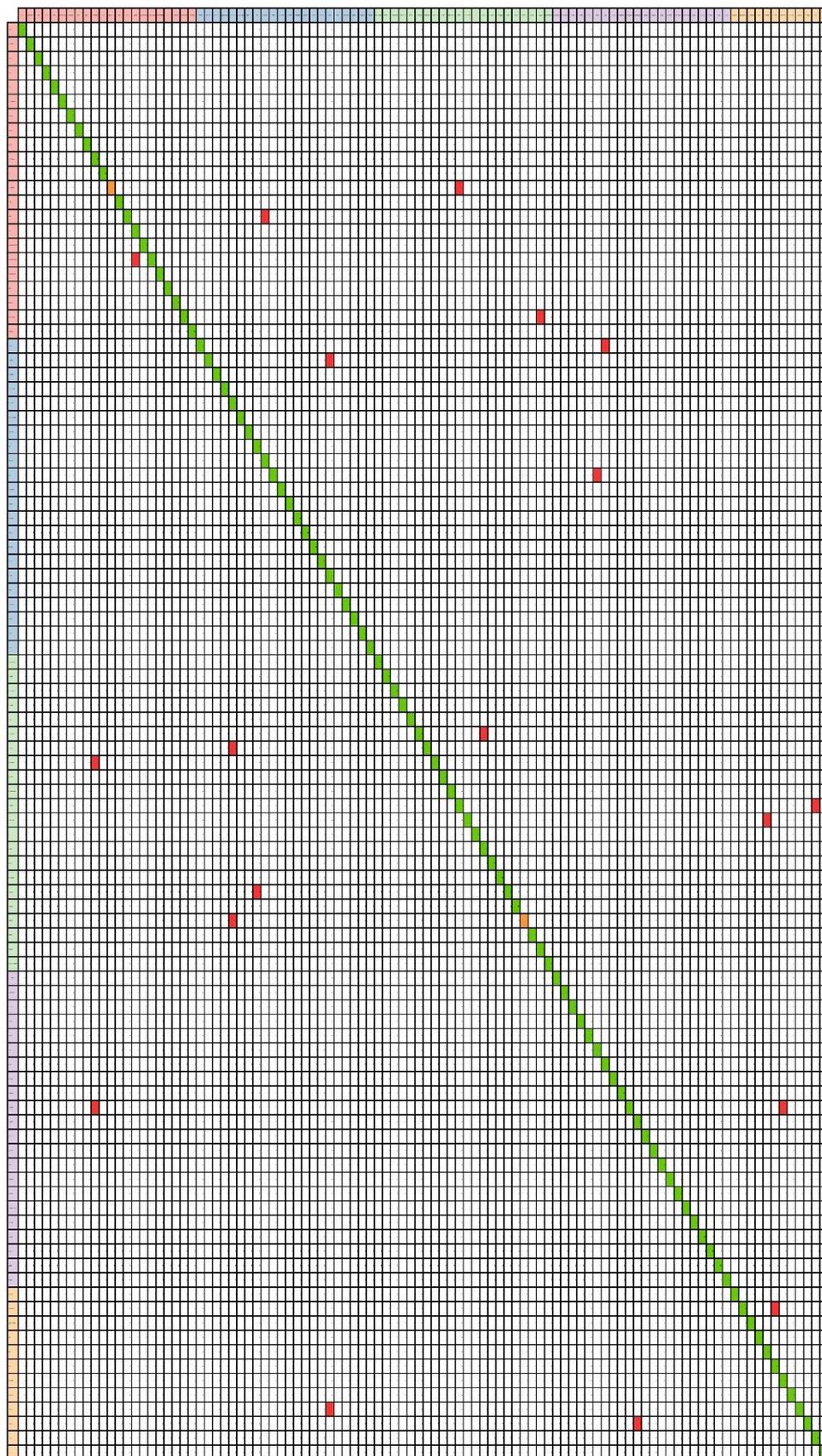


Figura AI.1. Matriz de confusión de 100 palabras conjunto 1

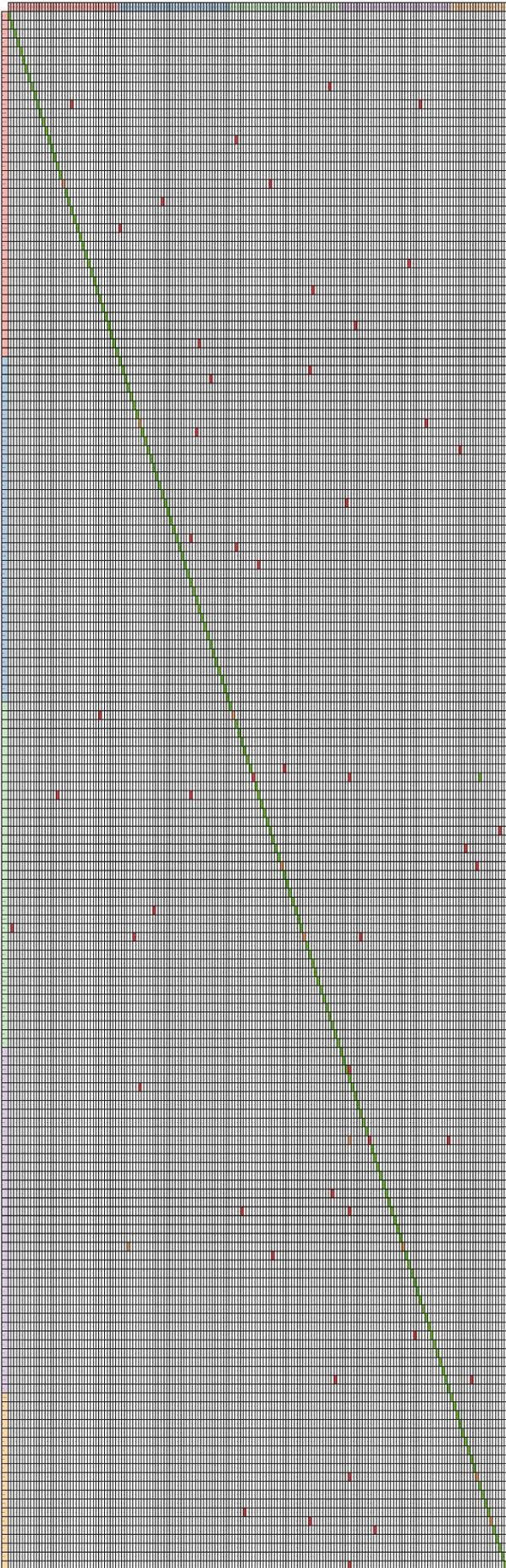


Figura AI.2. Matriz de confusión de 176 palabras conjunto 1

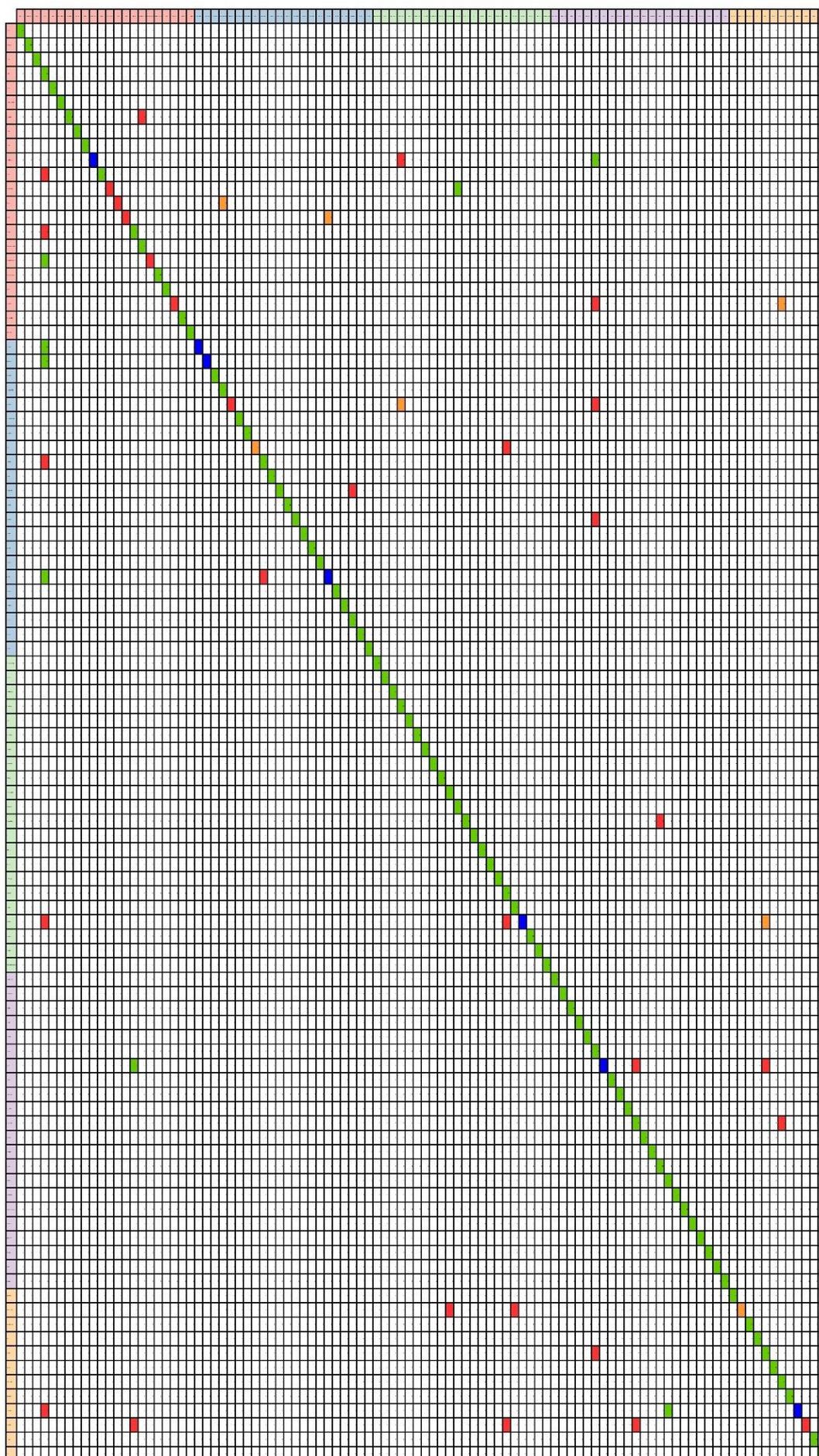


Figura A1.3. Matriz de confusión de 100 palabras conjunto 2

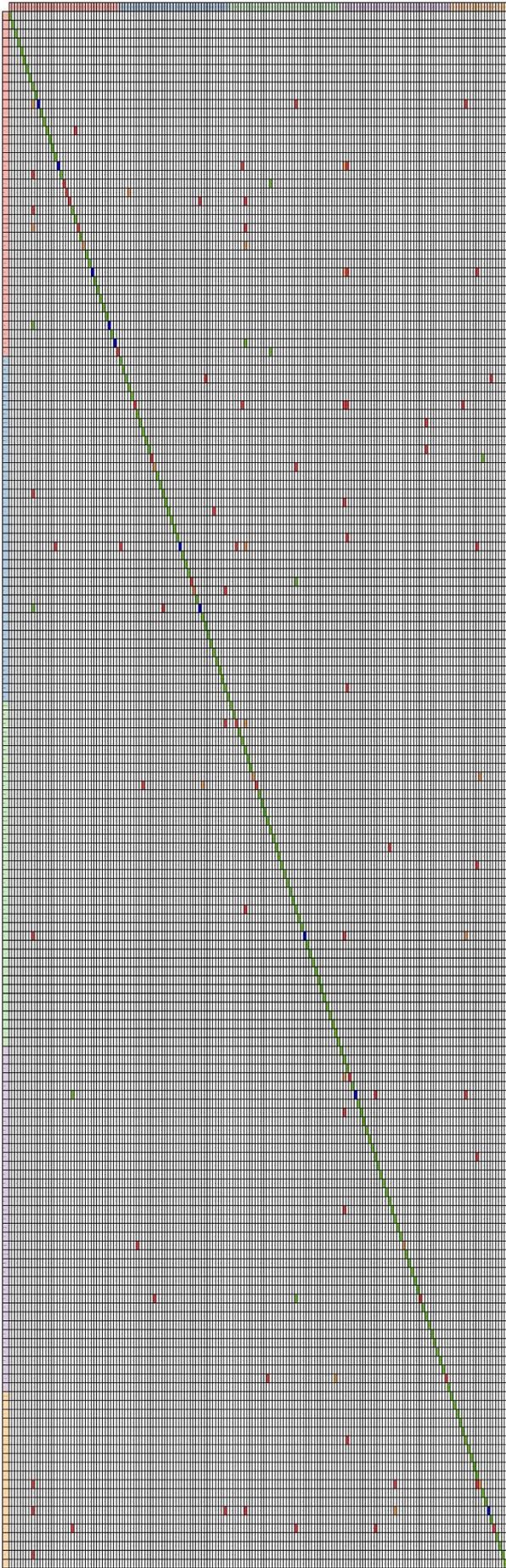


Figura AI.4. Matriz de confusión de 176 palabras conjunto 2

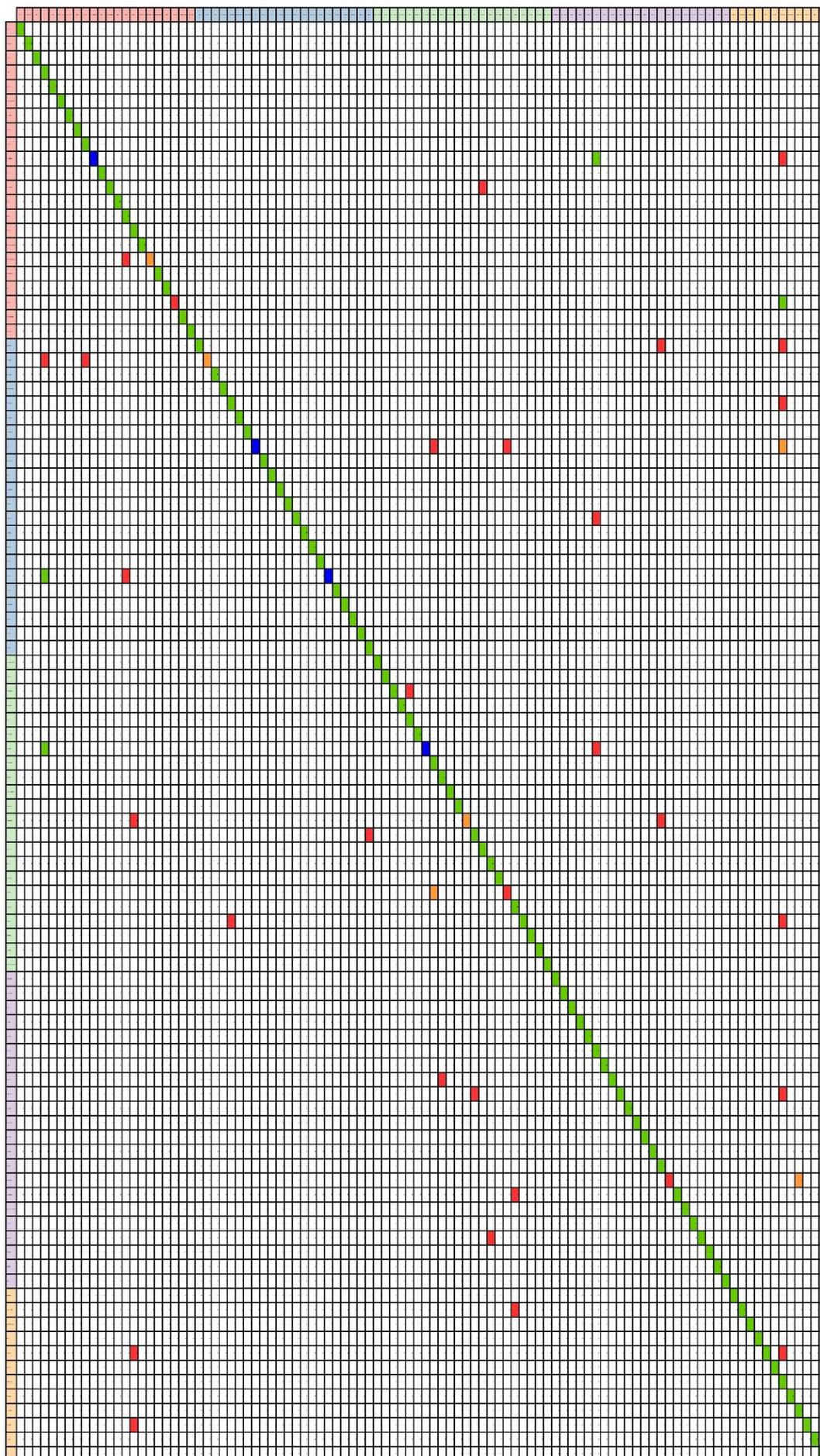


Figura A1.5. Matriz de confusión de 100 palabras conjunto 3

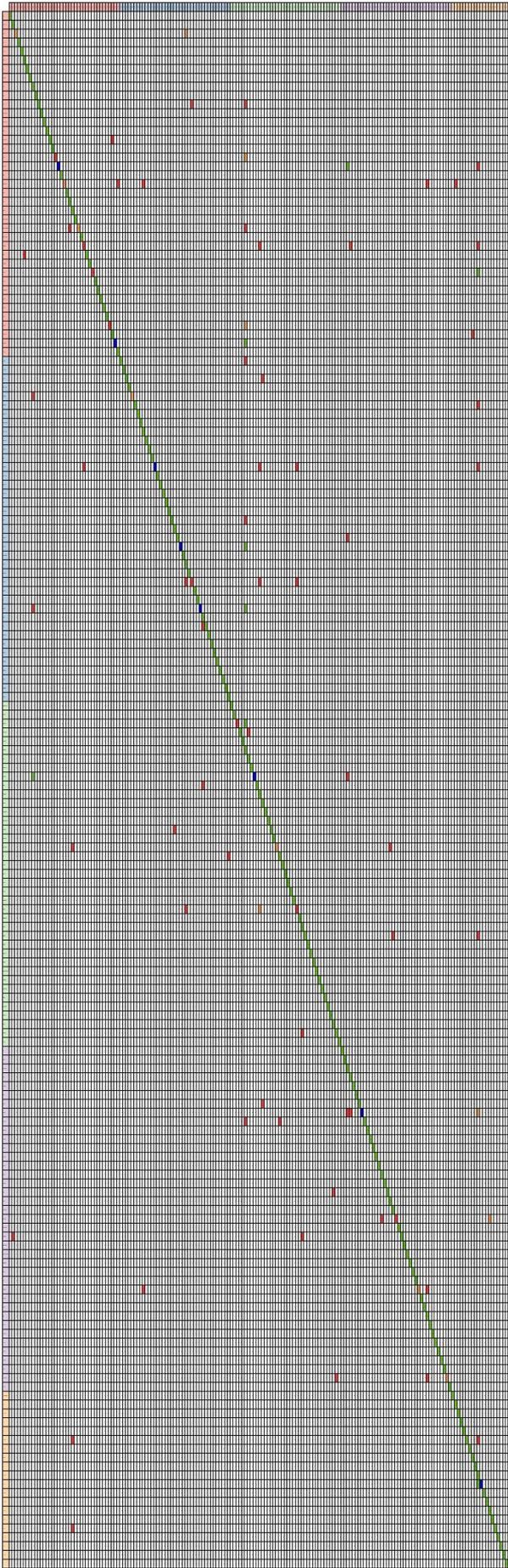


Figura AI.6. Matriz de confusión de 176 palabras conjunto 3

# Apéndice II



## Apéndice II

### Figuras

- Figura 1.1.** *Persona signando una palabra.* Fuente: <https://www.youtube.com/> (Recuperado 26 de septiembre de 2023)
- Figura 1.2.** *Dispositivos diseñados hasta el momento para el reconocimiento gestual.* Fuente: Elaboración Propia.
- Figura 2.1.** *Diferentes usos del Lenguaje de Signos.* Fuente: Imágenes públicas de internet.
- Figura 2.2.** *Escritura sobre Sketchpad en 1963.* Fuente: <http://augmented-photography.ch/documents/meaa-01-techniques-of-the-observer> (Recuperado 26 de septiembre de 2023).
- Figura 2.3.** *Guantes Pasivos que ayudan a diferenciar la posición de los dedos.* Fuente: <https://www.radiantvisionsystems.com/blog/wave-your-hand-3d-gesture-recognition-systems> (Recuperado 26 de septiembre de 2023).
- Figura 2.4.** *CyberGlove III usado para el control del Robots.* Fuente: <http://www.cyberglovesystems.com/cyberglove-iii> (Recuperado 26 de septiembre de 2023).
- Figura 2.5.** *Fases para el Tratamiento de las señales EMG.* Fuente: Elaboración Propia.
- Figura 2.6.** *Reconocimiento de diferentes posiciones de manos usando WiFi.* Fuente: <https://www.kurzweilai.net/using-your-wifi-for-gesture-recognition> (Recuperado 26 de septiembre de 2023).
- Figura 2.7.** *Etapas de los Sistemas Basados en la Visión.* Fuente: Elaboración Propia.
- Figura 2.8.** *Modo de Funcionamiento del Leap Motion.* Fuente: <https://proyectoidis.org/leap-motion/> (Recuperado 26 de septiembre de 2023).
- Figura 3.1.** *Diferentes componentes del sistema de hardware de Leap Motion.* Fuente: <https://www.allaboutcircuits.com/news/teardown-tuesday-leap-motion-controller/> y <https://www.mdpi.com/1424-8220/13/5/6380> (Recuperado 26 de septiembre de 2023).
- Figura 3.2.** *Área de interacción del Leap Motion.* Fuente: [https://developer-archive.leapmotion.com/documentation/v2/javascript/devguide/Leap\\_Coordinate\\_Mapping.html](https://developer-archive.leapmotion.com/documentation/v2/javascript/devguide/Leap_Coordinate_Mapping.html) (Recuperado 26 de septiembre de 2023).
- Figura 3.3.** *Objeto Frame del Leap Motion.* Fuente: <https://xiaotingrunning.wordpress.com/2015/03/23/leap-motion-tutorial-java2-general-introduction-and-hello-world/> (Recuperado 26 de septiembre de 2023).
- Figura 3.4.** *Seguimiento del modelo Esquelético de las Manos.* Fuente 1: [https://developer-archive.leapmotion.com/documentation/javascript/devguide/Leap\\_Overview.html](https://developer-archive.leapmotion.com/documentation/javascript/devguide/Leap_Overview.html). (Recuperado 26 de septiembre de 2023). Fuente 2: <https://ieeexplore.ieee.org/abstract/document/8538425> (Recuperado 26 de septiembre de 2023).
- Figura 3.5.** *Distorsión que las lentes pueden producir.* Fuente: <https://www.xatakafoto.com/guias/distorsion-de-lente-vs-distorsion-de-la-perspectiva> (Recuperado 26 de septiembre de 2023).
- Figura 3.6.** *Imagen con la Distorsión Corregida antes de llegar al Microcontrolador.* Fuente: [https://developer-archive.leapmotion.com/documentation/v2/python/devguide/Leap\\_Images.html](https://developer-archive.leapmotion.com/documentation/v2/python/devguide/Leap_Images.html) (Recuperado 26 de septiembre de 2023).
- Figura 3.7.** *Grabación de muestras.* Fuente: Elaboración Propia.
- Figura 3.8.** *Muestra del archivo del signo hola.* Fuente: Elaboración Propia.
- Figura 3.9.** *Movimiento de la mano.* Fuente: [https://developer-archive.leapmotion.com/documentation/javascript/devguide/Leap\\_Overview.html](https://developer-archive.leapmotion.com/documentation/javascript/devguide/Leap_Overview.html) (Recuperado 26 de septiembre de 2023).

**Figura 3.10.** *Ejes del Leap Motion.* Fuente: [https://developer-archive.leapmotion.com/documentation/javascript/devguide/Leap\\_Overview.html](https://developer-archive.leapmotion.com/documentation/javascript/devguide/Leap_Overview.html) (Recuperado 26 de septiembre de 2023).

**Figura 3.11.** *Ejemplo de una ruta de deformación en el tiempo.* Fuente: [https://www.cs.ucr.edu/~eamonn/pkdd99keogh\\_pazzani.pdf](https://www.cs.ucr.edu/~eamonn/pkdd99keogh_pazzani.pdf) (Recuperado 26 de septiembre de 2023).

**Figura 3.12.** *Comparador DTW.* Fuente: Elaboración Propia.

**Figura 3.13.** *Esquema del comparador.* Fuente: Elaboración Propia.

**Figura 4.1.** *Validación cruzada dejando a uno fuera.* Fuente: [https://es.wikipedia.org/wiki/Validaci%C3%B3n\\_cruzada](https://es.wikipedia.org/wiki/Validaci%C3%B3n_cruzada) (Recuperado 26 de septiembre de 2023).

**Figura 4.2.** *Procedimiento utilizado para el testeo de las muestras y ejecución de la validación cruzada.* Fuente: Elaboración Propia.

**Figura 4.3.** *Sistema Implementado.* Fuente: Elaboración Propia.

**Figura 4.4.** *Modelo gráfico de los parámetros.* Fuente: Elaboración Propia.

**Figura 4.5.** *Modelo gráfico del parámetro de precisión.* Fuente: Elaboración Propia.

**Figura 4.6.** *Modelo gráfico del parámetro de recuperación o "recall".* Fuente: Elaboración Propia.

**Figura 4.7.** *Matriz de confusión multiclase. (a) Clase I. (b) Clase N-1.* Fuente: Elaboración Propia.

**Figura 4.8.** *Matriz de confusión de 50 palabras.* Fuente: Elaboración Propia.

**Figura 5.1.** *Matriz de confusión de 50 palabras conjunto 1.* Fuente: Elaboración Propia.

**Figura 5.2.** *Matriz de confusión 50 palabras con muestras externas conjunto 1.* Fuente: Elaboración Propia.

**Figura 5.3.** *Matriz de confusión de 50 palabras conjunto 2.* Fuente: Elaboración Propia.

**Figura 5.4.** *Matriz de confusión 50 palabras con muestras externas conjunto 2.* Fuente: Elaboración Propia.

**Figura 5.5.** *Matriz de confusión de 50 palabras conjunto 3.* Fuente: Elaboración Propia.

**Figura 5.6.** *Matriz de confusión 50 palabras con muestras externas conjunto 3.* Fuente: Elaboración Propia.

**Figura AI.1.** *Matriz de confusión de 100 palabras conjunto 1.* Fuente: Elaboración Propia.

**Figura AI.2.** *Matriz de confusión de 176 palabras conjunto 1.* Fuente: Elaboración Propia.

**Figura AI.3.** *Matriz de confusión de 100 palabras conjunto 2.* Fuente: Elaboración Propia.

**Figura AI.2.** *Matriz de confusión de 176 palabras conjunto 2.* Fuente: Elaboración Propia.

**Figura AI.2.** *Matriz de confusión de 100 palabras conjunto 3.* Fuente: Elaboración Propia.

**Figura AI.2.** *Matriz de confusión de 176 palabras conjunto 3.* Fuente: Elaboración Propia.

# Apéndice III



## Apéndice III

indra

Transportes Tráfico Aéreo Defensa y Seguridad Minsait Talento @ Sostenibilidad Sobre Indra

# Facilitar la lectura a personas con discapacidad intelectual, dar voz a sordomudos y predecir crisis de autismo, más cerca gracias a Indra y Fundación Universia

Noticias

Eventos

Contenidos multimedia

Blog Neo

Contacto Prensa

3 Diciembre 2018 España

- Son los objetivos que persiguen las tres propuestas seleccionadas en la III Convocatoria de ayudas a proyectos de investigación en Tecnologías Accesibles, lanzada por Indra y Fundación Universia, a la que se han presentado decenas de iniciativas de universidades de toda España
- Como empresa socialmente responsable, Indra promueve el desarrollo de Tecnologías Accesibles que contribuyan a mejorar la calidad de vida, el acceso a la educación y la integración social y laboral de las personas con discapacidad



Coincidiendo con el Día Mundial de la Discapacidad, Indra, una de las principales compañías globales de tecnología y consultoría, y Fundación Universia, que cuenta con el apoyo de Banco Santander para el desarrollo de sus actividades, han dado a conocer las tres propuestas seleccionadas en la III Convocatoria de ayudas a proyectos de investigación en Tecnologías Accesibles, que lanzan anualmente.

**Figura C.1.** Noticia publicada en la página web de Indra. Fuente:

<https://www.indracompany.com/es/noticia/facilitar-lectura-personas-discapacidad-intelectual-voz-sordomudos-predecir-crisis-autismo> (Recuperado 26 de septiembre de 2023)

Open Access Review

## Perspective and Evolution of Gesture Recognition for Sign Language: A Review

by  Jesús Galván-Ruiz <sup>1,2</sup> ,  Carlos M. Travieso-González <sup>1,2,\*</sup>  ,  Acaymo Tejera-Fettmilch <sup>1</sup> ,  
 Alejandro Pinan-Roescher <sup>1</sup> ,  Luis Esteban-Hernández <sup>3</sup>   and  Luis Domínguez-Quintana <sup>2</sup>  

<sup>1</sup> IDeTIC, Universidad de Las Palmas de Gran Canaria, 35017 Las Palmas de G.C., Spain

<sup>2</sup> Signals and Communications Department, Universidad de Las Palmas de G.C. (ULPGC), 35017 Las Palmas de G.C., Spain

<sup>3</sup> Aries Research Center, Universidad Antonio Nebrija, 28015 Madrid, Spain

\* Author to whom correspondence should be addressed.

*Sensors* **2020**, *20*(12), 3571; <https://doi.org/10.3390/s20123571>

Received: 18 May 2020 / Revised: 22 June 2020 / Accepted: 22 June 2020 / Published: 24 June 2020

(This article belongs to the Section **State-of-the-Art Sensors Technologies**)

Download

Browse Figures

Review Reports

Versions Notes

### Abstract

This review analyses the different gesture recognition systems through a timeline, showing the different types of technology, and specifying which are the most important features and their achieved recognition rates. At the end of the review, Leap Motion sensor possibilities are described in detail, in order to consider its application on the field of sign language. This device has many positive characteristics that make it a good option for sign language. One of the most important conclusions is the ability of the Leap Motion sensor to provide 3D information from the hands for due identification.

**Keywords:** gesture recognition; algorithms; Leap Motion; pattern recognition; EMG; RFID; gloves; Wi-Fi

**Figura C.2.** Artículo publicado en la revista *Sensors*. Fuente: <https://www.mdpi.com/1424-8220/20/12/3571>

(Recuperado 26 de septiembre de 2023)

Open Access Article

## Robust Identification System for Spanish Sign Language Based on Three-Dimensional Frame Information

by  Jesús Galván-Ruiz<sup>1</sup>,  Carlos M. Travieso-González<sup>1,2,\*</sup>  ,  Alejandro Pinan-Roescher<sup>1</sup> and  Jesús B. Alonso-Hernández<sup>1,2</sup> 

<sup>1</sup> IDeTIC, Universidad de Las Palmas de G.C. (ULPGC), 35017 Las Palmas de Gran Canaria, Spain

<sup>2</sup> Signals and Communications Department, Universidad de Las Palmas de G.C. (ULPGC), 35017 Las Palmas de Gran Canaria, Spain

\* Author to whom correspondence should be addressed.

*Sensors* **2023**, *23*(1), 481; <https://doi.org/10.3390/s23010481>

Received: 6 December 2022 / Revised: 23 December 2022 / Accepted: 30 December 2022 /

Published: 2 January 2023

(This article belongs to the Topic *Advances in Intelligent Biosignals Processing and Analysis*)

[Download](#)[Browse Figures](#)[Review Reports](#)[Versions Notes](#)

### Abstract

Nowadays, according to the World Health Organization (WHO), of the world's population suffers from a hearing disorder that makes oral communication with other people challenging. At the same time, in an era of technological evolution and digitization, designing tools that could help these people to communicate daily is the base of much scientific research such as that discussed herein. This article describes one of the techniques designed to transcribe Spanish Sign Language (SSL). A Leap Motion volumetric sensor has been used in this research due to its capacity to recognize hand movements in 3 dimensions. In order to carry out this research project, an impaired hearing subject has collaborated in the recording of 176 dynamic words. Finally, for the development of the research, Dynamic Time Warping (DTW) has been used to compare the samples and predict the input with an accuracy of 95.17%.

**Keywords:** dynamic time warping; Spanish sign language; gesture recognition; pattern recognition

**Figura C.3.** Artículo publicado en la revista *Sensors*. Fuente: <https://www.mdpi.com/1424-8220/23/1/481>

(Recuperado 26 de septiembre de 2023)

## Doce iniciativas premiadas en la III edición del concurso de la Cátedra Telefónica de la ULPGC

12 JUN 2020

Compartir en las redes:   

La Cátedra Telefónica de la ULPGC ha premiado 12 iniciativas en la III edición de su concurso dirigido a los miembros de la comunidad universitaria para servir como estímulo a la comunidad universitaria y a la divulgación social de la actividad de Investigación, Innovación y Cultura Científica en la ULPGC. Las propuestas debían ir orientadas a ideas, soluciones, trabajos y proyectos cercanos a necesidades sociales y se han entregado premios en cuatro subgrupos diferentes.

**Figura C.4.** Imagen de la página de la Universidad de Las Palmas de G.C. relativa al III edición del concurso de la Cátedra de Telefónica. Fuente: <https://www.ulpgc.es/noticia/doce-iniciativas-premiadas-iii-edicion-del-concurso-catedra-telefonica-ulpgc>. (Recuperado 26 de septiembre de 2023)

# Un prototipo que da voz a las manos

**Ciencia.** Un grupo de investigadores de la ULPGC diseña un programa que convierte la lengua de signos en sonido

CANARIAS7

Un equipo de investigadores de la Universidad de Las Palmas de Gran Canaria (ULPGC), liderado por el catedrático Carlos M. Travieso-González, y compuesto por los investigadores Jesús Galván Ruiz, Alejandro Piñán Roescher, Acaymo Tejera Fetmilch y Joel José Gutiérrez Cortadellas, ha desarrollado un prototipo de dispositivo que «traduce» la lengua de signos a voz sintetizada.

El proyecto 'DaVoz' ha sido financiado dentro de la tercer convocatoria de ayudas a proyectos de investigación aplicados a la creación de nuevas tecnologías accesibles, que cuenta con el apoyo económico de Indra y la Fundación Univerisia, y cuyo objetivo es promover el desarrollo de soluciones tecnológicas innovadoras que mejoren la calidad de vida y la integración social y laboral de las personas con discapacidad, según señala la ULPGC en un comunicado.

'DaVoz' se basa en un sistema de interpretación automática de lengua de signos, mediante el uso de un sensor volumétrico e inteligencia artificial, capaz de identificar los signos y transformarlos en voz sintetizada. Uno de los aspectos fuertes de esta novedosa propuesta es el uso de componentes tecnológicos de bajo coste, así como su fácil manejo y su portabilidad, que le permite ser utilizado en cualquier ordenador portátil estándar. Resulta, pues, una solución eficiente que



Imagen del equipo de investigadores. Abajo, una de las pruebas del «traductor».



permite al colectivo de personas sordas una mayor autonomía al romper con la dependencia constante de la figura del intérprete.

Cabe destacar la fiabilidad

del sensor volumétrico, al emplearse para convertir el movimiento de tres falanges, el carpio y el metacarpiano de cada dedo a series temporales tridimensionales. De igual forma,

ha sido clave el uso de inteligencia artificial para que esa señal multidimensional generada por el sensor pueda ser interpretada y transformada a voz sintetizada. El dispositivo muestra una interesante versatilidad, ya que puede emplearse así para otros idiomas introduciendo los parámetros adecuados.

El equipo investigador que ha desarrollado 'DaVoz' ha contado con la colaboración continua de la Asociación de Personas Sordas de la Provincia de Las Palmas (ASOR), tanto en la captación de información del lenguaje de signos, como en los comentarios que han aportado para el diseño, mejora y adaptabilidad del prototipo hasta lograr la versión actual.

Figura C.5. Artículo publicado en el diario Canarias7 edición impresa el 21 de noviembre de 2020.



Figura C.6. Telenoticias I de la Radio Televisión Canaria (RTVC) del 11 de diciembre de 2020 (del minuto 40:20 al minuto 42:00). Fuente <https://www.youtube.com/watch?v=e4tPaHiNjos&t=2444s> (Recuperado 26 de septiembre de 2023)

**Otros enlaces de interés donde se hace eco de la investigación desarrollada:**

- Diario la Vanguardia digital  
<https://www.lavanguardia.com/vida/20201120/49554187160/la-ulpgc-disena-un-prototipo-que-traduce-lengua-de-signos-a-voz-sintetizada.html> (Recuperado 26 de septiembre de 2023)
- EsRadio (con Carmen Tomás)  
[https://www.ivoox.com/entrevista-a-carlos-manuel-travieso-audios-mp3\\_rf\\_69863849\\_1.html](https://www.ivoox.com/entrevista-a-carlos-manuel-travieso-audios-mp3_rf_69863849_1.html)  
(Recuperado 26 de septiembre de 2023)
- Diario La Provincia  
<https://www.laprovincia.es/sociedad/2020/11/20/ulpgc-disena-prototipo-traduce-lengua-23869794.html> (Recuperado 26 de septiembre de 2023)
- Diario El Plural  
[https://www.elplural.com/el-telescopio/tech/davoz-solucion-inteligencia-artificial-interpretar-lenguaje-signos\\_262188102](https://www.elplural.com/el-telescopio/tech/davoz-solucion-inteligencia-artificial-interpretar-lenguaje-signos_262188102) (Recuperado 26 de septiembre de 2023)
- Diario Europa Press  
<https://www.europapress.es/epsocial/responsables/noticia-dispositivo-da-voz-personas-sordas-interpretar-lengua-signos-20210318100635.html> (Recuperado 26 de septiembre de 2023)
- Diario Siglo XXI

- 
- <https://www.diariosigloxxi.com/texto-ep/mostrar/20210318100638/indra-fundacion-universia-impulsan-dispositivo-da-voz-personas-sordas-interpretar-lengua-signos> (Recuperado 26 de septiembre de 2023)
- ULPGC  
<https://www.ulpgc.es/noticia/indra-y-fundacion-universia-impulsan-dispositivo-creado-ulpgc-que-da-voz-personas-sordas-al> (Recuperado 26 de septiembre de 2023)
  - ULPGC  
<https://www.ulpgc.es/noticia/ulpgc-disena-prototipo-que-traduce-lengua-signos-voz-sintetizada> (Recuperado 26 de septiembre de 2023)
  - El Digital (ULPGC)  
<https://eldigital.ulpgc.es/noticia/2021/01/15/dando-voz-las-personas-sordas> (Recuperado 26 de septiembre de 2023)
  - CIVAT del Cabildo de Tenerife  
<https://civat.es/noticia/un-nuevo-dispositivo-da-voz-a-las-personas-sordas-al-interpretar-la-lengua-de-signos/> (Recuperado 26 de septiembre de 2023)
  - Cadena Valor  
<https://cadenadevalor.es/indra-fundacion-universia-davoz/> (Recuperado 26 de septiembre de 2023)
  - Sol Senior  
<https://www.sol senior.com/proyecto-davoz> (Recuperado 26 de septiembre de 2023)
  - Tododisca  
<https://www.tododisca.com/dispositivo-da-voz-personas-sordas/> (Recuperado 26 de septiembre de 2023)
  - INDRA  
<https://www.indracompany.com/es/noticia/indra-fundacion-universia-impulsan-dispositivo-creado-ulpgc-voz-personas-sordas-interpretar> (Recuperado 26 de septiembre de 2023)
  - GN Diario  
<https://www.gndiario.com/davoz-sordas-personas-lengua-signos> (Recuperado 26 de septiembre de 2023)
  - Itusers  
<https://itusers.today/impulsan-dispositivo-davoz-que-utiliza-inteligencia-artificial-para-brindar-voz-a-personas-sordas/> (Recuperado 26 de septiembre de 2023)
  - Servimedia  
<https://itusers.today/impulsan-dispositivo-davoz-que-utiliza-inteligencia-artificial-para-brindar-voz-a-personas-sordas/> (Recuperado 26 de septiembre de 2023)
  - Business Empresarial
-

- <https://www.businessempresarial.com.pe/e/davoz/> (Recuperado 26 de septiembre de 2023)
- Technocio  
<https://technocio.com/dispositivo-da-voz-a-personas-sordas-al-interpretar-la-lengua-de-signos/>  
(Recuperado 26 de septiembre de 2023)
- Website Editor  
[https://cdn.website-editor.net/856d2dfc2a4c404eb89f72273584c8c4/files/uploaded/Savanah%2520n%25C2%25BA%2520200\\_gNr26tY3QfOoSvebiKEz.pdf](https://cdn.website-editor.net/856d2dfc2a4c404eb89f72273584c8c4/files/uploaded/Savanah%2520n%25C2%25BA%2520200_gNr26tY3QfOoSvebiKEz.pdf) (Recuperado 26 de septiembre de 2023)
- Ser Peruano  
<https://www.serperuano.com/2021/03/indra-y-fundacion-universia-impulsan-un-dispositivo-creado-por-la-ulpgc-que-da-voz-a-las-personas-sordas-al-interpretar-la-lengua-de-signos/>  
(Recuperado 26 de septiembre de 2023)
- Cio Perú  
<https://cioperu.pe/articulo/32045/presentan-dispositivo-capaz-de-interpretar-la-lengua-de-signos/>  
(Recuperado 26 de septiembre de 2023)
- Red Universidad y Discapacidad  
[http://red-universidadydiscapacidad.org/documentos/2021\\_04\\_boletin\\_abril\\_2021.pdf](http://red-universidadydiscapacidad.org/documentos/2021_04_boletin_abril_2021.pdf)  
(Recuperado 26 de septiembre de 2023)
- Market Screener  
<https://es.marketscreener.com/cotizacion/accion/INDRA-SISTEMAS-S-A-413470/noticias/Indra-Sistemas-S-A-y-Fundaci-n-Universia-impulsan-un-dispositivo-creado-por-la-ULPGC-que-ldquo-da-32717829/> (Recuperado 26 de septiembre de 2023)
- Acis (Colombia)  
<https://acis.org.co/portal/content/noticiasdelsector/indra-y-fundaci%C3%B3n-universia-impulsan-un-dispositivo-creado-por-la-ulpgc-que-%E2%80%9Cda-voz%E2%80%9D-las> (Recuperado 26 de septiembre de 2023)
- Comunitaria  
<https://comunitaria.com/indra-y-fundacion-universia-impulsan-un-dispositivo-da-voz-personas-sordas/> (Recuperado 26 de septiembre de 2023)
- Spains News  
<https://spainsnews.com/the-ulpgc-designs-a-prototype-that-translates-sign-language-into-synthesized-speech/> (Recuperado 26 de septiembre de 2023)
- Technopatas  
<https://www.technopatas.com/indra-y-fundacion-universia-impulsan-un-dispositivo-creado-por-la-ulpgc-que-da-voz-a-las-personas-sordas-al-interpretar-la-lengua-de-signos/davoz1/> (Recuperado 26 de septiembre de 2023)

- Publicación de la VI Edición de la Convocatoria Indra y Fundación Universia, como modelo o ejemplo <https://www.indracompany.com/es/noticia/indra-fundacion-universia-apoyan-desarrollo-tecnologias-inclusion-personas-discapacidad> (Recuperado 26 de septiembre de 2023)

**Enlaces de su repercusión en redes sociales:**

- Twitter  
<https://twitter.com/funiversia/status/1372129972198989825> (Recuperado 26 de septiembre de 2023)  
<https://www.trendsmap.com/twitter/tweet/1372861456920756224> (Recuperado 26 de septiembre de 2023)  
<https://www.trendsmap.com/twitter/tweet/1372134832101277696> (Recuperado 26 de septiembre de 2023)
- Facebook  
<https://m.facebook.com/economiadehoy.es/posts/2940383279541006> (Recuperado 26 de septiembre de 2023)  
<https://www.facebook.com/fundacionuniversia/photos/a.219894641354641/4172142916129774/?type=3> (Recuperado 26 de septiembre de 2023)
- Otros  
<http://prensainfosord.blogspot.com/2021/01/dando-voz-las-personas-sordas.html> (Recuperado 26 de septiembre de 2023)