



**Desarrollo de un sistema de rastreo de puntos clave de las manos en  
expresiones básicas de la lengua de señas ecuatoriana.**

Aguiar Vega, Mayra Carolina

Departamento de Eléctrica, Electrónica y Telecomunicaciones

Carrera de Telecomunicaciones

Trabajo de integración curricular, previo a la obtención del título de  
Ingeniera en Telecomunicaciones

Ing. Larco Bravo, Julio César. Msc

25 de febrero del 2026

# AguiarVegaCarolina\_Trabajo\_TitulacionV2

**9%**  
Textos  
sospechosos



**< 1%** Similitudes  
 0 % similitudes entre comillas  
 0 % entre las fuentes mencionadas  
 0 % Idiomas no reconocidos  
 9 % Textos potencialmente generados por la IA

Nombre del documento: AguiarVegaCarolina\_Trabajo\_TitulacionV2.pdf  
 ID del documento: 0eaa69c36f09123e4db1ee7cced9168373ee4ff1  
 Tamaño del documento original: 1,58 MB

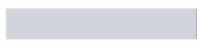
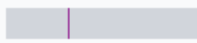
Depositante: JULIO CESAR LARCO BRAVO  
 Fecha de depósito: 25/2/2026  
 Tipo de carga: interface  
 fecha de fin de análisis: 25/2/2026

Número de palabras: 23.425  
 Número de caracteres: 163.847

Ubicación de las similitudes en el documento:



## Fuentes con similitudes fortuitas

| N° | Descripciones                                                                                                                                                                                                                                                                                                                       | Similitudes | Ubicaciones                                                                          | Datos adicionales                      |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|--------------------------------------------------------------------------------------|----------------------------------------|
| 1  |  <b>Alexander Pilamunga.pdf</b>   Alexander Pilamunga #d8cf58<br>Viene de de mi biblioteca                                                                                                                                                         | < 1%        |    | Palabras idénticas: < 1% (20 palabras) |
| 2  |  <b>Vasquez Anthony AvatarISEC.pdf</b>   Vasquez Anthony AvatarISEC #5e036a<br>Viene de de mi grupo                                                                                                                                                | < 1%        |    | Palabras idénticas: < 1% (12 palabras) |
| 3  |  <b>iaexpertos.com</b>   MediaPipe crea apps de IA para visión y audio - IA Expertos<br><a href="https://iaexpertos.com/mediapipe-crea-apps-de-ia-para-vision-y-audio/">https://iaexpertos.com/mediapipe-crea-apps-de-ia-para-vision-y-audio/</a> | < 1%        |   | Palabras idénticas: < 1% (12 palabras) |
| 4  |  <b>github.com</b>   GitHub - fpetitjean/DBA: DBA: Averaging for Dynamic Time Warping<br><a href="https://github.com/fpetitjean/DBA">https://github.com/fpetitjean/DBA</a>                                                                       | < 1%        |  | Palabras idénticas: < 1% (11 palabras) |

## Fuentes mencionadas (sin similitudes detectadas) Estas fuentes han sido citadas en el documento sin encontrar similitudes.

-  <https://drive.google.com/drive/folders/1ZdHjoC7KswLjn5Nr5dn5TUIUbwUI963I>
-  <https://pypi.org/project/dtw/>

**Ing. Larco Bravo, Julio César. Msc**

**Director**



## **Departamento de Eléctrica, Electrónica y Telecomunicaciones**

### **Carrera de Telecomunicaciones**

#### **Certificación**

Certifico que el trabajo de integración curricular: “**Desarrollo de un sistema de rastreo de puntos clave de las manos en expresiones básicas de la lengua de señas ecuatoriana**” fue realizado por la señorita **Aguiar Vega, Mayra Carolina**, el mismo que cumple con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, además fue revisado y analizada en su totalidad por la herramienta de prevención y/o verificación de similitud de contenidos; razón por la cual me permito acreditar y autorizar para que se lo sustente públicamente.

**Sangolquí, 25 de febrero del 2026**

**Ing. Larco Bravo, Julio César Msc**

**Director**

**C.C 1710638808**



## **Departamento de Eléctrica, Electrónica y Telecomunicaciones**

### **Carrera de Telecomunicaciones**

#### **Responsabilidad de Autoría**

Yo, **Aguiar Vega, Mayra Carolina**, con cédula de ciudadanía No. 1752052025 declaro que el contenido, ideas y criterios del trabajo de integración curricular: **“Desarrollo de un sistema de rastreo de puntos clave de las manos en expresiones básicas de la lengua de señas ecuatoriana”** es de mi autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos, y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

**Sangolquí, 25 de febrero del 2026**

**Aguiar Vega, Mayra Carolina**

C.C. 1752052025



**Departamento de Eléctrica, Electrónica y Telecomunicaciones**

**Carrera de Telecomunicaciones**

**Autorización de Publicación**

Yo **Aguiar Vega, Mayra Carolina**, con cédula de ciudadanía **1752052025**, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de integración curricular: **“Desarrollo de un sistema de rastreo de puntos clave de las manos en expresiones básicas de la lengua de señas ecuatoriana”** en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi responsabilidad.

**Sangolquí, 25 de febrero del 2026**

**Aguiar Vega, Mayra Carolina**

C.C. 1752052025

### **Dedicatoria**

A mis padres, por ser mi apoyo incondicional desde el inicio de este camino, brindándome amor, esfuerzo y confianza en cada etapa de mi vida; por ustedes y para ustedes realizo cada uno de mis esfuerzos. A mi madre, Nancy, mi luz y mi pilar fundamental, por ser mi fuerza constante y la motivación que impulsa mis ganas de superarme cada día. A mi hermana, Fernanda, por su apoyo diario y por no permitirme rendirme en los momentos difíciles. A mi abuelita, Gloria, por sus bendiciones y su amor incondicional. A mi sobrina, Eimmie Sofía, quien fue mi mayor inspiración en la etapa final de mi carrera. Y a todas las personas que formaron parte de este proceso, gracias por ser mi motor y la razón por la que cada día busco superarme.

**Mayra Carolina Aguiar Vega**

## **Agradecimiento**

A lo largo de este camino, he contado con el apoyo incondicional de muchas personas, a quienes deseo expresar mi más profundo agradecimiento.

Agradezco a Dios por darme la fortaleza y la sabiduría necesarias para no rendirme, y por acompañarme en cada desafío hasta alcanzar esta meta.

A mi familia, por su amor incondicional, su paciencia y su apoyo constante. Gracias por estar a mi lado en los momentos difíciles y por motivarme a seguir adelante cuando más lo necesitaba.

A mis amigos, por sus palabras de ánimo y por hacer más llevadero este proceso con su compañía. De manera especial, a Kevin, mi mejor amigo, por su apoyo incondicional, por escucharme, aconsejarme y sostenerme en los momentos en que sentía que no podía continuar.

Al, Ing. Julio César Larco, por su guía constante, su disposición y el tiempo dedicado durante el desarrollo de este trabajo. Gracias por compartir sus conocimientos, por cada observación constructiva y por su paciencia en cada etapa del proceso. Su orientación académica y profesional fue fundamental para fortalecer este proyecto y para mi crecimiento tanto personal como profesional.

**Mayra Carolina Aguiar Vega**

## Índice de Contenidos

|                                         |    |
|-----------------------------------------|----|
| Certificación .....                     | 3  |
| Responsabilidad de Autoría .....        | 4  |
| Autorización de Publicación .....       | 5  |
| Dedicatoria.....                        | 6  |
| Agradecimiento.....                     | 7  |
| Índice de Contenidos .....              | 8  |
| Índice de Tablas.....                   | 11 |
| Índice de Figuras .....                 | 14 |
| Resumen .....                           | 22 |
| Abstract.....                           | 23 |
| Abreviaciones .....                     | 24 |
| Capítulo 1 .....                        | 25 |
| Introducción .....                      | 25 |
| Antecedentes .....                      | 25 |
| Motivación .....                        | 27 |
| Justificación.....                      | 29 |
| Objetivos .....                         | 30 |
| Objetivo General.....                   | 30 |
| Objetivos Específicos.....              | 30 |
| Alcance .....                           | 31 |
| Estructura del documento.....           | 32 |
| Capítulo 2 .....                        | 33 |
| Fundamentos Teóricos .....              | 33 |
| Lengua de Señas Ecuatoriana (LSEC)..... | 33 |

|                                                                         |     |
|-------------------------------------------------------------------------|-----|
| Detección y rastreo de puntos clave de la mano (Keypoints Hands). ..... | 34  |
| MediaPipe. ....                                                         | 38  |
| Representación de gestos mediante series temporales.....                | 39  |
| Normalización de coordenadas. ....                                      | 41  |
| Firma gestual y modelo base de un gesto. ....                           | 44  |
| Distancia euclidiana. ....                                              | 47  |
| Deformación Temporal Dinámica (DTW). ....                               | 49  |
| Capítulo 3 .....                                                        | 54  |
| Implementación del Sistema .....                                        | 54  |
| Metodología .....                                                       | 54  |
| Extracción del modelo. ....                                             | 57  |
| Fuente de datos.....                                                    | 57  |
| Detección y extracción de puntos clave de la mano.....                  | 58  |
| Normalización de coordenadas.....                                       | 63  |
| Creación de firma individual.....                                       | 69  |
| Creación Modelo promedio euclidiana. ....                               | 80  |
| Creación modelo promedio DTW (DBA). ....                                | 87  |
| Validación de modelos promedio de gestos.....                           | 95  |
| Capítulo 4 .....                                                        | 116 |
| Validación del sistema .....                                            | 116 |
| Escenario Pruebas .....                                                 | 116 |
| Métricas de evaluación.....                                             | 118 |
| Resultados obtenidos de la evaluación. ....                             | 130 |
| Análisis de resultados.....                                             | 141 |
| Capítulo 5 .....                                                        | 143 |
| Conclusiones y Recomendaciones .....                                    | 143 |

|                        |     |
|------------------------|-----|
| Conclusiones.....      | 143 |
| Recomendaciones.....   | 144 |
| Trabajos Futuros ..... | 145 |
| Referencias.....       | 146 |

## Índice de Tablas

|                                                                                                                                     |     |
|-------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>Tabla 1</b> Resultados Distancia Euclidiana respecto a las 7 componentes de desplazamiento.                                      | 103 |
| <b>Tabla 2</b> Resultados Similitud Distancia Euclidiana respecto a las 7 componentes de desplazamiento.....                        | 104 |
| <b>Tabla 3</b> Resultados Distancia Euclidiana con respecto a las 10 componentes de movimiento (punta de los dedos). ....           | 104 |
| <b>Tabla 4</b> Resultados Similitud Distancia Euclidiana con respecto a las 10 componentes de movimiento (punta de los dedos). .... | 105 |
| <b>Tabla 5</b> Distancia componentes de desplazamiento para DTW. ....                                                               | 113 |
| <b>Tabla 6</b> Distancia componentes de desplazamiento para DTW - Similitud.....                                                    | 113 |
| <b>Tabla 7</b> Distancia componentes de movimiento para DTW.....                                                                    | 114 |
| <b>Tabla 8</b> Distancia componentes de movimiento para DTW - Similitud. ....                                                       | 114 |
| <b>Tabla 9</b> Descripción de los escenarios de prueba.....                                                                         | 117 |
| <b>Tabla 10</b> Resultados pruebas Distancia componentes de desplazamiento - Método Euclidiano. ....                                | 131 |
| <b>Tabla 11</b> Resultados en similitud, pruebas componentes de desplazamiento - Método Euclidiano. ....                            | 131 |
| <b>Tabla 12</b> Resultados de las pruebas distancia de los componentes de movimiento - Método Euclidiano. ....                      | 131 |
| <b>Tabla 13</b> Resultados en similitud, pruebas componentes de movimiento - Método Euclidiano. ....                                | 132 |
| <b>Tabla 14</b> Resultados pruebas componentes de desplazamiento - Método DTW. ....                                                 | 132 |
| <b>Tabla 15</b> Resultados en similitud de las pruebas de los componentes de desplazamiento - Método DTW.....                       | 132 |
| <b>Tabla 16</b> Resultados Pruebas componentes de movimiento - Método DTW.....                                                      | 133 |

|                                                                                                                 |     |
|-----------------------------------------------------------------------------------------------------------------|-----|
| <b>Tabla 17</b> Resultados en similitud de las pruebas de los componentes de movimiento - Método DTW. ....      | 133 |
| <b>Tabla 18</b> Resultados pruebas distancia componentes de desplazamiento - Método Euclidiano. ....            | 133 |
| <b>Tabla 19</b> Resultados en similitud, pruebas de los componentes de desplazamiento - Método Euclidiano. .... | 134 |
| <b>Tabla 20</b> Resultados pruebas distancia componentes de movimiento - Método Euclidiano...                   | 134 |
| <b>Tabla 21</b> Resultados en similitud, pruebas componentes de movimiento - Método Euclidiano. ....            | 134 |
| <b>Tabla 22</b> Resultados de las pruebas de los componentes de desplazamiento - Método DTW. ....               | 135 |
| <b>Tabla 23</b> Resultados en similitud de las pruebas de los componentes de desplazamiento - Método DTW.....   | 135 |
| <b>Tabla 24</b> Resultados pruebas de los componentes de movimiento - Método DTW.....                           | 135 |
| <b>Tabla 25</b> Resultados en similitud de las pruebas componentes de movimiento - Método DTW. ....             | 136 |
| <b>Tabla 26</b> Resultados, pruebas distancia de los componentes de desplazamiento - Método Euclidiano. ....    | 136 |
| <b>Tabla 27</b> Resultados en similitud, pruebas de las componentes de desplazamiento - Método Euclidiano. .... | 136 |
| <b>Tabla 28</b> Resultados pruebas distancia componentes de movimiento - Método Euclidiano...                   | 137 |
| <b>Tabla 29</b> Resultados en similitud, pruebas de los componentes de movimiento - Método Euclidiano. ....     | 137 |
| <b>Tabla 30</b> Resultados de las pruebas de los componentes de desplazamiento - Método DTW. ....               | 137 |

|                                                                                                                 |     |
|-----------------------------------------------------------------------------------------------------------------|-----|
| <b>Tabla 31</b> Resultados en similitud de las pruebas componentes de desplazamiento - Método DTW. ....         | 138 |
| <b>Tabla 32</b> Resultados de las pruebas de los componentes de movimiento - Método DTW. ....                   | 138 |
| <b>Tabla 33</b> Resultados en similitud de las pruebas de los componentes de movimiento - Método DTW. ....      | 138 |
| <b>Tabla 34</b> Resultados de las pruebas de los componentes de desplazamiento - Método Euclidiano. ....        | 139 |
| <b>Tabla 35</b> Resultados en similitud, pruebas de las componentes de desplazamiento - Método Euclidiano. .... | 139 |
| <b>Tabla 36</b> Resultados de las pruebas de los componentes de movimiento - Método Euclidiano. ....            | 139 |
| <b>Tabla 37</b> Resultados en similitud, pruebas de las componentes de movimiento - Método Euclidiano. ....     | 140 |
| <b>Tabla 38</b> Resultados de las pruebas de las componentes de desplazamiento - Método DTW. ....               | 140 |
| <b>Tabla 39</b> Resultados en similitud de las pruebas de los componentes de desplazamiento - Método DTW. ....  | 140 |
| <b>Tabla 40</b> Resultados en similitud de las pruebas de los componentes de movimiento - Método DTW. ....      | 141 |
| <b>Tabla 41</b> Resultados en similitud de las pruebas de los componentes de movimiento - Método DTW. ....      | 141 |

## Índice de Figuras

|                  |                                                                                                                                       |    |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figura 1</b>  | Representación de los 21 puntos clave de una mano para el análisis gestual.....                                                       | 36 |
| <b>Figura 2</b>  | <i>Landmarks (21 puntos) en varios ejemplos de posiciones de la mano.</i> .....                                                       | 38 |
| <b>Figura 3</b>  | Normalización espacial de coordenadas de puntos clave en gestos manuales. ....                                                        | 41 |
| <b>Figura 4</b>  | Normalización espacial de coordenadas de puntos clave en gestos manuales. ....                                                        | 42 |
| <b>Figura 5</b>  | Proceso de normalización escalar de puntos clave de la mano.....                                                                      | 44 |
| <b>Figura 6</b>  | Promedio de series temporales mediante Promedio Aritmético.....                                                                       | 46 |
| <b>Figura 7</b>  | Promedio de series temporales mediante DTW Barycenter Averaging (DBA). ....                                                           | 47 |
| <b>Figura 8</b>  | Comparación de series temporales mediante distancia euclidiana.....                                                                   | 48 |
| <b>Figura 9</b>  | Matriz de costos acumulados y camino óptimo de alineamiento en DTW.....                                                               | 50 |
| <b>Figura 10</b> | Comparación entre la distancia DTW y la Distancia Euclidiana.....                                                                     | 53 |
| <b>Figura 11</b> | Diagrama general para el Método de Distancia Euclidiana.....                                                                          | 56 |
| <b>Figura 12</b> | Diagrama general para el Método DTW. ....                                                                                             | 56 |
| <b>Figura 13</b> | Participantes en la posición ideal de inicio del gesto. ....                                                                          | 58 |
| <b>Figura 14</b> | Detección de los 21 puntos claves de cada mano de un frame de video.....                                                              | 60 |
| <b>Figura 15</b> | Eje de coordenadas inicialmente en un frame de video.....                                                                             | 61 |
| <b>Figura 16</b> | Almacenamiento de las series temporales 126 datos por frame.....                                                                      | 63 |
| <b>Figura 17</b> | Normalización espacial de la mano .....                                                                                               | 65 |
| <b>Figura 18</b> | Normalización escalar de la mano.....                                                                                                 | 66 |
| <b>Figura 19</b> | Normalización escalar de la mano.....                                                                                                 | 68 |
| <b>Figura 20</b> | Proceso de normalización temporal. ....                                                                                               | 69 |
| <b>Figura 21</b> | Firma individual del desplazamiento global en el eje x de todos los puntos de la<br>mano derecha – gesto gracias de la persona 1..... | 77 |
| <b>Figura 22</b> | Firma individual del desplazamiento global en el eje y de todos los puntos de la<br>mano derecha – gesto gracias de la persona 1..... | 78 |

|                                                                                                                                              |    |
|----------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figura 23</b> Firma individual del desplazamiento del dedo anular derecho – gesto gracias de la persona 1.....                            | 78 |
| <b>Figura 24</b> Firma individual del desplazamiento del dedo índice derecho – gesto gracias de la persona 1.....                            | 78 |
| <b>Figura 25</b> Firma individual del desplazamiento del dedo medio derecho – gesto gracias de la persona 1.....                             | 79 |
| <b>Figura 26</b> Firma individual del desplazamiento del dedo meñique derecho – gesto gracias de la persona 1.....                           | 79 |
| <b>Figura 27</b> Firma individual del desplazamiento del dedo pulgar derecho – gesto gracias de la persona 1.....                            | 79 |
| <b>Figura 28</b> Firma individual del movimiento en x del dedo índice derecho – gesto gracias de la persona 1.....                           | 80 |
| <b>Figura 29</b> Firma individual del movimiento en y del dedo índice derecho – gesto gracias de la persona 1.....                           | 80 |
| <b>Figura 30</b> Modelo promedio del desplazamiento global en el eje x los 21 puntos de la mano derecha con IC del 95 % – gesto gracias..... | 83 |
| <b>Figura 31</b> Modelo promedio del desplazamiento global en el eje y de los 21 puntos la mano derecha con IC del 95 % – gesto gracias..... | 83 |
| <b>Figura 32</b> Modelo promedio del desplazamiento del dedo anular derecho con IC del 95 % – gesto gracias. ....                            | 84 |
| <b>Figura 33</b> Modelo promedio del desplazamiento del dedo índice derecho con IC del 95 % – gesto gracias. ....                            | 84 |
| <b>Figura 34</b> Modelo promedio del desplazamiento del dedo meñique derecho con IC del 95 % – gesto gracias. ....                           | 85 |
| <b>Figura 35</b> Modelo promedio del desplazamiento del dedo pulgar derecho con IC del 95 % – gesto gracias. ....                            | 85 |

|                                                                                                                                                        |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figura 36</b> Modelo promedio del desplazamiento del dedo medio derecho con IC del 95 % – gesto gracias. ....                                       | 86 |
| <b>Figura 37</b> Modelo promedio del movimiento en $x$ del dedo índice derecho con IC del 95 % – gesto gracias. ....                                   | 86 |
| <b>Figura 38</b> Modelo promedio del movimiento en $y$ del dedo índice derecho con IC del 95 % – gesto gracias. ....                                   | 87 |
| <b>Figura 39</b> Ejemplo de alineación temporal de 2 señales mediante DBA. ....                                                                        | 88 |
| <b>Figura 40</b> Modelo promedio obtenido del desplazamiento en $x$ de los 21 puntos de la mano derecha del gesto “gracias” mediante DTW (DBA). ....   | 90 |
| <b>Figura 41</b> Modelo promedio obtenido del desplazamiento en $y$ de los 21 puntos de la mano derecha del gesto “gracias” mediante DTW (DBA). ....   | 90 |
| <b>Figura 42</b> Modelo promedio obtenido del desplazamiento del dedo anular de la mano derecha del gesto “gracias” mediante DTW (DBA). ....           | 91 |
| <b>Figura 43</b> Modelo promedio obtenido del desplazamiento del dedo índice de la mano derecha del gesto “gracias” mediante DTW (DBA). ....           | 91 |
| <b>Figura 44</b> Modelo promedio obtenido del desplazamiento del dedo medio de la mano derecha del gesto “gracias” mediante DTW (DBA). ....            | 92 |
| <b>Figura 45</b> Modelo promedio obtenido del desplazamiento del dedo meñique de la mano derecha del gesto “gracias” mediante DTW (DBA). ....          | 92 |
| <b>Figura 46</b> Modelo promedio obtenido del desplazamiento del dedo pulgar de la mano derecha del gesto “gracias” mediante DTW (DBA). ....           | 93 |
| <b>Figura 47</b> Modelo promedio obtenido del movimiento en el eje $x$ del dedo índice de la mano derecha del gesto “gracias” mediante DTW (DBA). .... | 93 |
| <b>Figura 48</b> Modelo promedio obtenido del movimiento en el eje $y$ del dedo índice de la mano derecha del gesto “gracias” mediante DTW (DBA). .... | 94 |

|                                                                                                                                                                                             |           |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>Figura 49</b> Comparación de la firma test vs modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento en el eje $x$ global 21 puntos de la mano derecha..... | 97        |
| <b>Figura 50</b> Comparación de la firma test vs modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento en el eje $y$ global 21 puntos de la mano derecha..... | 97        |
| <b>Figura 51</b> Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento del dedo anular de la mano derecha.....            | 98        |
| <b>Figura 52</b> Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento del dedo índice de la mano derecha. ....           | 98        |
| <b>Figura 53</b> <i>Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento del dedo medio de la mano derecha. ....</i>     | <i>99</i> |
| <b>Figura 54</b> Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento del dedo meñique de la mano derecha. ....          | 99        |
| <b>Figura 55</b> Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento del dedo pulgar de la mano derecha.....            | 100       |
| <b>Figura 56</b> Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el movimiento del dedo medio en $x$ de la mano derecha.....          | 100       |
| <b>Figura 57</b> Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el movimiento del dedo medio en $y$ de la mano derecha.....          | 101       |
| <b>Figura 58</b> Comparación de la firma test y del modelo del gesto “gracias” mediante distancia euclidiana para el desplazamiento del dedo meñique de la mano derecha. ....               | 101       |
| <b>Figura 59</b> Comparación de la firma test y del modelo del gesto “gracias” mediante distancia euclidiana para el movimiento en $x$ del dedo índice de la mano derecha.....              | 102       |
| <b>Figura 60</b> Comparación de 10 firmas test vs modelo del gesto “gracias” mediante distancia euclidiana para el movimiento en $x$ del dedo índice de la mano derecha. ....               | 102       |

|                                                                                                                                                                                                                            |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>Figura 61</b> Comparación de 10 firmas test y del modelo del gesto “gracias” mediante distancia euclidiana para el movimiento en y del dedo índice de la mano derecha.....                                              | 103 |
| <b>Figura 62</b> Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento en el eje x de la mano derecha.....      | 107 |
| <b>Figura 63</b> Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento en el eje y de la mano derecha.....      | 107 |
| <b>Figura 64</b> Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo anular de la mano derecha. .... | 108 |
| <b>Figura 65</b> Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo índice de la mano derecha.....  | 108 |
| <b>Figura 66</b> Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo medio de la mano derecha.....   | 109 |
| <b>Figura 67</b> Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo meñique de la mano derecha..... | 109 |
| <b>Figura 68</b> Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo pulgar de la mano derecha. .... | 110 |
| <b>Figura 69</b> Comparación de la firma individual de prueba (test) con el modelo del gesto “gracias” obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo meñique de la mano derecha.....       | 111 |

|                                                                                                                                                                                                                         |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>Figura 70</b> Comparación de la firma individual de prueba (test) con el modelo del gesto “gracias” obtenidos mediante Dynamic Time Warping (DTW) para el movimiento en $y$ del dedo pulgar de la mano derecha. .... | 111 |
| <b>Figura 71</b> Comparación de 10 firmas test (videos de prueba) con el modelo del gesto “gracias” mediante el método DTW para el movimiento en $x$ del dedo índice de la mano derecha. ....                           | 112 |
| <b>Figura 72</b> Comparación de 10 firmas test (videos de prueba) con el modelo del gesto “gracias” mediante el método DTW para el movimiento en $y$ del dedo índice de la mano derecha. ....                           | 112 |
| <b>Figura 73</b> Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento en $x$ de la mano derecha – gesto gracias. ....                          | 119 |
| <b>Figura 74</b> Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento en $y$ de la mano derecha – gesto gracias. ....                          | 120 |
| <b>Figura 75</b> Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento del dedo anular de la mano derecha – gesto gracias. ....                 | 120 |
| <b>Figura 76</b> Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento del dedo índice la mano derecha – gesto gracias. ....                    | 121 |
| <b>Figura 77</b> Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento del dedo medio la mano derecha – gesto gracias. ....                     | 121 |
| <b>Figura 78</b> Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento del dedo meñique de la mano derecha – gesto gracias. ....                | 122 |

|                                                                                                                                                                                                           |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>Figura 79</b> Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento del dedo pulgar de la mano derecha – gesto gracias. ....   | 122 |
| <b>Figura 80</b> Comparación de la firma test y del modelo del gesto “gracias” mediante distancia euclidiana para el desplazamiento del dedo meñique de la mano derecha. ....                             | 123 |
| <b>Figura 81</b> Comparación de la firma test y del modelo del gesto “gracias” mediante distancia euclidiana para el movimiento en $x$ del dedo índice de la mano derecha. ....                           | 123 |
| <b>Figura 82</b> Comparación entre la firma de prueba (test) los modelos de cada gesto mediante Distancia Euclidiana, componente de movimiento en $x$ de la punta del dedo índice de la mano derecha..... | 124 |
| <b>Figura 83</b> Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento en $x$ de la mano derecha. ....                                     | 125 |
| <b>Figura 84</b> Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento en $y$ de la mano derecha.....                                      | 125 |
| <b>Figura 85</b> Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento del dedo anular de la mano derecha. ....                            | 126 |
| <b>Figura 86</b> Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento del dedo índice de la mano derecha. ....                            | 126 |
| <b>Figura 87</b> Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento del dedo medio de la mano derecha.....                              | 127 |
| <b>Figura 88</b> Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento del dedo meñique de la mano derecha.....                            | 127 |
| <b>Figura 89</b> Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento del dedo pulgar de la mano derecha. ....                            | 128 |

**Figura 90** Comparación de la firma individual de prueba (test) con el modelo del gesto

“gracias” obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo meñique de la mano derecha..... 128

**Figura 91** Comparación de la firma individual de prueba (test) con el modelo del gesto

“gracias” obtenidos mediante Dynamic Time Warping (DTW) para el movimiento en  $y$  del dedo pulgar de la mano derecha. .... 129

**Figura 92** Comparación entre la firma de prueba (test) con los modelos promedio de los gestos mediante DTW, componente de movimiento en  $x$  del dedo índice de la mano derecha..... 130

## Resumen

La Lengua de Señas Ecuatoriana (LSEC) constituye el principal medio de comunicación de la comunidad con discapacidad auditiva en el Ecuador; sin embargo, su conocimiento en la población general sigue siendo limitado, lo que genera barreras de inclusión en ámbitos educativos, sociales y laborales. En este contexto, el presente trabajo de integración curricular propone el desarrollo de un sistema de rastreo de puntos clave de las manos (*keypoint tracking*) aplicado a expresiones básicas de la LSEC, con el fin de generar representaciones temporales precisas que sirvan como base para futuras aplicaciones de clasificación automática y herramientas accesibles de enseñanza y aprendizaje. El proyecto emplea la técnica de un modelo de inteligencia artificial preentrenado, como MediaPipe, la cual permite detectar hasta 21 puntos anatómicos por mano y segmentar la región de interés en videos de señas. A partir de material gestual de uso libre como lo son videos, los cuales se procesan cuadro a cuadro para extraer y almacenar las coordenadas relevantes de puntos clave de las manos como series temporales. Estas trayectorias representan la dinámica espacial y temporal de gestos básicos de la LSEC. Los datos obtenidos se organizan en conjuntos de modelamiento y validación, que se evalúan utilizando métricas como Distancia Euclidiana y Deformación Temporal Dinámica (DTW), para medir la similitud entre diferentes ejecuciones de un mismo gesto. Este proceso permite caracterizar cada gesto de forma consistente y obtener las bases para modelos futuros de reconocimiento automático. El resultado del trabajo constituye un avance importante en la representación digital de la LSEC y un aporte pionero para la construcción de sistemas inteligentes orientados a la accesibilidad y la comunicación inclusiva en el Ecuador.

Palabras clave: rastreo de puntos clave, Lengua de Señas Ecuatoriana, MediaPipe, Distancia Euclidiana, DTW.

## Abstract

Ecuadorian Sign Language (LSEC) constitutes the primary means of communication for the deaf community in Ecuador; however, its knowledge among the general population remains limited, creating barriers to inclusion in educational, social, and occupational contexts. Within this framework, this curricular integration project proposes the development of a hand keypoint tracking system applied to basic LSEC expressions, with the aim of generating precise temporal representations that can serve as a foundation for future automatic classification applications and accessible teaching and learning tools. In this project, a pre-trained MediaPipe model is employed to detect up to 21 anatomical keypoints on each hand and to locate the region of interest within sign language videos. Each recording is processed sequentially, extracting the hand landmark coordinates from every frame and storing them as time-series data. The resulting trajectories describe how the hand moves over time, providing a structured representation of basic LSEC gestures. Subsequently, the collected data are organized into modeling and validation sets, which are evaluated using metrics such as Euclidean Distance and Dynamic Time Warping (DTW) in order to measure the similarity between different executions of the same gesture. This process enables a robust characterization of each gesture and establishes the foundation for future automatic recognition models. The outcomes of this study provide a structured digital description of Ecuadorian Sign Language gestures and offer a technical basis for future systems aimed at improving accessibility and communication support within the Ecuadorian context.

*Keywords:* keypoint tracking, Ecuadorian Sign Language, MediaPipe, Euclidean Distance, Dynamic Time Warping.

### **Abreviaciones**

LSEC: Lengua de Señas Ecuatoriana.

DTW: Dynamic Time Warping (Deformación Temporal Dinámica).

## Capítulo 1

### Introducción

#### Antecedentes

La comunicación gestual ha sido una forma fundamental de expresión humana desde tiempos ancestrales, utilizada para transmitir ideas, emociones y significados incluso antes del desarrollo del lenguaje oral. A lo largo de la historia, las manos han desempeñado un papel central en esta forma de comunicación, evidenciándose en manifestaciones artísticas, culturales y sociales. En este contexto, las lenguas de señas surgen como sistemas lingüísticos completos, de naturaleza visual-gestual, que permiten a las personas con discapacidad auditiva interactuar y participar activamente en la sociedad (Stokoe, 1960).

En el Ecuador, la Lengua de Señas Ecuatoriana (LSEC) es el principal medio de comunicación de la comunidad con discapacidad auditiva y cuenta con el reconocimiento oficial como parte del patrimonio lingüístico y cultural desde la Constitución de 2008 (Asamblea Nacional del Ecuador, 2008). A partir de esto, diferentes iniciativas institucionales han impulsado su fortalecimiento y difusión (Vicepresidencia de la República del Ecuador, 2013). El conocimiento y dominio de la LSEC por parte de la población oyente sigue siendo limitado, lo que genera barreras significativas de inclusión en ámbitos educativos, laborales y de acceso a servicios. Esta situación evidencia la necesidad de desarrollar herramientas tecnológicas que faciliten el aprendizaje, la interpretación y la difusión de la LSEC de manera accesible y efectiva.

En los últimos años, los avances en visión por computador e inteligencia artificial han permitido el desarrollo de sistemas capaces de analizar el movimiento humano a partir de imágenes y videos. Entre estas técnicas, el rastreo de puntos clave o *Keypoint Tracking* se ha consolidado como una alternativa eficiente para la representación del movimiento corporal, ya

que permite identificar y seguir automáticamente puntos anatómicos específicos, como articulaciones y extremidades, proporcionando una descripción más detallada y precisa que los métodos tradicionales basados en cajas delimitadoras (Bounding Box) (Cao et al., 2017). Esta capacidad resulta importante para el análisis de lenguas de señas, donde la posición, trayectoria y sincronización de las manos son determinantes para el significado de cada expresión.

Varios estudios internacionales han analizado el uso de *Keypoint Tracking* para la interpretación automática de lenguas de señas, especialmente en sistemas desarrollados como la *American Sign Language* (ASL) y la *British Sign Language* (BSL). Estos trabajos muestran que el seguimiento de puntos clave en manos y brazos permite modelar el movimiento como secuencias temporales, que luego pueden ser analizadas mediante métricas de similitud para identificar patrones gestuales (Alsharif et al., 2025; Koller et al., 2017).

El desarrollo de *frameworks* especializados han favorecido estas investigaciones. Herramientas como *MediaPipe* desarrollada por *Google*, permiten detectar hasta 21 puntos clave por mano, ofreciendo un alto nivel de precisión incluso en condiciones variables de iluminación y fondo (Lugaresi et al., 2019). De manera complementaria, modelos de detección de objetos como *YOLO* han evolucionado a versiones más eficientes y precisas, como *YOLO11*, facilitando la identificación de la región de interés del usuario dentro del video y optimizando el análisis posterior de los movimientos de las manos (Ultralytics, 2024).

A pesar de estos avances, la mayoría de los estudios se han enfocado en lenguas de señas de países angloparlantes, existiendo aún una limitada producción científica orientada específicamente a la LSEC. Los recursos digitales disponibles en el Ecuador, como diccionarios audiovisuales y canales educativos, constituyen un punto de partida importante,

pero no ofrecen representaciones estructuradas ni modelos computacionales que permitan analizar de forma sistemática la dinámica de los gestos (Consejo Nacional para la Igualdad de Discapacidades (CONADIS), s.f.; Franco-Segovia, 2022). Esta falta de trabajos especializados posiciona al rastreo de puntos clave aplicado a la LSEC como un campo de investigación emergente y necesario.

En este contexto, el presente trabajo se plantea como una respuesta a la necesidad de aportar bases técnicas y metodológicas orientadas a ayudar a la representación digital de expresiones básicas de la LSEC mediante el rastreo de puntos clave de las manos. El análisis de estas trayectorias, utilizando métricas como la Deformación Temporal Dinámica (DTW) y Distancia Euclidiana, permitirá establecer las bases para futuros sistemas de clasificación automática, enseñanza asistida y herramientas inclusivas, ayudando al desarrollo tecnológico y social del país, así como al fortalecimiento de la inclusión digital de la comunidad con discapacidad auditiva.

## **Motivación**

La comunicación es un elemento fundamental para la participación activa de las personas en la sociedad. En el caso de las personas con discapacidad auditiva, la Lengua de Señas Ecuatoriana (LSEC) constituye el principal medio de interacción social, educativa y cultural. Sin embargo, el limitado conocimiento de esta lengua por parte de la población oyente continúa generando dificultades de inclusión que dificultan el acceso equitativo a la educación, al empleo y a distintos servicios (Stokoe, 1960; Vicepresidencia de la República del Ecuador, 2013). Esta situación evidencia la necesidad de desarrollar herramientas tecnológicas que ayuden a facilitar la comprensión, el análisis y la enseñanza de la LSEC.

En los últimos años, los avances en Visión por Computador e Inteligencia Artificial han permitido el desarrollo de técnicas capaces de analizar de forma automática el movimiento humano a partir de imágenes y videos (Szeliski, 2022). Entre estas técnicas, el rastreo de puntos clave de las manos se ha consolidado como una alternativa eficaz para representar la dinámica gestual, ya que permite capturar con precisión la posición y trayectoria de las articulaciones a lo largo del tiempo (Cao et al., 2017). Este enfoque resulta especialmente relevante en el análisis de lenguas de señas, donde el significado de una expresión depende no solo de la posición de la mano, sino también de su movimiento y sincronización.

A pesar del crecimiento de investigaciones orientadas al reconocimiento automático de lenguas de señas, la mayoría de los estudios se han centrado en lenguas extranjeras, como la *American Sign Language* (ASL), dejando en segundo plano a lenguas locales como la LSEC (Alsharif et al., 2025). Esta situación ha limitado la disponibilidad de modelos, bases de datos y representaciones computacionales adaptadas al contexto lingüístico y cultural ecuatoriano, donde los recursos existentes se concentran principalmente en material audiovisual educativo y no en modelos estructurados para análisis automático (Consejo Nacional para la Igualdad de Discapacidades (CONADIS), s.f.). Por lo que, existe una necesidad evidente de generar estudios específicos que permitan caracterizar de manera sistemática las expresiones básicas de la LSEC a partir de herramientas tecnológicas actuales.

La motivación de este trabajo nace, de la necesidad de aportar al desarrollo de soluciones tecnológicas accesibles que permitan representar y analizar gestos básicos de la LSEC a partir del rastreo de puntos clave de las manos. El uso de modelos preentrenados y métodos computacionales eficientes facilita la implementación de sistemas con bajo costo y potencial de aplicación en contextos educativos y sociales.

El presente trabajo busca integrar el interés académico por el análisis del movimiento humano con un compromiso social orientado a la inclusión. La generación de representaciones temporales de los gestos de la LSEC no solo aporta al ámbito de la investigación en Visión por Computador, sino que también ayuda a dar un paso importante hacia el desarrollo de herramientas que faciliten la comunicación entre personas oyentes y no oyentes, apoyando a una sociedad más inclusiva y equitativa.

### **Justificación**

El presente proyecto de integración curricular busca aportar al desarrollo de una herramienta tecnológica que ayude a disminuir las barreras de comunicación entre personas oyentes y personas con discapacidad auditiva, mediante el análisis de expresiones básicas de la Lengua de Señas Ecuatoriana (LSEC). La limitada disponibilidad de soluciones tecnológicas enfocadas particularmente en esta lengua evidencia la necesidad de generar propuestas que permitan representar de manera estructurada la dinámica gestual de las manos, dentro del contexto lingüístico y cultural ecuatoriano (Stokoe, 1960; Consejo Nacional para la Igualdad de Discapacidades (CONADIS), s.f.).

Desde el punto de vista técnico, el rastreo de puntos clave de las manos permite capturar información espacial y temporal con mayor precisión en comparación con los métodos tradicionales basados en regiones de interés, ya que describe de forma explícita la estructura y el movimiento de las articulaciones a lo largo del tiempo (Cao et al., 2017). El uso de modelos de Visión por Computador basados en Inteligencia Artificial facilita la detección automática de puntos anatómicos relevantes, permitiendo obtener representaciones consistentes de las expresiones gestuales, incluso cuando existen variaciones individuales en su ejecución (Szeliski, 2022).

En el ámbito académico, este trabajo propone un procedimiento para la extracción, almacenamiento y análisis de trayectorias de puntos clave, utilizando métricas de similitud como la Deformación Temporal Dinámica (DTW) y Distancia Euclidiana. Esta técnica ha resultado ser apropiada para la comparación de secuencias temporales que presentan variaciones en duración y velocidad, aspectos propios de la ejecución natural de gestos en lenguas de señas (Berndt & Clifford, 1994; Keogh & Ratanamahatana, 2005). Los resultados obtenidos podrán ser usados como base para futuros desarrollos orientados al reconocimiento automático de señas y a la creación de recursos educativos accesibles.

En cuanto al impacto social, este proyecto pretende apoyar a la inclusión digital, mediante el desarrollo de herramientas que ayuden al aprendizaje y la difusión de la LSEC. Al facilitar el análisis de los gestos, se abre la posibilidad de mejorar los mecanismos de interacción entre personas oyentes y personas con discapacidad auditiva, contribuyendo a reducir las limitaciones en el acceso a la comunicación.

## **Objetivos**

### ***Objetivo General***

Realizar el rastreo de puntos clave de las manos en expresiones básicas de la lengua de señas ecuatoriana.

### ***Objetivos Específicos***

- Investigar y seleccionar una base de datos gestual adecuada para el entrenamiento y validación del sistema.
- Definir el conjunto mínimo de puntos clave en la mano relevantes para las expresiones a analizar.

- Utilizar un modelo de inteligencia artificial preentrenado para la detección y seguimiento automático de estos puntos clave en los frames de un video.
- Almacenar la información de los puntos clave extraídos de cada frame y construir un modelo que represente cada gesto.
- Realizar pruebas funcionales del sistema para evaluar su rendimiento y sus resultados obtenidos.

## **Alcance**

El proyecto se centra en la construcción de un modelo base para la representación de expresiones básicas de la Lengua de Señas Ecuatoriana (LSEC), mediante la extracción y el rastreo de puntos clave de las manos en secuencias de video. El análisis del movimiento gestual se realiza con técnicas de Visión por Computador, con el objetivo de capturar y modelar la información espacial y temporal que caracteriza cada seña básica.

Por ello, se realiza la selección de videos que muestran expresiones básicas de la LSEC, a partir de un material audiovisual de uso libre y de fuentes reconocidas. Estos videos serán procesados frame a frame para la detección y extracción de las coordenadas de los puntos clave que representan las articulaciones y segmentos principales de los dedos y la muñeca. La información obtenida se organizará y almacenará en forma de series temporales que describen la evolución del gesto a lo largo del tiempo.

A partir de estas series temporales, se construirá un modelo base representativo de las expresiones básicas de la LSEC, considerando la trayectoria y el comportamiento dinámico de los puntos clave de las manos. Para el análisis y la comparación de diferentes ejecuciones de una misma seña, se usarán métricas de similitud como la Deformación Temporal Dinámica

(DTW) y Distancia Euclidiana, lo que permitirá evaluar la consistencia y la repetibilidad de los gestos registrados.

El proyecto contempla la realización de pruebas funcionales orientadas a verificar la estabilidad del rastreo de puntos clave y la coherencia de las representaciones generadas frente a variaciones naturales en la ejecución de las señas, tales como cambios en la velocidad de ejecución o ligeras diferencias de postura. Estas pruebas se limitarán a un conjunto previamente definido de expresiones básicas.

Es importante indicar que este trabajo no incluye la implementación de un sistema final de clasificación automática ni la traducción de la LSEC a lenguaje oral o escrito. El alcance se limita a la creación de un modelo base fundamentado en la extracción y el análisis de puntos clave de las manos, que podrá servir como referencia para investigaciones y desarrollos futuros relacionados con el reconocimiento automático, la enseñanza asistida o el desarrollo de aplicaciones inclusivas.

## **Estructura del documento**

El presente documento describe en secuencia el desarrollo del proyecto de integración curricular, cuya organización se estructura en varios capítulos. En el Capítulo 1 se presentan la introducción dentro del cual se detalla la justificación y el alcance del proyecto, así como los objetivos que se planteó ejecutar. En el Capítulo 2 se presentan los conceptos y fundamentos teóricos que sustentan el trabajo y permiten contextualizar el enfoque adoptado. El Capítulo 3 presenta la metodología empleada y el proceso continuo para la construcción del modelo base a partir de la extracción de puntos clave de las manos. En el Capítulo 4 se muestran las pruebas realizadas, así como el análisis de los resultados obtenidos. Finalmente, en el Capítulo 5 se presentan las conclusiones del trabajo y posibles líneas de investigación futuras.

## **Capítulo 2**

### **Fundamentos Teóricos**

En este capítulo se desarrollan los fundamentos técnicos y conceptuales relacionados con la construcción de un modelo base para la representación de gestos básicos de la Lengua de Señas Ecuatoriana (LSEC), a partir del análisis de secuencias de video. Se describen las técnicas de Visión por Computador empleadas para la detección y el rastreo de puntos clave de la mano, haciendo énfasis en el uso de herramientas como MediaPipe para la extracción de coordenadas espaciales asociadas a las articulaciones de los dedos y la muñeca. Asimismo, se abordan los principios de la normalización de datos y la modelación de gestos mediante series temporales, necesarios para reducir la variabilidad entre diferentes ejecuciones de una misma seña. Finalmente, se presentan los fundamentos de las métricas de similitud utilizadas para la construcción y comparación de firmas gestuales, incluyendo la distancia euclidiana y la Deformación Temporal Dinámica (DTW), las cuales permiten evaluar la consistencia y representatividad del modelo base propuesto.

#### **Lengua de Señas Ecuatoriana (LSEC)**

La Lengua de Señas Ecuatoriana (LSEC) es el sistema de comunicación visual–gestual utilizado por la comunidad con discapacidad auditiva en el Ecuador, la misma que se basa en el uso coordinado de movimientos de las manos, expresiones corporales y referencias espaciales para transmitir significado, y no constituye una traducción directa del español, sino una lengua natural con estructura y reglas propias. Su reconocimiento y difusión resultan fundamentales para garantizar el acceso equitativo a la educación, la comunicación y la inclusión social de las personas sordas en el país (Stokoe, 1960) (Federación Nacional de Personas Sordas del Ecuador, s.f.).

En la LSEC, las manos cumplen un rol central en la construcción de las expresiones, ya que a través de su configuración, orientación y movimiento se codifica la mayor parte de la información semántica. La trayectoria y la dinámica temporal de los gestos manuales permiten diferenciar una seña de otra, incluso cuando la trayectoria de la mano es similar. Esta característica convierte al movimiento de las manos en un elemento clave para su análisis y representación computacional, especialmente en sistemas que modelan gestos a partir de coordenadas espaciales y su evolución en el tiempo.

La representación de los gestos manuales mediante puntos clave facilita el estudio sistemático de las expresiones básicas de la LSEC. Al capturar las posiciones de las articulaciones de la mano a lo largo del tiempo, es posible construir modelos base que describan el comportamiento gestual de una seña, sentando las bases para aplicaciones posteriores de análisis, comparación y apoyo al aprendizaje mediante técnicas de visión por computador (Lugaresi et al., 2019).

### **Detección y rastreo de puntos clave de la mano (Keypoints Hands).**

La detección de puntos clave es una técnica de Visión por Computador para localizar posiciones anatómicas específicas de un objeto dentro de una imagen o secuencia de video. Cuando se trabaja con la mano humana, estos puntos se asocian con articulaciones y segmentos importantes, como la muñeca y las falanges de los dedos. A diferencia de los métodos basados en regiones de interés o cajas delimitadoras (Bounding Box), la detección de puntos clave permite describir la estructura interna del objeto de manera precisa y detallada (Cao et al., 2017).

Este tipo de representación resulta útil para el análisis gestual, ya que los puntos clave aportan información directa sobre la forma y postura de la mano. A partir de esa información es

posible reconstruir la configuración del gesto y, posteriormente, analizar cómo cambia a lo largo del tiempo.

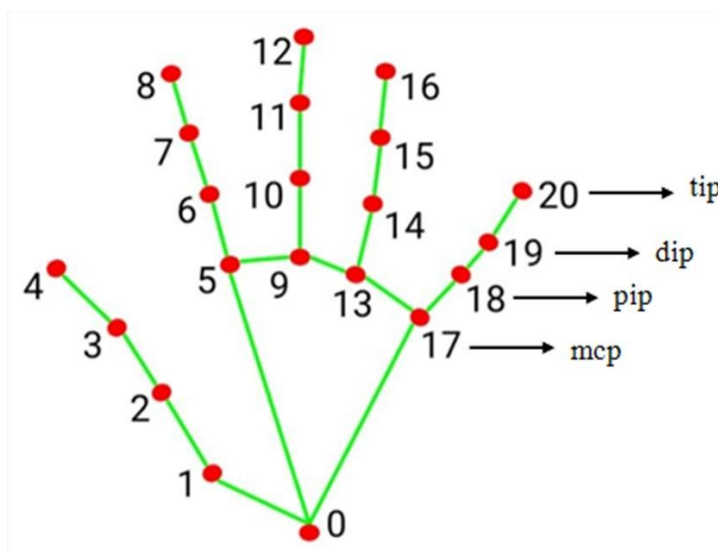
**Rastreo de puntos clave en secuencias de video.** El rastreo de puntos clave lleva la detección estática al dominio temporal, permitiendo seguir la evolución de cada punto a lo largo de los diferentes cuadros de un video (frames). Este proceso genera trayectorias que describen cómo cambian las posiciones de las articulaciones en el tiempo, lo que es importante para modelar movimientos gestuales (Szeliski, 2022).

Cuando se analiza el movimiento en gestos, el seguimiento de puntos clave permite describir cada ejecución como una secuencia de coordenadas organizadas en el tiempo, manteniendo la correspondencia de cada punto entre los distintos instantes. De esta forma, es posible comparar distintas ejecuciones de un mismo gesto y observar diferencias relacionadas con la velocidad, la amplitud o la sincronización del movimiento, variaciones que se presentan de manera natural en las lenguas de señas.

**Keypoints.** Los Keypoints o puntos clave son ubicaciones anatómicas o estructurales importantes de un objeto que se identifican en imágenes o fotogramas y se representan mediante coordenadas espaciales. En la mano, los puntos clave corresponden a la muñeca y a las articulaciones de cada dedo como se muestra en la Figura 1. Su detección permite describir la forma y posición de la mano de manera consistente y reproducible (Lugaresi et al., 2019) (Szeliski, 2022). Modelar un gesto mediante la secuencia temporal de estos puntos es una forma natural de representar la dinámica gestual sin depender de la información de píxeles crudos.

**Figura 1**

*Representación de los 21 puntos clave de una mano para el análisis gestual.*



*Nota.* Ubicación de puntos clave para representar la estructura de la mano en el modelo MediaPipe. Adaptado de (Ultralytics, 2024)

**Keypoints Tracking.** El *Keypoint Tracking* identifica y sigue puntos anatómicos específicos a lo largo de los frames o cuadros de video, obteniendo una representación característica de la estructura de la mano y su evolución temporal como se muestra en la Figura 2. De manera que combina la localización precisa con la descripción interna de la pose, lo que facilita la normalización, el análisis temporal y la comparación entre ejecuciones del mismo gesto (Cao et al., 2017).

Algunas de las ventajas que presentan son:

- Permite analizar individualmente la posición y movimiento de falanges y muñeca, esenciales en la semántica de las señas.

- Las coordenadas de los puntos clave son datos ligeros y fáciles de almacenar y procesar.
- Al trabajar con *landmarks*, disminuye la influencia del fondo o la silueta con respecto a la segmentación.
- Al organizar las coordenadas como series temporales, es posible aplicar métricas como la Distancia Euclidiana o la Deformación Temporal Dinámica (DTW) para comparar distintas ejecuciones de un mismo gesto (Berndt & Clifford, 1994; Lugaresi et al., 2019).

**Keypoints representación en coordenadas.** Cada punto clave se almacena típicamente como un par  $(x, y)$  (o tripleta  $(x, y, z)$  si se dispone de profundidad/estimación 3D) por fotograma. La secuencia de estos pares a través del tiempo conforma una serie temporal multivariada, definida como *STM*:

$$STM = \{(x_{1,t}, y_{1,t}, x_{2,t}, y_{2,t}, \dots, x_{n,t}, y_{n,t})\}_{t=1}^T$$

donde:

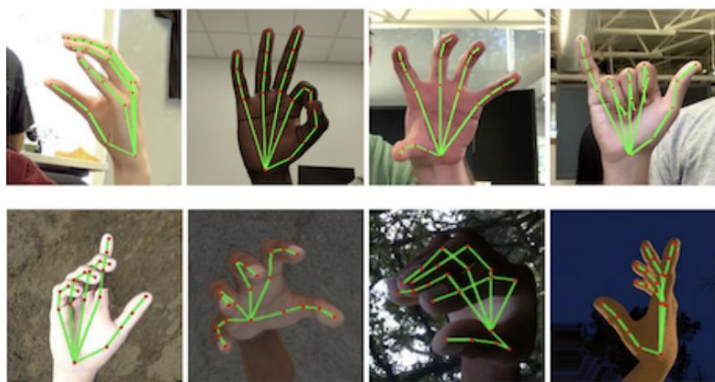
- $n$  es el número de landmarks (por ejemplo, 21 en muchos modelos de mano),
- $T$  la longitud en fotogramas los cuales componen el video completo a analizar.
- $t$  corresponde al índice temporal,
- $x_{i,t}$  y  $y_{i,t}$  indican las coordenadas del punto clave  $i$  en el instante  $t$ .

Antes de realizar el análisis, las secuencias pasan por una etapa de preprocesamiento que incluye filtrado e interpolación de frames faltantes, así como un ajuste de escala y espacio. Estas operaciones buscan reducir diferencias asociadas a la distancia respecto a la cámara o

al tamaño aparente de la mano, de modo que las ejecuciones puedan compararse bajo condiciones más estables (Szeliski, 2022; Keogh & Ratanamahatana, 2005).

## Figura 2

*Landmarks (21 puntos) en varios ejemplos de posiciones de la mano.*



*Nota.* Representación de los 21 puntos clave de la mano usados en estimación de pose.

Tomado de (Xataka, s.f.)

## MediaPipe.

MediaPipe es un marco de trabajo de código abierto desarrollado por Google para la creación de soluciones de Visión por Computador y aprendizaje automático orientadas al procesamiento de datos multimedia, como imágenes y secuencias de video. Este marco de trabajo proporciona un conjunto de herramientas, bibliotecas y pipelines preconstruidos que permiten integrar capacidades de análisis visual de forma eficiente, sin la necesidad de entrenar modelos desde cero (Google, 2023).

Dentro del ecosistema de MediaPipe se incluyen soluciones específicas para tareas de Visión por Computador, entre las cuales se encuentra la detección de puntos de referencia de la mano (*Hand Landmarker*). Estas soluciones ofrecen la funcionalidad principal necesaria para identificar estructuras anatómicas relevantes a partir de imágenes o video, permitiendo extraer

información estructurada de manera directa. La guía de soluciones de MediaPipe resalta que cada tarea integra modelos optimizados y procesos de inferencia que facilitan la detección y el rastreo de puntos clave en tiempo real, lo que resulta adecuado para aplicaciones de análisis gestual (Google, 2023).

### **Representación de gestos mediante series temporales.**

Se denomina serie temporal a un conjunto de valores que se registran de manera sucesiva en distintos instantes de tiempo. Cada observación está asociada a un momento específico, lo que permite estudiar cómo cambia una variable conforme transcurre el tiempo. Este tipo de representación resulta útil cuando se analizan fenómenos dinámicos y se busca identificar tendencias o patrones en su comportamiento (Shumway & Stoffer, 2017).

En el análisis de gestos, una serie temporal permite describir cómo cambian ciertas características como la posición de una articulación a lo largo de una secuencia de video para que el gesto deje de ser una imagen estática y pase a representarse como un proceso dinámico dependiente del tiempo.

**Serie temporales multivariadas.** Cuando varias variables se registran al mismo tiempo a lo largo de una secuencia temporal, se habla de una serie temporal multivariada. En el contexto de este trabajo, cada punto clave de la mano puede considerarse una variable independiente, de modo que un gesto completo queda descrito por el conjunto de todas estas variables evolucionando simultáneamente. Este enfoque permite observar cómo interactúan las distintas partes de la mano y analizar el movimiento como un fenómeno global, más allá del comportamiento aislado de cada punto (Shumway & Stoffer, 2017).

Por ejemplo, la serie temporal de un gesto puede representarse como  $STM$ :

$$STM = \{(x_{1,t}, y_{1,t}, x_{2,t}, y_{2,t}, \dots, x_{n,t}, y_{n,t})\}_{t=1}^T$$

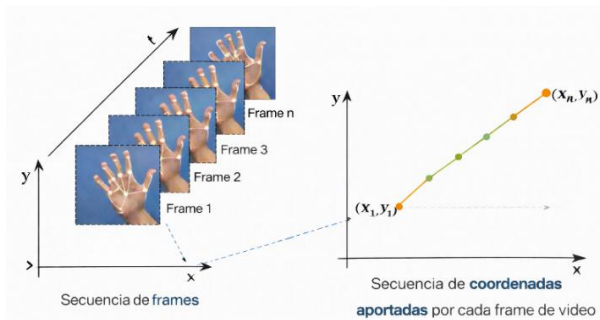
donde  $x_{i,t}$  y  $y_{i,t}$  corresponden a las coordenadas del punto clave  $i$  en el frame  $t$ , y  $n$  es el número de puntos clave por mano.

**Relación entre coordenadas y tiempo.** Para describir el movimiento como una serie temporal, las coordenadas espaciales de los puntos clave se extraen en cada fotograma y se organizan según el orden en que aparecen en el video. En cada instante se obtiene un vector que contiene las posiciones de todos los puntos, y al unir estos vectores de forma consecutiva se construye una secuencia que refleja la evolución del gesto.

De esta manera, un movimiento originalmente visual puede representarse como datos organizados en el tiempo, tal como se muestra en la Figura 3. Esta estructura facilita el uso de métodos de análisis temporal, por ejemplo, técnicas de alineamiento o métricas de similitud para comparar distintas ejecuciones.

**Figura 3**

*Normalización espacial de coordenadas de puntos clave en gestos manuales.*



*Nota.* Secuencia de frames de video que es transformada en una secuencia temporal de coordenadas, cada frame aporta un par ordenado  $(x_i, y_i)$  que genera una serie temporal continua en el tiempo. Adaptado de (Szeliski, 2022).

### **Normalización de coordenadas.**

La extracción de puntos clave en imágenes o video generan coordenadas que dependen de varios factores como el tamaño de la mano en la imagen o frame, la distancia entre la mano y la cámara, o la resolución de captura. Si estas variables no se ajustan previamente, pueden influir en la comparación entre gestos, ya que las coordenadas reflejan tanto la configuración real como esas variaciones externas que no forman parte del movimiento en sí (Szeliski, 2022).

Por ejemplo, algunos de estos errores significativos son:

- **Diferente tamaño de mano.** Dos personas con manos de diferente tamaño generarán rangos de coordenadas distintos, aunque realicen el mismo gesto.

- **Distancia a la cámara.** La distancia de la mano a la cámara afecta la escala de las coordenadas detectadas, manos más cercanas tendrán valores mayores que manos lejanas.

**Normalización espacial.** La normalización espacial transforma las coordenadas de los puntos clave a un sistema de referencia común para todos los gestos. Para ello se centran las coordenadas con respecto a un punto base como la muñeca, de manera que todos los puntos clave estén expresados en un espacio común como se muestra en la Figura 4 (Szeliski, 2022).

#### Figura 4

*Normalización espacial de coordenadas de puntos clave en gestos manuales.*



*Nota.* Todas las manos toman la base de coordenadas en el mismo punto  $x(0,0)$ , eliminando la ubicación en el espacio de las manos para tener referencia común antes del análisis. Adaptado de (Freepik, s.f.).

**Normalización de escala.** La normalización escalar consiste en transformar los valores para que estén en un rango uniforme, como [0,1] establecidos entre dos puntos de referencia de modo que todos los puntos clave estén expresados en este marco común. Las variables se ajustan para que contribuyan de forma comparable en modelos basados en distancias (Bishop, 2006).

Una normalización común es el escalamiento Min–Max, que transforma una coordenada  $X$  a una coordenada normalizada de  $X$  a un rango acotado, generalmente [0, 1] (Bishop, 2006). La transformación se define como:

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

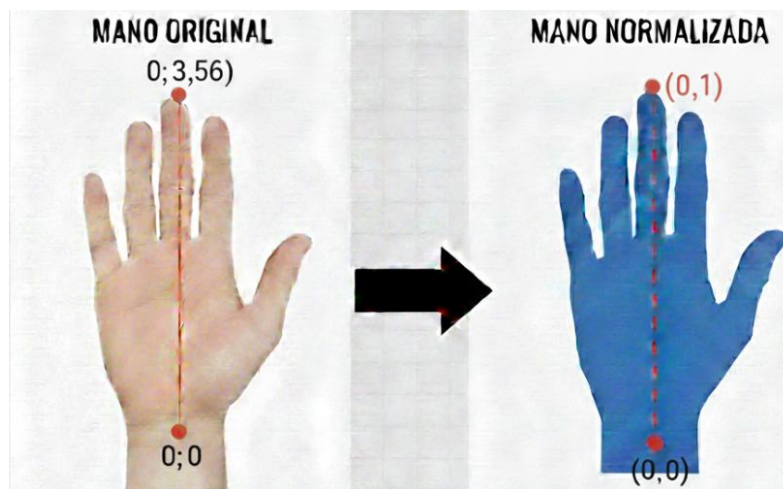
donde:

- $X$  es el valor original,
- $X_{\min}$  es el valor mínimo del conjunto de datos,
- $X_{\max}$  es el valor máximo del conjunto de datos,
- $X_{\text{norm}}$  es el valor normalizado.

Esto asegura que los valores normalizados estén en el intervalo [0, 1] como se muestra en la Figura 5.

**Figura 5**

*Proceso de normalización escalar de puntos clave de la mano.*



*Nota.* En la mano original, el punto distal se identifica en  $(0; 3.56)$ , el cual se reescala a  $(0; 1)$  en la mano normalizada. Adaptado de (Freepik, s.f.).

**Ventajas de la normalización de coordenadas.** La normalización de coordenadas da varios beneficios técnicos para su procesamiento como:

- Permite comparar señales o gestos de diferentes personas o capturados en distintas condiciones porque se elimina la dependencia de tamaño y posición.
- Las métricas que dependen de distancias (por ejemplo, distancia euclidiana o DTW) funcionan mejor cuando las características están en la misma escala, evitando que una sola dimensión domine la medida de similitud.

### **Firma gestual y modelo base de un gesto.**

Una firma gestual es una representación o un patrón de referencia de un gesto que resume sus características espacio-temporales esenciales. Normalmente es una serie temporal

multivariante derivada de las coordenadas de los puntos clave en este caso de la mano que busca representar la “forma típica” del gesto para su uso en comparación, verificación o enseñanza (Berndt & Clifford, 1994; Bishop, 2006). La firma actúa como un modelo que facilita comparar nuevas ejecuciones contra un patrón conocido.

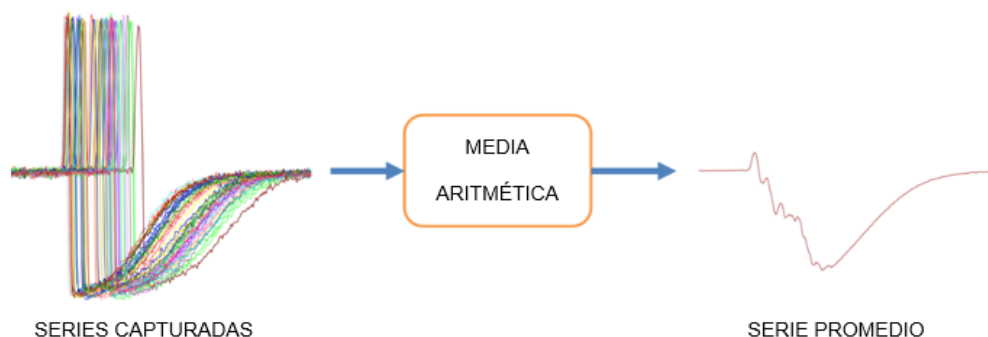
Para construir una firma gestual se suele partir de varias ejecuciones del mismo gesto realizadas por una o varias personas. El uso de múltiples ejecuciones del mismo gesto es esencial para construir una firma robusta. Estas ejecuciones permiten:

- Capturar la variabilidad natural (velocidad, amplitud, diferencias de trayectoria).
- Generar un modelo que describa el comportamiento promedio de las ejecuciones, menorando el efecto de diferencias individuales entre los gestos.
- Aumentar la estabilidad frente a errores de detección o frames perdidos a través de aplicar filtrado e interpolación antes realizar el promedio.

**Promedio aritmético (media punto a punto).** Consiste en alinear las series en longitud (por resampling o truncamiento) y calcular la media aritmética de cada componente o coordenada en cada tiempo como se muestra en la Figura 6. Este método es simple, pero sensible a desalineaciones temporales, es decir si las ejecuciones varían en velocidad o fase, la media puede cambiar o modificar características temporales importantes.

**Figura 6**

*Promedio de series temporales mediante Promedio Aritmético.*

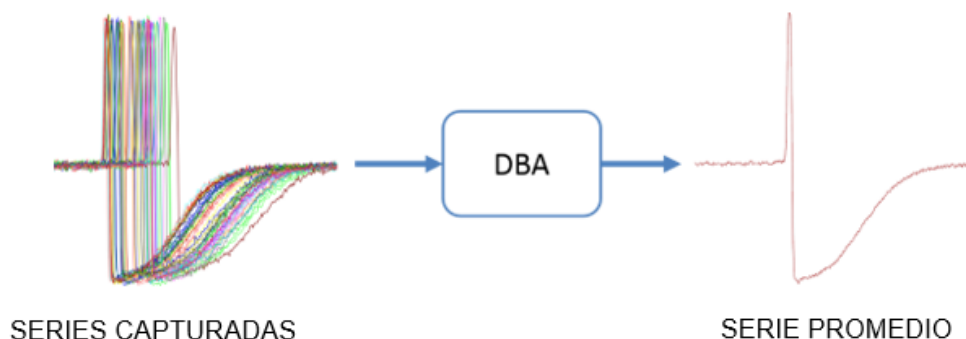


*Nota. Cuando las secuencias no están alineadas temporalmente, el promedio obtenido puede distorsionar la forma real de las trayectorias originales. Adaptado de (Petitjean, Dynamic Time Warping Barycenter Averaging, 2011).*

**Promedio DBA (DTW Barycenter Averaging).** DBA (*Dynamic Time Warping Barycenter Averaging*) es un método diseñado para promediar series temporales bajo la métrica DTW. DBA realiza un alineamiento temporal implícito entre series y calcula un modelo o prototipo que minimiza una medida de discrepancia consistente con DTW, conservando mejor la forma dinámica del gesto frente a variaciones en velocidad o fase entre ejecuciones como se muestra en la Figura 7 (Petitjean et al., 2011). Por tanto, DBA suele generar firmas más representativas para gestos con variabilidad temporal.

**Figura 7**

*Promedio de series temporales mediante DTW Barycenter Averaging (DBA).*



*Nota. Proceso de obtención de una serie promedio mediante el método DBA, donde múltiples series temporales son alineadas con DTW antes de calcular el modelo representativo.*

*Adaptado de (Petitjean, 2011).*

### **Distancia euclidiana.**

La distancia euclidiana es una medida clásica de similitud que cuantifica la distancia directa entre dos vectores en un espacio n-dimensional. Para dos vectores  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  y  $\mathbf{y} = (y_1, y_2, \dots, y_n)$ , la distancia euclidiana se define como:

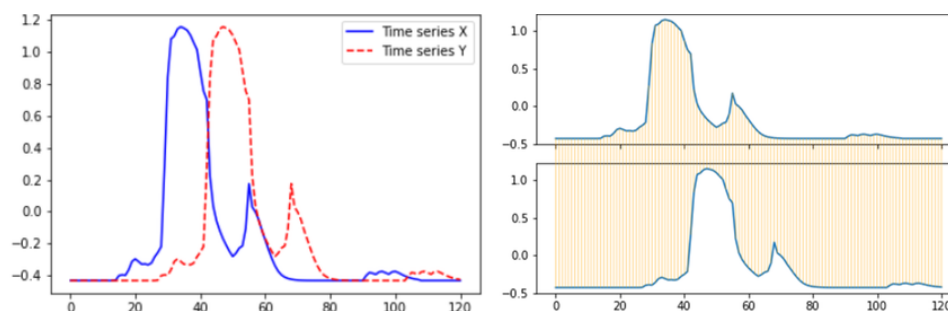
$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Esta métrica es ampliamente utilizada en análisis de patrones y aprendizaje automático debido a su simplicidad y eficiencia computacional (Bishop, 2006).

En la Figura 8 se muestra cómo se realiza la comparación de dos series temporales mediante la distancia euclidiana y como se comparan punto a punto.

**Figura 8**

*Comparación de series temporales mediante distancia euclidiana.*



*Nota. Se muestra la comparación punto a punto de dos series temporales, muestra cómo pequeñas desalineaciones temporales pueden generar distancias elevadas cuando se utiliza la métrica euclidiana. Tomado de (Keogh & Ratanamahatana, 2005).*

**Distancia euclidiana aplicada a vectores de gestos de la mano.** Cada gesto se describe mediante una serie temporal multivariada construida a partir de las coordenadas normalizadas de los puntos clave. Para poder emplear la distancia euclidiana, las secuencias deben tener la misma duración, ya que la comparación se realiza punto a punto.

Cuando ambas series tienen la misma longitud, cada gesto puede reorganizarse como un vector concatenado de coordenadas, y la Distancia Euclidiana se usa para medir la similitud entre una ejecución de prueba y el modelo promedio del gesto. Esta métrica es útil como referencia inicial en procesos de comparación y evaluación.

**Ventajas y limitaciones de la Distancia Euclidiana.** Las principales ventajas de la distancia euclidiana en la comparación de gestos son:

- Simplicidad de implementación, no requiere procesos adicionales de alineación temporal.
- Bajo costo computacional, lo que la hace adecuada para evaluaciones rápidas y sistemas en tiempo real.
- Interpretación directa, los valores menores indican mayor similitud entre gestos.
- No es robusta frente a desfasajes temporales, es decir cuando dos ejecuciones de un mismo gesto se realizan a distintas velocidades o presentan ligeros desplazamientos en el tiempo.
- Aunque la forma general del gesto sea similar, la comparación punto a punto puede arrojar distancias elevadas y conducir a conclusiones erróneas (Keogh & Ratanamahatana, 2005).

### **Deformación Temporal Dinámica (DTW).**

La Deformación Temporal Dinámica (*Dynamic Time Warping, DTW*) es una técnica de alineación temporal que para comparar secuencias que no necesariamente tienen la misma duración o que presentan diferencias en la velocidad de ejecución. En lugar de contrastarlas punto a punto, el algoritmo busca un alineamiento que reduzca el costo acumulado entre ambas, estableciendo correspondencias entre segmentos que pueden aparecer en momentos distintos dentro del tiempo (Sakoe & Chiba, 1978; Keogh & Ratanamahatana, 2005).

**Deformación Temporal Dinámica aplicada a gestos de la mano.** La Deformación Temporal Dinámica (DTW) compara dos series temporales construyendo una matriz de costos que evalúa la similitud entre todos los pares de instantes de ambas secuencias temporales. A partir de esta matriz, DTW encuentra un camino óptimo de alineación que minimiza el costo acumulado, respetando la monotonía y continuidad temporal del gesto como se muestra en la

Figura 9. Este camino permite alinear ejecuciones de un mismo gesto realizadas a diferentes velocidades, haciendo posible comparar patrones gestuales similares aun cuando existan desfasajes temporales significativos (Sakoe & Chiba, 1978; Keogh & Ratanamahatana, 2005).

**Figura 9**

*Matriz de costos acumulados y camino óptimo de alineamiento en DTW.*

| Y \ X | X |   |   |   |   |   |
|-------|---|---|---|---|---|---|
|       | 1 | 2 | 3 | 3 | 2 | 4 |
| 1     | 0 | 0 | 2 | 2 | 1 | 3 |
| 3     | 2 | 2 | 0 | 0 | 1 | 1 |
| 2     | 1 | 1 | 1 | 1 | 0 | 2 |
| 2     | 1 | 1 | 1 | 1 | 0 | 2 |
| 4     | 3 | 3 | 1 | 1 | 2 | 0 |
| 4     | 3 | 3 | 1 | 1 | 2 | 0 |

*Nota. Cada celda de la matriz muestra el costo mínimo para alinear los instantes de ambas series hasta esa posición específica. El recorrido resaltado es el camino del alineamiento temporal que mejor ajusta una serie con respecto a la otra. Tomado de (Sakoe & Chiba, 1978).*

Sean dos gestos representados por las series temporales  $X, Y$ :

$$X = (x_1, x_2, \dots, x_n), \quad Y = (y_1, y_2, \dots, y_m)$$

Donde cada elemento  $x_i$  y  $y_j$  corresponde a un vector de coordenadas normalizadas de un punto clave de la mano en el instante  $i$  o  $j$ .

Inicialmente se calcula una matriz de costos locales  $D$ , donde cada elemento se define como la distancia euclidiana entre los vectores de coordenadas:

$$D(i, j) = \|x_i - y_j\|_2$$

A partir de esta matriz  $D(i, j)$ , DTW construye una matriz de costos acumulados  $C$ , definida recursivamente como:

$$C(i, j) = D(i, j) + \min \begin{cases} C(i-1, j), \\ C(i, j-1), \\ C(i-1, j-1) \end{cases}$$

Con la condición inicial  $C(1,1) = D(1,1)$  (Sakoe & Chiba, 1978). El valor final de la distancia DTW entre los dos gestos se obtiene como:

$$DTW(X, Y) = C(n, m)$$

Un valor bajo de esta distancia muestra que existe un alineamiento temporal con bajo costo acumulado, lo que indica que ambos gestos presentan trayectorias similares, aunque hayan sido ejecutados a diferentes velocidades. Por el contrario, valores elevados indican que no es posible alinear adecuadamente las secuencias, lo que corresponde a gestos distintos (Sakoe & Chiba, 1978).

### **Normalización de la distancia DTW a una medida de similitud en el intervalo [0,1].**

Una forma sencilla y ampliamente utilizada de convertir la distancia DTW en una medida de similitud consiste en aplicar una transformación inversa acotada. En este caso, la similitud  $S$  entre dos gestos  $(X, Y)$  se define como:

$$S(\mathbf{X}, \mathbf{Y}) = \frac{1}{1 + \text{DTW}(\mathbf{X}, \mathbf{Y})}$$

Donde  $\text{DTW}(\mathbf{X}, \mathbf{Y})$  representa la distancia obtenida mediante *Dynamic Time Warping* entre las series temporales  $(\mathbf{X}, \mathbf{Y})$ . Con lo que garantiza:

- $S \in [0,1]$ ,
- $S \rightarrow 1$  la similitud tiende a 1 cuando la distancia DTW es menor, indican que los gestos comparados presentan trayectorias temporales muy similares y sugiere que probablemente corresponde a la misma señal.
- $S \rightarrow 0$  la similitud tiende a 0 cuando la distancia DTW es mayor, indican baja similitud y sugieren que los gestos comparados corresponden a señales distintas.

Transformaciones de este tipo han sido utilizadas para convertir métricas de distancia en puntuaciones de similitud comparables y estables numéricamente (Müller, 2007; Petitjean et al., 2011).

**Ventajas de DTW frente a la Distancia Euclidiana.** Las principales ventajas de DTW (DBA) frente a la Distancia Euclidiana son:

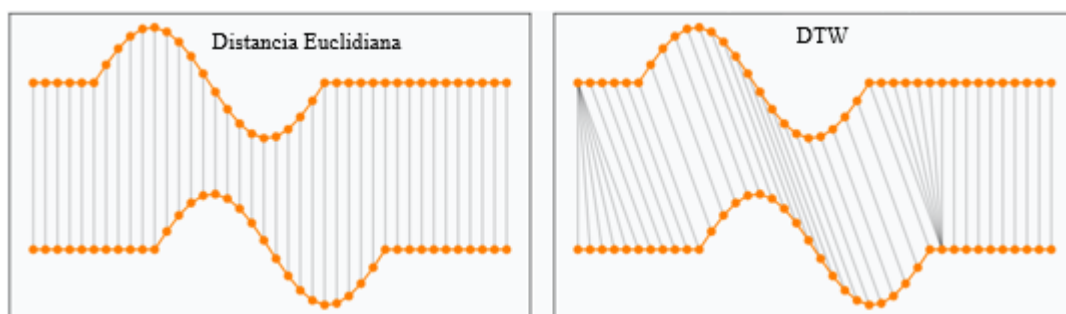
- Tolerancia a variaciones temporales, alineando gestos con diferentes velocidades.
- Este método maneja mejor los desfases temporales, que suelen aparecer cuando un mismo gesto se ejecuta a distintas velocidades.

- El promedio obtenido conserva mejor la dinámica del movimiento, lo que permite construir una firma que refleje con mayor fidelidad el comportamiento general del gesto.

En la Figura 10 se muestra la comparación entre el método DTW y Distancia Euclidiana.

**Figura 10**

*Comparación entre la distancia DTW y la Distancia Euclidiana.*



*Nota. Ejemplo del alineamiento flexible de dos series temporales con distinta duración mediante Distancia Euclidiana y DTW. Tomado de (Sakoe & Chiba, 1978).*

Estas características hacen de DTW el método principal de comparación en sistemas de reconocimiento y análisis de gestos (Sakoe & Chiba, 1978).

## Capítulo 3

### Implementación del Sistema

Luego de la presentación de los fundamentos teóricos, en el presente capítulo se describe el proceso de implementación del sistema propuesto. Este capítulo se centra en detallar la metodología empleada y las herramientas utilizadas para la construcción de un modelo base de gestos correspondientes a expresiones básicas de la Lengua de Señas Ecuatoriana (LSEC).

Inicialmente, se realiza la selección y preparación del material audiovisual utilizado como entrada del sistema, para después continuar con el proceso de detección y extracción de puntos clave de la mano a partir de secuencias de video utilizando técnicas de Visión por Computador. Así mismo, se describen las etapas de preprocesamiento aplicadas a las coordenadas obtenidas, incluyendo los procesos de normalización espacial y de escala, necesarios para disminuir la variabilidad entre distintas ejecuciones de un mismo gesto.

En esta sección se describe cómo se construye el modelo base del gesto a partir de varias ejecuciones. Para ello, se consideran dos formas de representación: un promedio punto a punto apoyado en la distancia euclidiana y un promedio alineado mediante Deformación Temporal Dinámica (DTW). Posteriormente, se analizan y contrastan ambos enfoques utilizando métricas de similitud, con el fin de examinar el comportamiento y la estabilidad de los modelos obtenidos.

### Metodología

La metodología propuesta se estructura en dos etapas principales, las cuales permiten organizar de manera clara el proceso de construcción y evaluación del modelo base de gestos de la Lengua de Señas Ecuatoriana (LSEC). Estas etapas corresponden a la etapa de

construcción del modelo base y la etapa de comparación junto con la validación, ambas estrechamente relacionadas y ejecutadas de forma secuencial.

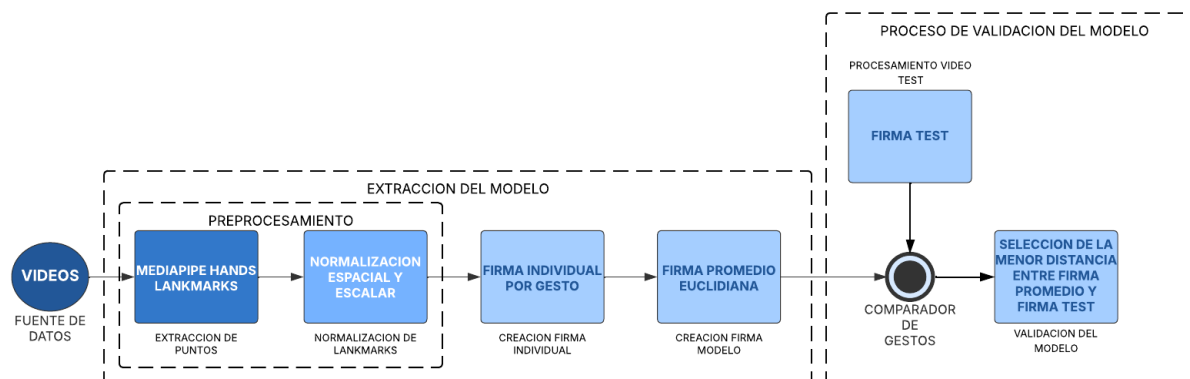
La primera etapa, correspondiente a la construcción del modelo base, se desarrolla en un entorno de procesamiento implementado en *Python* y tiene como finalidad generar una firma representativa de gestos básicos. Esta etapa incluye la selección el preprocesamiento del material audiovisual, la detección y extracción de los puntos clave de la mano a partir de secuencias de video, y el manejo de las coordenadas obtenidas mediante procesos de normalización espacial y de escala. Posteriormente, las ejecuciones normalizadas de un mismo gesto se agrupan para construir el modelo base o firma gestual, utilizando dos enfoques: un promedio punto a punto basado en Distancia Euclidiana y un promedio compatible con Deformación Temporal Dinámica (DTW).

La segunda etapa se realiza la comparación de los modelos promedio construidos. Para ello, se toma una nueva ejecución del gesto y se procesa bajo el mismo esquema de extracción y normalización de puntos clave. Posteriormente, esta ejecución se contrasta con el modelo base utilizando métricas de similitud basadas en distancia euclidiana y DTW. El resultado de esta comparación se expresa en un valor dentro del intervalo  $[0,1]$ , donde valores más altos reflejan mayor correspondencia entre la muestra evaluada y el modelo de referencia.

En Figura 11, se presenta el diagrama de bloques general del sistema usando el Método Euclidiano, mientras que en la Figura 12 se presenta el diagrama de bloques general del sistema usando el método DTW (DBA). En dichos diagramas se muestra el flujo del proceso, desde el video hasta la obtención de la medida de similitud, así como la relación entre la etapa de extracción del modelo base y la etapa de validación del modelo.

Figura 11

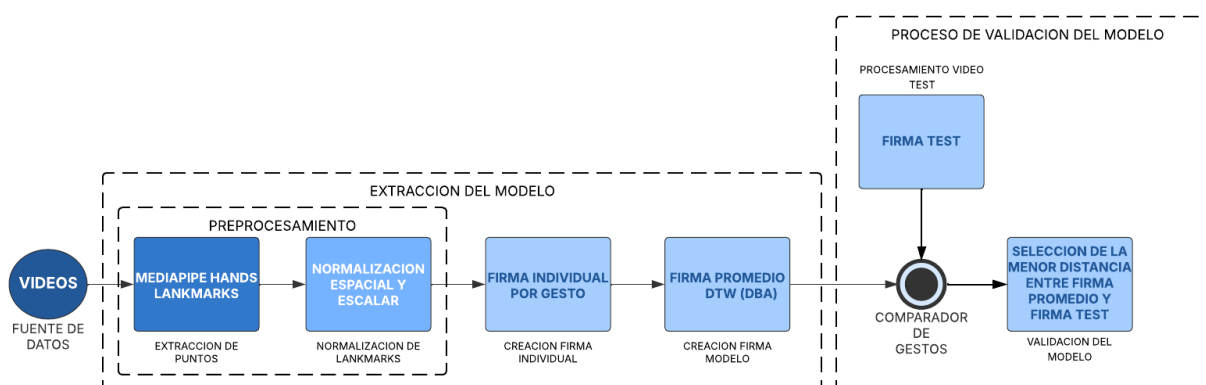
Diagrama general para el Método de Distancia Euclidiana.



*Nota.* Diagrama de bloques del procesamiento mediante el método Euclidiano. Elaboración propia.

Figura 12

Diagrama general para el Método DTW.



*Nota.* Diagrama de bloques del procesamiento mediante el método DTW (DBA). Elaboración propia.

## **Extracción del modelo.**

### ***Fuente de datos.***

Para este estudio se recopilamos videos de acceso libre que contienen expresiones básicas de la Lengua de Señas Ecuatoriana (LSEC). El material seleccionado permitió contar con registros adecuados para extraer y analizar el desplazamiento y movimiento de las manos a partir de secuencias visuales.

Los videos de uso libre usados corresponden a 4 personas con discapacidad auditiva entrenados en LSEC, cada una de las cuales ejecutó diez repeticiones de expresiones básicas de la LSEC. Las expresiones seleccionadas como señales de prueba son los gestos “feliz”, “gracias”, “hola” y “por favor”, ya que son las consideradas como básicas por parte de profesionales de LSEC. Esta cantidad de muestras de video permitió contar con múltiples ejecuciones de un mismo gesto, capturando variaciones naturales en la forma y velocidad de ejecución propias de cada participante.

Cada video corresponde a secuencias controladas teniendo en cuenta que siempre se visualicen ambas manos, por lo que cada gesto inicia y finaliza en una posición similar de la mano, previamente definida como posición de referencia. Esta condición facilita la identificación de los límites temporales del gesto y contribuye a reducir la variabilidad no deseada en las representaciones temporales generadas a partir de las coordenadas de los puntos clave.

Los videos obtenidos son la entrada principal de datos del sistema y fueron organizados por tipo de gesto y por participante, permitiendo su utilización tanto en la construcción del modelo base de cada gesto como en la etapa de validación.

En Figura 13, se presentan las fotografías de las cuatro personas participantes ubicadas en la posición ideal de inicio del gesto, las cuales muestran la postura de referencia de inicio y fin empleada durante la etapa de adquisición de datos, la base de datos se encuentra disponibles en: <https://drive.google.com/drive/folders/1ZdHJoC7KswLJn5Nrdsn5TUIUbwUI963I>

### Figura 13

*Participantes en la posición ideal de inicio del gesto.*



*Nota. Videos de uso libre para la etapa de adquisición de datos.*

### ***Detección y extracción de puntos clave de la mano.***

La detección y extracción de los puntos clave de la mano constituye la etapa principal en la implementación del sistema, ya que permite transformar las secuencias de video en datos numéricos estructurados que describen el movimiento gestual. En esta etapa, cada video es procesado cuadro a cuadro para identificar ambas manos del participante y extraer las coordenadas espaciales asociadas a sus articulaciones.

Para este propósito se utilizó MediaPipe, una herramienta desarrollada por Google y empleada durante la ejecución del sistema para la detección y el rastreo de los puntos clave de

la mano en tiempo real a partir de imágenes y secuencias de video (Google, 2023). Esta herramienta permite obtener de forma automática la localización de las articulaciones de la mano sin necesidad de marcadores físicos ni dispositivos adicionales, lo que resulta adecuado para aplicaciones basadas exclusivamente en información visual.

El proceso de detección y extracción se aplica de forma independiente a cada uno de los videos de uso libre utilizados, garantizando que todas las ejecuciones del gesto sean tratadas bajo las mismas condiciones de procesamiento.

**Extracción de puntos clave de la mano.** Para cada frame de video, se realiza la detección y el rastreo de los puntos clave de las manos mediante un modelo de Visión por Computador especializado. La duración de cada video está en un rango de 3 a 4 segundos, extrayendo 1 frame cada 1.5 microsegundo aproximadamente obteniendo un total aproximado de 170 frames por video de igual forma se realiza un filtrado para que únicamente se consideren válidos aquellos frames en los que se detectan simultáneamente ambas manos, derecha e izquierda. Los frames en los que no se logra identificar las dos manos son descartados, ya que no contienen información completa del gesto por lo que se toma el siguiente frame que contenga la información de ambas manos.

A continuación, en la Figura 14, se presenta una imagen ilustrativa del proceso de extracción de los puntos clave de ambas manos aplicado a la persona 1, en la cual se visualiza la distribución de las articulaciones detectadas para un frame representativo del video.

**Figura 14**

*Detección de los 21 puntos claves de cada mano de un frame de video.*



Una vez detectadas las manos, se extraen 21 puntos clave por cada mano, los cuales representan la muñeca y las articulaciones de los dedos. Cada punto clave se describe mediante un conjunto de tres coordenadas espaciales  $(x, y, z)$  como se muestra en la Figura 15, donde las coordenadas:

- $x$  y  $y$  indican la posición relativa del punto en el plano de la imagen,
- $z$  representa la profundidad relativa respecto a la muñeca.

**Figura 15**

*Eje de coordenadas inicialmente en un frame de video.*



Durante el proceso de extracción se aplican criterios de filtrado con el objetivo de garantizar la calidad de las secuencias obtenidas. En particular, se eliminan frames repetidos en los que el desplazamiento de los puntos clave entre instantes consecutivos es inferior a un umbral de 0.5 milisegundos (ms), evitando incluir información redundante. Asimismo, se descartan los frames iniciales en los que no se detecta movimiento significativo de la mano, de modo que la secuencia almacenada corresponda únicamente a la ejecución efectiva del gesto, por lo que en promedio cada video tiene alrededor de 130 frames o cuadros validos con lo que se garantiza la captura de la naturaleza del gesto completa.

**Almacenamiento de las coordenadas.** Las coordenadas  $(x, y, z)$  obtenidas para cada punto clave son almacenadas de forma estructurada al finalizar el procesamiento de cada

video. Para cada ejecución del gesto, las coordenadas correspondientes a todos los frames válidos se organizan como una serie temporal multivariada, donde cada instante temporal contiene la información espacial completa de ambas manos.

Para cada frame representado en un instante de tiempo  $t$ , se extraen los puntos clave de ambas manos:

- 21 puntos por mano
- 2 manos
- 3 coordenadas por punto  $(x, y, z)$

El vector de características asociado al frame  $t$  se define como:

$$\mathbf{f}_t = [x_{t,1}^{(R)}, y_{t,1}^{(R)}, z_{t,1}^{(R)}, \dots, x_{t,21}^{(R)}, y_{t,21}^{(R)}, z_{t,21}^{(R)}, x_{t,1}^{(L)}, y_{t,1}^{(L)}, z_{t,1}^{(L)}, \dots, x_{t,21}^{(L)}, y_{t,21}^{(L)}, z_{t,21}^{(L)}]$$

donde:

- $R$  y  $L$  describen la mano derecha e izquierda, respectivamente,
- $\mathbf{f}_t \in \mathbb{R}^{126}$ .

La ejecución completa del gesto se representa como una serie temporal multivariada, definida de forma matricial como:

$$\mathbf{G} \in \mathbb{R}^{N \times 126}$$

Donde:

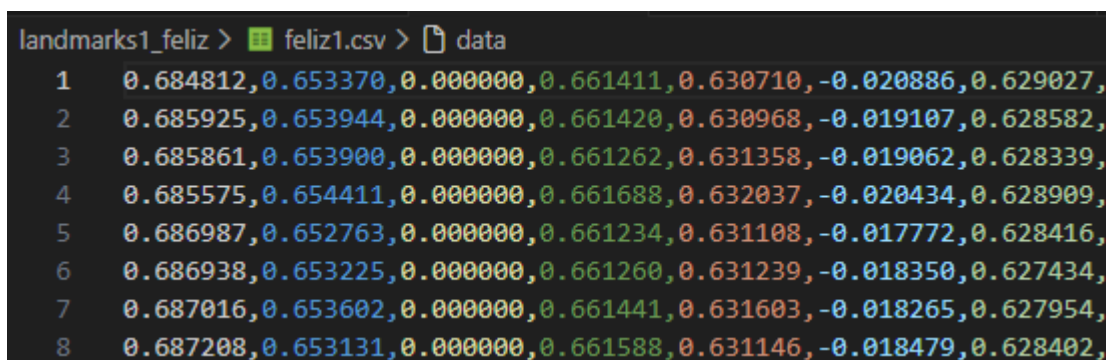
- $N$  es el número de frames válidos del video, cada fila corresponde a un instante temporal.

- El valor 126 corresponde al número total de columnas, donde cada una representa una coordenada específica asociada a un punto clave.

En la Figura 16, se muestra un ejemplo de cómo se encuentran almacenados cada serie temporal.

### Figura 16

*Almacenamiento de las series temporales 126 datos por frame.*



```
landmarks1_feliz > feliz1.csv > data
1 0.684812,0.653370,0.000000,0.661411,0.630710,-0.020886,0.629027,
2 0.685925,0.653944,0.000000,0.661420,0.630968,-0.019107,0.628582,
3 0.685861,0.653900,0.000000,0.661262,0.631358,-0.019062,0.628339,
4 0.685575,0.654411,0.000000,0.661688,0.632037,-0.020434,0.628909,
5 0.686987,0.652763,0.000000,0.661234,0.631108,-0.017772,0.628416,
6 0.686938,0.653225,0.000000,0.661260,0.631239,-0.018350,0.627434,
7 0.687016,0.653602,0.000000,0.661441,0.631603,-0.018265,0.627954,
8 0.687208,0.653131,0.000000,0.661588,0.631146,-0.018479,0.628402,
```

Cada serie temporal se guarda en formato .npy que permite su posterior análisis y reutilización. De esta manera, cada video genera un conjunto de datos individual del gesto, compuesto por una secuencia ordenada de vectores de coordenadas, que constituye la entrada directa para la siguiente etapa de preprocesamiento.

### **Normalización de coordenadas.**

Las coordenadas obtenidas en la etapa detección y extracción de puntos clave presentan variaciones asociadas a la posición absoluta de la mano dentro de la imagen y a la escala del gesto. Estas variaciones pueden afectar la comparación entre distintas ejecuciones

de una misma seña. Por este motivo, se aplica un proceso de normalización espacial y de escala a las coordenadas extraídas.

En este caso, la normalización se realiza de manera independiente para cada mano y para cada cuadro del video, considerando únicamente las coordenadas bidimensionales  $(x, y)$ . La componente de profundidad  $z$  no se incluye, ya que el análisis posterior se centra en la dinámica planar del gesto debido a que los videos son bidimensionales.

**Normalización espacial.** La normalización espacial nos ayuda a eliminar la dependencia de la posición absoluta de la mano dentro del encuadre. Para ello, se utilizó la muñeca como punto de referencia, dado que constituye una articulación estable y común a todas las configuraciones de la mano.

Sea una mano representada por el conjunto de puntos clave en un frame  $\mathbf{f}_t$ :

$$\mathbf{f}_t = \{\mathbf{p}_{t,1}, \mathbf{p}_{t,2}, \dots, \mathbf{p}_{t,21}\}, \quad \mathbf{p}_{t,i} = (x_{t,i}, y_{t,i})$$

donde  $t$  denota el instante temporal. La normalización espacial se realiza restando las coordenadas de la muñeca  $\mathbf{p}_{t,0}$  a todos los puntos de la mano:

$$\mathbf{p}'_{t,i} = \mathbf{p}_{t,i} - \mathbf{p}_{t,0}, \quad i = 0, \dots, 20$$

De esta manera, la muñeca se traslada al origen del sistema de coordenadas, y la configuración de la mano queda expresada de forma relativa a este punto de referencia, esta

normalización se aplica para cada una de las manos de los frames de cada video que representa un gesto.

**Figura 17**

*Normalización espacial de la mano*



*Nota. Proceso de normalización espacial para compensar la ubicación en el espacio de las manos en un instante de tiempo  $t$ . Adaptado de (Freepik, s.f.).*

Por ejemplo, como se observa en la Figura 17, para obtener la coordenada normalizada  $\mathbf{p}'_{1,12}$  en el espacio del punto 12 de la mano (punta del dedo medio)  $\mathbf{p}_{1,12}$  en un frame en el instante  $t = 1$ , se obtiene de la siguiente, asumiendo que  $\mathbf{p}_{1,12} = (4.65, 10.47)$ ,  $\mathbf{p}_{1,0} = (4.55, 6.15)$  :

$$\mathbf{p}'_{1,12} = \mathbf{p}_{1,12} - \mathbf{p}_{1,0}, \quad i = 12$$

$$\mathbf{p}'_{1,12} = (4.65, 10.47) - (4.55, 6.15)$$

$$\mathbf{p}'_{1,12} = (0.10, 4.32)$$

**Normalización de escala.** Una vez aplicada la normalización espacial, realizamos la normalización de escala, aquí eliminamos las diferencias en el tamaño aparente de las manos entre distintas ejecuciones. Para ello, se utiliza como referencia la distancia entre los puntos normalizados espacialmente de la muñeca  $\mathbf{p}'_{t,0}$  y la punta del dedo medio  $\mathbf{p}'_{t,12}$ , debido a su estabilidad y relevancia anatómica.

**Figura 18**

*Normalización escalar de la mano.*



*Nota. Proceso de normalización de escala para compensar la diferencia de tamaño de las manos en un instante de tiempo  $t$ . Adaptado de (Freepik, s.f.).*

La escala de la mano en el instante  $t$  se define como:

$$s_t = \|\mathbf{p}'_{t,12}\|_2$$

donde:

- $\mathbf{p}'_{t,12}$  corresponde al punto clave normalizado espacialmente de la punta del dedo medio.
- $\|\cdot\|_2$  denota la norma euclidiana.

Las coordenadas normalizadas espacialmente se escalan de la siguiente manera:

$$\mathbf{p}''_{t,i} = \frac{\mathbf{p}'_{t,i}}{s_t}, i = 1, \dots, 21$$

Por ejemplo, como se observa en la Figura 18, para el punto 12 de la mano, donde  $s_t = 4.32$  que es la distancia desde la muñeca (punto 0) a la punta del dedo medio (punto 12) tomando el ejemplo anterior de la normalización espacial, entonces:

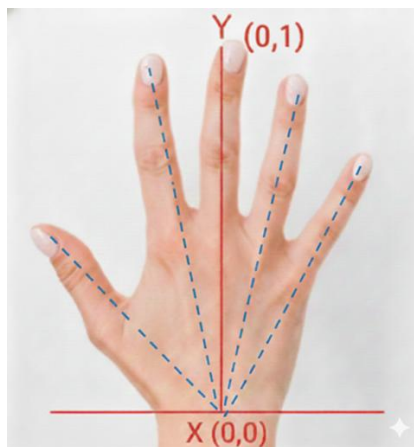
$$\mathbf{p}''_{1,12} = \left( \frac{0.10}{4.32}, \frac{4.32}{4.32} \right)$$

$$\mathbf{p}''_{1,12} = (0.0231, 1)$$

Este procedimiento garantiza que todas las ejecuciones del gesto queden representadas en una escala comparable, teniendo el punto clave de la muñeca como el valor 0 y el punto clave de la punta del dedo medio como el valor de 1 como se muestra en la Figura 19, por lo que todos los frames quedan comparables independientemente de la distancia de la mano a la cámara o del tamaño relativo de la mano de cada persona.

**Figura 19**

*Normalización escalar de la mano*



*Nota. Ejemplo de escalamiento para compensar las diferencias en la distancia de la cámara y en el tamaño de cada mano. Adaptado de (Freepik, s.f.).*

**Normalización por cuadro y por mano.** El proceso de normalización descrito se aplica de manera independiente a cada cuadro o frame del video y para cada mano, derecha e izquierda. Para cada frame válido se obtienen dos conjuntos de puntos normalizados, correspondientes a ambas manos, los cuales se concatenan para formar el vector normalizado del instante temporal.

Como resultado, cada frame queda representado por un vector de características de dimensión 84, correspondiente a 21 puntos  $\times$  2 manos  $\times$  2 coordenadas  $(x, y)$ . La secuencia completa del gesto se representa entonces como una serie temporal normalizada:

$$\mathbf{G}_{\text{norm}} \in \mathbb{R}^{N \times 84}$$

Donde  $N$  es el número frames válidos del video.

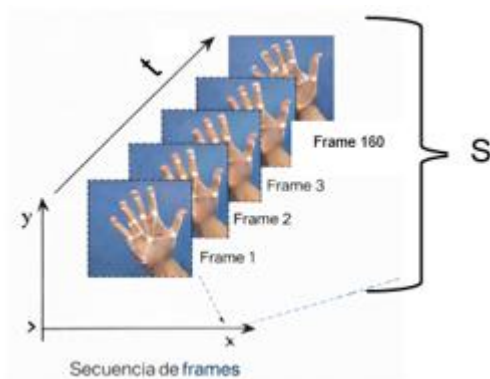
### ***Creación de firma individual.***

Una vez obtenidas las secuencias de coordenadas normalizadas espacial y escalarmente, se procede a la construcción de una firma representativa del gesto basada en Distancia Euclidiana, esta firma individual es la salida de datos obtenido al procesar un video de un gesto completo.

Debido a que las ejecuciones de un mismo gesto pueden presentar duraciones distintas, se aplica un proceso de normalización temporal para llevar todas las secuencias a una longitud común. Cada video de un gesto completo se reescala a un número fijo de  $N=100$  frames, ya que cada uno de los videos tiene duración y longitud " $S$ " diferente, por lo que se hace un muestreo igualmente espaciado ya que gracias a pruebas realizadas verificamos que con 100 frames se mantiene la naturaleza del gesto, con lo que se sigue manteniendo la estructura completa del gesto.

**Figura 20**

*Proceso de normalización temporal.*



**Nota.** Secuencia de frames de longitud  $S$  de un video que es transformado en una secuencia temporal de longitud 100 frames. Adaptado de (Szeliski, 2022).

La normalización temporal se realiza usando un re-muestreando de cada conjunto de frames de diferente longitud  $S$  a 100 frames como se muestra en la Figura 20. Para ello, la señal normalizada se definió como  $\tilde{s}(k)$ :

$$\tilde{s}(k) = s\left(\frac{k}{99}(T - 1)\right), \quad k = 0, 1, \dots, 99$$

donde  $T$  corresponde al número de frames originales de un video de un gesto completo y  $k$  es el índice de normalización que va de 0 a 99.

Este procedimiento preserva la evolución relativa de la señal en el tiempo y permite la comparación directa entre ejecuciones de distinta duración para que sean comparables en un mismo marco temporal normalizado a 100 frames.

Por ejemplo, para obtener las coordenadas normalizadas de los puntos clave del frame  $k = 33$  del conjunto de 100 frames normalizados de una secuencia de video de longitud de 160 frames se realiza de la siguiente manera:

$$\tilde{s}(33) = s\left(\frac{33}{99} \cdot (160 - 1)\right)$$

$$\tilde{s}(33) = s\left(\frac{33}{99} \cdot 159\right)$$

$$\tilde{s}(33) = s\left(\frac{1}{3} \cdot 159\right)$$

$$\tilde{s}(33) = s(53)$$

Por lo que se toma los valores de las coordenadas normalizadas del frame 53 de la secuencia original  $S$  y se asocia al frame 33 dentro del conjunto normalizado de 100 frames.

De igual forma en otro ejemplo para obtener las coordenadas normalizadas para el frame  $k = 25$  del conjunto normalizado de 100 frames, a partir de una señal original  $S$  de un video de 160 frames, se procede de la siguiente manera:

$$\tilde{s}(25) = s\left(\frac{25}{99} \cdot 159\right)$$

$$\tilde{s}(25) = s(0.2525 \cdot 159)$$

$$\tilde{s}(25) = s(40.15)$$

Dado que el resultado no es un número entero, se realiza una interpolación lineal entre las coordenadas normalizadas de los frames 40 y 41 del conjunto original  $S$ . Por lo tanto, los valores de las coordenadas normalizadas correspondientes al frame 25 del conjunto normalizado se obtiene interpolando proporcionalmente entre los frames originales 40 y 41, ya que entre dos frames consecutivos, el cambio es lo suficientemente pequeño como para aproximarlos de manera lineal.

**Desplazamiento euclidiano.** El análisis del movimiento de la mano se realiza a partir de las variaciones espaciales de los puntos clave detectados en frames consecutivos. Dado que una mano al moverse a cualquier dirección desplaza todos sus puntos en conjunto lo que describe el movimiento de un gesto, su movimiento se descompondrá en un desplazamiento global en  $(x, y)$ , asociado a la traslación de la mano como conjunto, y en movimientos locales, asociados principalmente a la dinámica individual de cada uno de los dedos.

Por lo que, a partir de las coordenadas normalizadas, se calcula el desplazamiento euclidiano entre 2 frames consecutivos como una métrica para medir el movimiento de la mano, es decir cuánto se movió la mano entre 2 frames consecutivos.

- **Desplazamiento global en el eje  $x$ .** Para cada eje espacial, este desplazamiento se calcula a partir de las diferencias entre las coordenadas de los puntos clave en dos frames consecutivos. En particular, para el eje  $x$ , se consideran las coordenadas horizontales de los 21 puntos de la mano en los instantes  $t$  y  $t + 1$ , y se calcula la magnitud del desplazamiento conjunto mediante la norma euclidiana:

$$D_x(t) = \sqrt{\sum_{i=1}^{21} (x_{t+1,i} - x_{t,i})^2}$$

Por ejemplo, para 3 puntos las coordenadas horizontales entre los instantes  $t$  y  $t + 1$  son:

$$\begin{aligned} x(t, 1) &= 0.20, & x(t + 1, 1) &= 0.25 \\ x(t, 2) &= -0.10, & x(t + 1, 2) &= -0.05 \\ x(t, 3) &= 0.40, & x(t + 1, 3) &= 0.35 \end{aligned}$$

Entonces, sustituyendo directamente en la ecuación:

$$D_x(t) = \sqrt{(0.25 - 0.20)^2 + (-0.05 - (-0.10))^2 + (0.35 - 0.40)^2}$$

$$D_x(t) = \sqrt{(0.05)^2 + (0.05)^2 + (-0.05)^2}$$

$$D_x(t) = \sqrt{0.0025 + 0.0025 + 0.0025}$$

$$D_x(t) = \sqrt{0.0075}$$

$$D_x(t) = 0.0866$$

El valor obtenido  $D_x(t)$  representa la magnitud del desplazamiento horizontal en de la mano entre los instantes  $t$  y  $t + 1$ . De esa misma forma se realiza considerando los 21 puntos clave de la mano, proporcionando una medida global del movimiento en el eje  $x$ .

- **Desplazamiento global en los ejes  $y$ .** De la misma forma que se realizó el desplazamiento global en el eje  $x$  se realiza en desplazamiento global en el eje  $y$ , se obtiene a partir de las coordenadas verticales, la cual resume el movimiento global de la mano entre cuadros consecutivos. El uso de todos los puntos clave garantiza que la magnitud obtenida represente el comportamiento conjunto de la mano, de referencia se puede usar el ejemplo anterior que se realizó con las coordenadas en  $x$  y aplicarlo con las coordenadas en  $y$ :

$$D_y(t) = \sqrt{\sum_{i=1}^{21} (y_{t+1,i} - y_{t,i})^2}$$

- **Desplazamiento global de cada dedo.** Así como el desplazamiento por cada eje se analiza el desplazamiento de cada dedo de manera independiente, con el fin de capturar la dinámica local del gesto. Cada dedo se encuentra representado por un conjunto de cuatro puntos clave correspondientes a sus articulaciones. Para un dedo específico, se calculan las diferencias entre las coordenadas  $x$  e  $y$  de sus puntos clave en dos frames consecutivos, y dichas variaciones se agrupan para obtener una única magnitud de desplazamiento mediante la norma euclidiana combinada:

$$D_{\text{dedo}}(t) = \sqrt{\sum_{j=1}^4 ((x_{t+1,j} - x_{t,j})^2 + (y_{t+1,j} - y_{t,j})^2)}$$

Por ejemplo, si se analiza el dedo índice de la mano derecha, suponiendo que está compuesto por cuatro coordenadas de sus puntos claves correspondientes a sus articulaciones entre los instantes  $t$  y  $t + 1$ :

$$\begin{aligned} (x(t, 1), y(t, 1)) &= (0.30, 0.50); & (x(t + 1, 1), y(t + 1, 1)) &= (0.35, 0.55) \\ (x(t, 2), y(t, 2)) &= (0.40, 0.60); & (x(t + 1, 2), y(t + 1, 2)) &= (0.42, 0.63) \\ (x(t, 3), y(t, 3)) &= (0.45, 0.65); & (x(t + 1, 3), y(t + 1, 3)) &= (0.48, 0.68) \\ (x(t, 4), y(t, 4)) &= (0.50, 0.70); & (x(t + 1, 4), y(t + 1, 4)) &= (0.53, 0.74) \end{aligned}$$

Entonces se tiene:

$$D_{\text{índice } D}(t) = \sqrt{(0.35 - 0.30)^2 + (0.55 - 0.50)^2 + \dots + (0.53 - 0.50)^2 + (0.74 - 0.70)^2}$$

$$D_{\text{índice } D}(t) = \sqrt{(0.05)^2 + (0.05)^2 + (0.02)^2 + (0.03)^2 + \dots + (0.03)^2 + (0.04)^2}$$

$$D_{\text{índice } D}(t) = \sqrt{0.0106}$$

$$D_{\text{índice } D}(t) = 0.1030$$

El valor  $D_{\text{índice } D}(t)$  representa el movimiento global del dedo índice como una unidad, considerando simultáneamente todas sus articulaciones (4 puntos clave). De esta manera, se obtiene una firma individual independiente para cada dedo (pulgar, índice, medio, anular y meñique), la cual describe su comportamiento dinámico a lo largo del gesto y nos ayuda a distinguir entre

diferentes gestos según su la medida del desplazamiento, de la misma forma se realiza para la mano izquierda y para cada video.

- **Movimiento espacial en  $x$ .** El movimiento espacial en el eje  $x$  describe la trayectoria horizontal de la punta de cada dedo de la mano en el tiempo con respecto al eje de coordenadas normalizado con centro (0,0) en la muñeca. Esta representación permite modelar la configuración del movimiento del gesto como se muestra en la Figura 28.
- **Movimiento espacial en  $y$ .** El movimiento espacial en el eje  $y$  describe la trayectoria vertical de la punta de cada dedo en el tiempo con respecto al eje de coordenadas normalizado con centro (0,0) en la muñeca como muestra la Figura 29. Esta representación es importante en gestos que implican desplazamientos ascendentes o descendentes predominantes.

**Construcción de las firmas individuales.** Para cada ejecución del gesto, se construye una firma individual a partir de un conjunto de componentes temporales que describen el desplazamiento global y el movimiento espacial de la punta de cada dedo en los ejes ( $x$ ,  $y$ ) de ambas manos. Estas componentes son:

- Desplazamiento global en el eje  $x$  para mano derecha e izquierda por separado.
- Desplazamiento global en el eje  $y$  para mano derecha e izquierda por separado.
- Desplazamiento combinado de los puntos clave correspondientes a cada dedo (pulgar, índice, medio, anular y meñique), calculado de forma independiente para cada mano y para cada dedo.

- Movimiento espacial en  $x$  de las puntas de cada dedo (pulgar, índice, medio, anular y meñique) para mano derecha e izquierda y para cada dedo por separado.
- Movimiento espacial en  $y$  de las puntas de cada dedo (pulgar, índice, medio, anular y meñique) para mano derecha e izquierda y para cada dedo por separado.

Obteniendo así 17 componentes  $x$  2 manos. Cada una de estas señales es normalizada temporalmente a  $N = 100$  frames y posteriormente agrupada para formar una matriz:

$$\mathbf{F}_i \in \mathbb{R}^{100 \times 34}$$

donde  $\mathbf{F}_i$  representa la firma individual de la ejecución  $i$  y 34 corresponde al número total de componentes consideradas distribuidas en 17 componentes por mano, se consideran los 34 componentes ya que describen de forma robusta el comportamiento del gesto, los primeros 7 componentes nos indican cuanto se movió la mano a lo largo del gesto mientras que las 10 componentes restantes nos indican el movimiento relativo de las puntas de los dedos respecto al eje normalizado de coordenadas, separando sus coordenadas para tener un análisis del movimiento relativo en  $x$  y en  $y$ .

Los diecisiete componentes por cada mano permiten capturar tanto el desplazamiento global como el movimiento específico de los dedos en ambos ejes  $(x, y)$ . Considerando simultáneamente la mano derecha y la mano izquierda, la firma completa del gesto está compuesta por 34 componentes, organizadas como una señal multivariada que describe la evolución del gesto en el tiempo normalizado.

La matriz final de la firma del gesto queda expresada de la siguiente manera:

$$\mathbf{F}_{gesto} = \begin{bmatrix} m_{1,1} & m_{1,2} & m_{1,3} & \dots & m_{1,34} \\ m_{2,1} & m_{2,2} & m_{2,3} & \dots & m_{2,34} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{100,1} & m_{100,2} & m_{100,3} & \dots & m_{100,34} \end{bmatrix}$$

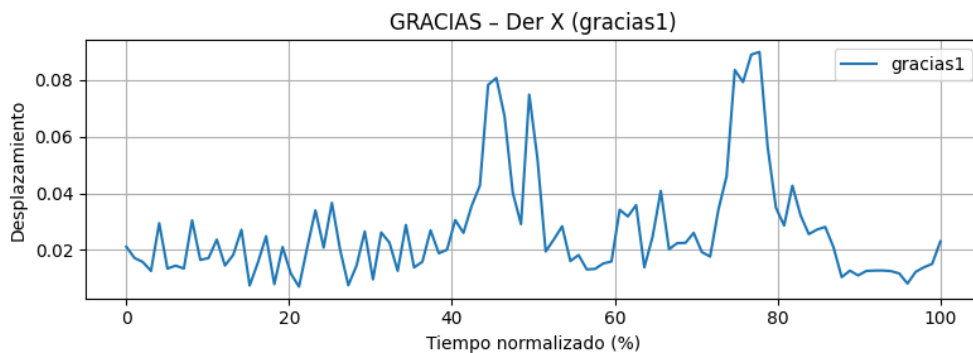
Donde:

- $\mathbf{F}_{gesto}$  es la matriz final de la firma del gesto para un video.
- Cada fila representa un instante temporal normalizado [1, 100].
- Cada columna representa una componente específica (34 componentes).

De la Figura 21 a la Figura 27 se muestran las 7 primeras firmas individuales para las 7 componentes de la mano derecha, mientras que en la Figura 28 se muestra en movimiento en el eje  $x$ , así como en el eje  $y$  que se muestra en la Figura 29, para la punta del dedo índice, de igual forma se procesó para cada una de las puntas de los dedos de ambas manos y para la mano izquierda, esto se realiza para cada uno de los 10 videos de cada gesto de cada persona.

## Figura 21

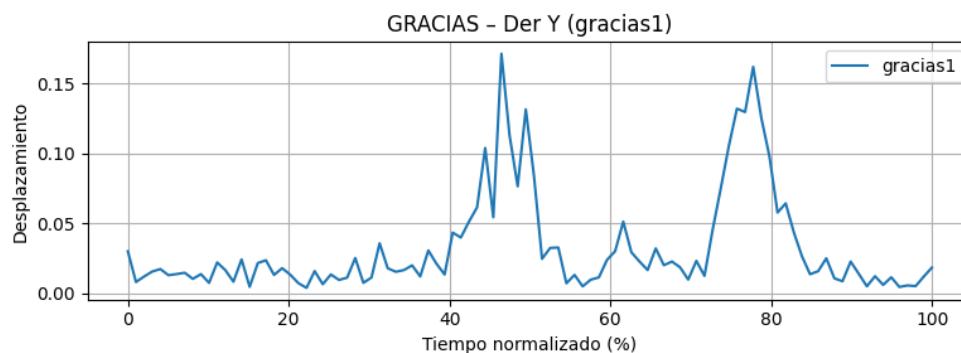
*Firma individual del desplazamiento global en el eje  $x$  de todos los puntos de la mano derecha – gesto gracias de la persona 1.*



## Figura 22

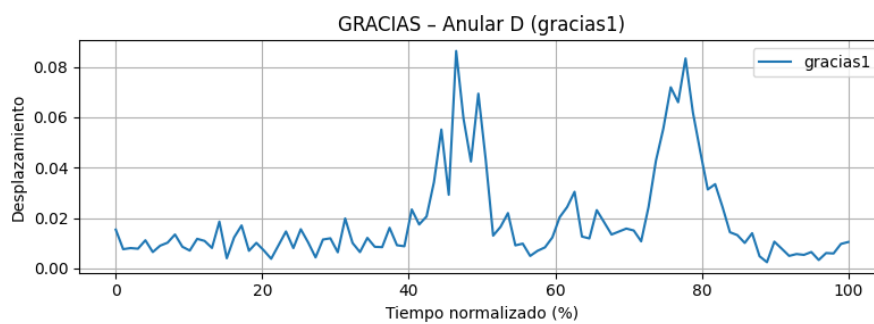
*Firma individual del desplazamiento global en el eje y de todos los puntos de la mano derecha*

*– gesto gracias de la persona 1.*



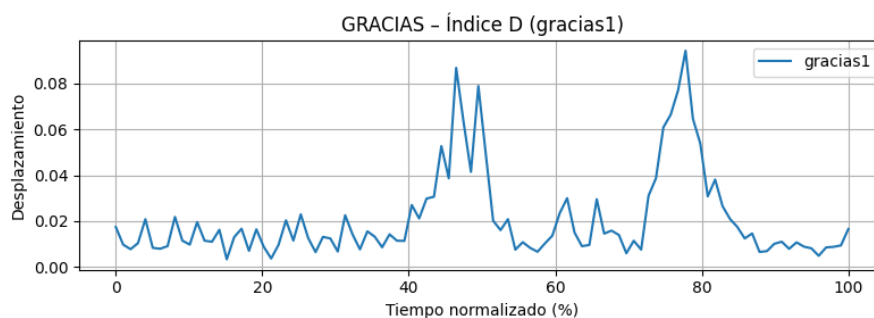
## Figura 23

*Firma individual del desplazamiento del dedo anular derecho – gesto gracias de la persona 1.*



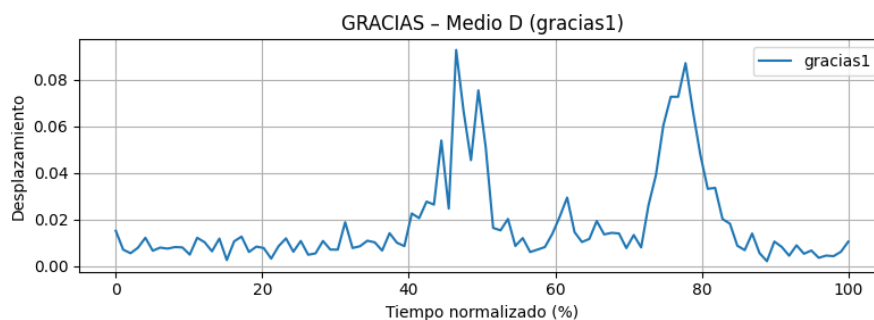
## Figura 24

*Firma individual del desplazamiento del dedo índice derecho – gesto gracias de la persona 1.*

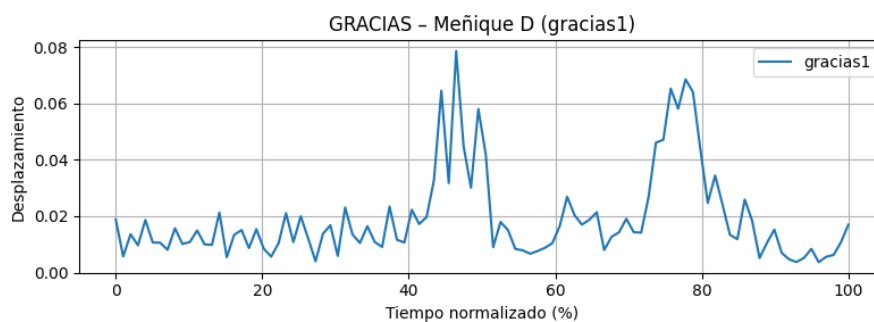


**Figura 25**

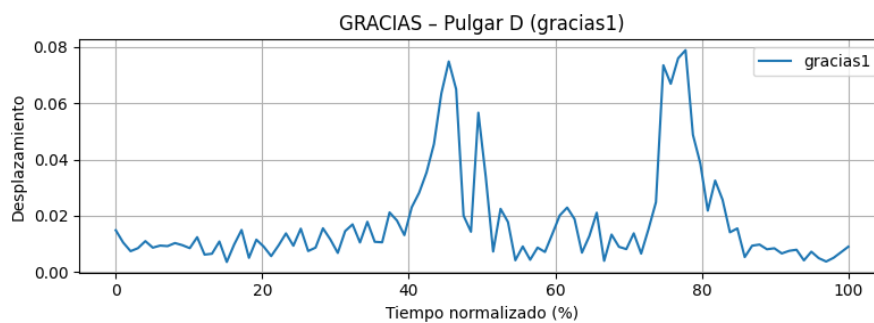
*Firma individual del desplazamiento del dedo medio derecho – gesto gracias de la persona 1.*

**Figura 26**

*Firma individual del desplazamiento del dedo meñique derecho – gesto gracias de la persona 1.*

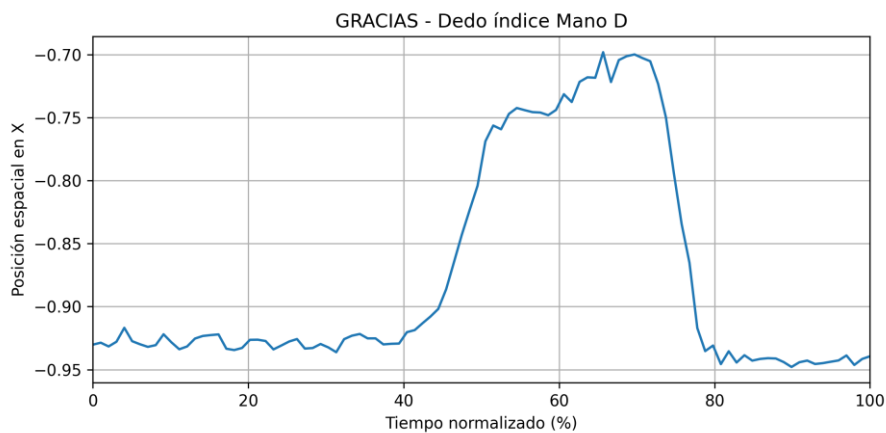
**Figura 27**

*Firma individual del desplazamiento del dedo pulgar derecho – gesto gracias de la persona 1.*

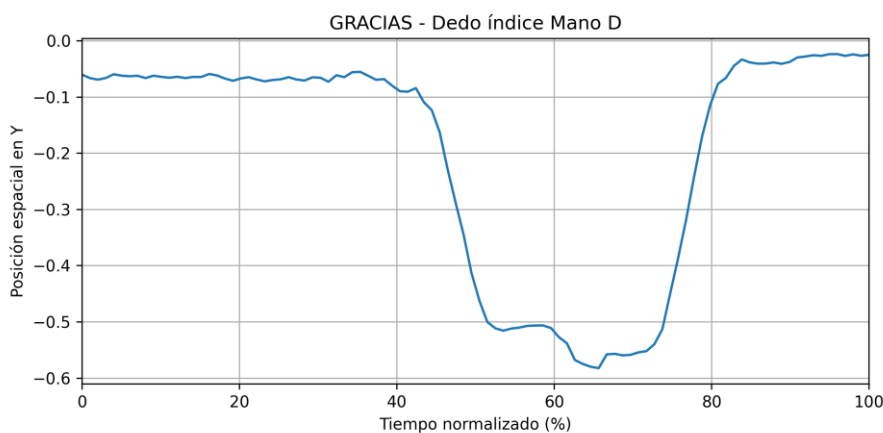


**Figura 28**

*Firma individual del movimiento en x del dedo índice derecho – gesto gracias de la persona 1.*

**Figura 29**

*Firma individual del movimiento en y del dedo índice derecho – gesto gracias de la persona 1.*



### ***Creación Modelo promedio euclidiana.***

La modelo promedio del gesto se obtiene a partir del conjunto de firmas individuales de cada gesto que realizó cada una de las personas en video. Cada firma individual corresponde a una ejecución del gesto y se encuentra representada como una matriz de dimensión  $100 \times 34$ , donde las filas corresponden a cada uno de los frame y las columnas corresponden a las 34 distintos

componentes que describen el desplazamiento y el movimiento de los puntos clave de las manos.

Las firmas individuales se organizan en una estructura tridimensional de tamaño  $K \times 100 \times 34$ , donde  $K$  representa el número de ejecuciones disponibles para cada gesto. Con esta organización, el modelo promedio se obtiene calculando el promedio correspondiente en cada punto de la secuencia a lo largo de todas las ejecuciones consideradas.

El cálculo se realiza utilizando la función *mean* de la librería *NumPy*, aplicada sobre el eje correspondiente a las ejecuciones. De esta forma, para cada frame y para cada componente de desplazamiento y movimiento, se obtiene el valor promedio entre todas las firmas individuales. El resultado es una matriz de dimensión  $100 \times 34$ , que constituye el modelo promedio del gesto como se muestra a continuación.

La matriz final del modelo del gesto queda expresada de la siguiente manera:

$$\mathbf{M}_{gesto} = \begin{bmatrix} m_{1,1} & m_{1,2} & m_{1,3} & \dots & m_{1,34} \\ m_{2,1} & m_{2,2} & m_{2,3} & \dots & m_{2,34} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{100,1} & m_{100,2} & m_{100,3} & \dots & m_{100,34} \end{bmatrix}$$

Donde:

- $\mathbf{M}_{gesto}$  es la matriz final del modelo del gesto.
- Cada fila representa un instante temporal normalizado  $[1, 100]$ .
- Cada columna representa una componente específica (34 componentes).

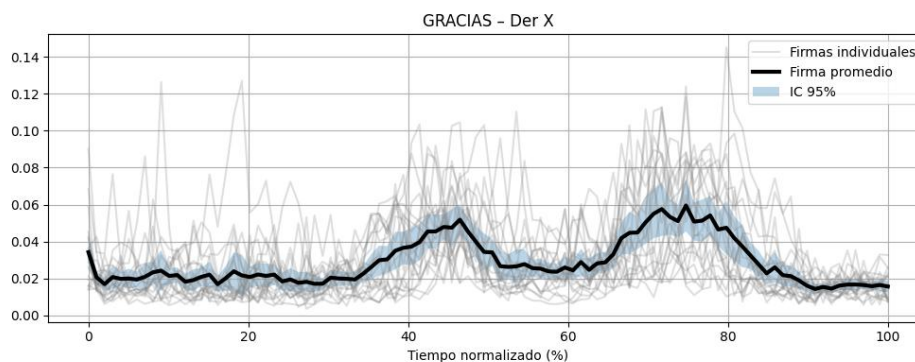
Adicionalmente, para cuantificar la variabilidad entre las distintas ejecuciones, se aborda un intervalo de confianza del 95 % para el modelo promedio de cada componente. Este nivel de confianza fue seleccionado debido a que constituye el valor estándar recomendado en estadística aplicada e ingeniería, al ofrecer un equilibrio adecuado entre confiabilidad y amplitud del intervalo. Diversos textos clásicos de estadística para ingeniería, como *Applied Statistics and Probability for Engineers* de Montgomery y Runger, así como las directrices del *Engineering Statistics Handbook* del *National Institute of Standards and Technology (NIST)*, establecen que el intervalo de confianza del 95 % es el más comúnmente utilizado en aplicaciones científicas y de ingeniería, permitiendo una caracterización representativa de la dispersión de los datos sin sobreestimar la variabilidad. Este proceso se realiza también de manera punto a punto y permite obtener una banda de variación alrededor del valor medio, lo que nos da consistencia del gesto entre ejecuciones.

La matriz de datos  $\mathbf{M}_{gesto}$  formada se utilizan como modelo base del gesto y sirve como referencia para los procesos de validación mostrados posteriormente. Desde la Figura 30 hasta la Figura 36 se observa los modelos promedio de los componentes de desplazamiento obtenidos de la mano derecha del gesto “gracias” a partir de 10 ejecuciones realizadas por cada una de las 3 diferentes personas, de igual forma se realiza para la mano izquierda y el resto de los gestos.

Así mismo en la Figura 37 se observa el modelo promedio del movimiento en  $x$  obtenido de la punta del dedo índice de la mano derecha mientras que en la Figura 38 se observa el modelo promedio del movimiento en  $y$  obtenido de la punta del dedo índice de la mano derecha, el proceso se repite para cada punta del dedo de la mano, para la mano derecha e izquierda, así mismo se aplica para el modelamiento del resto de gestos.

**Figura 30**

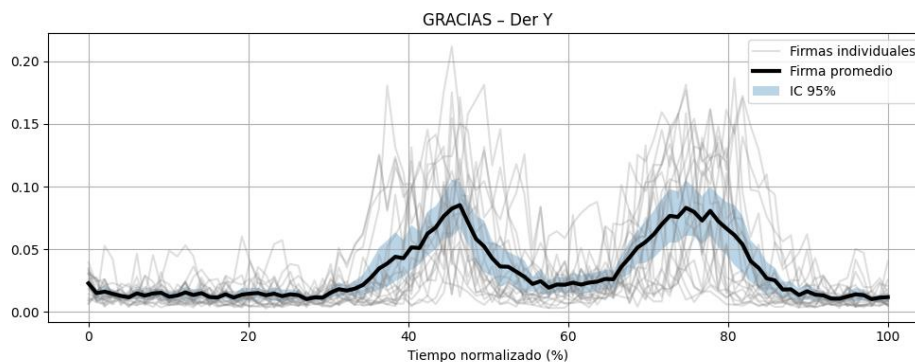
*Modelo promedio del desplazamiento global en el eje x los 21 puntos de la mano derecha con IC del 95 % – gesto gracias.*



*Nota. Las señales en gris muestran las firmas individuales del desplazamiento global de los 21 puntos de la mano en el eje x de cada video usado para sacar el modelo promedio.*

**Figura 31**

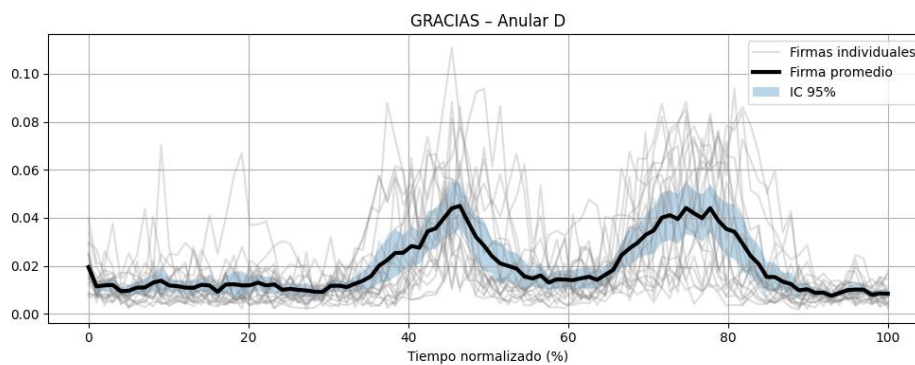
*Modelo promedio del desplazamiento global en el eje y de los 21 puntos la mano derecha con IC del 95 % – gesto gracias.*



*Nota. Las señales en gris muestran las firmas individuales del desplazamiento global de los 21 puntos de la mano en el eje y de cada video usado para sacar el modelo promedio.*

**Figura 32**

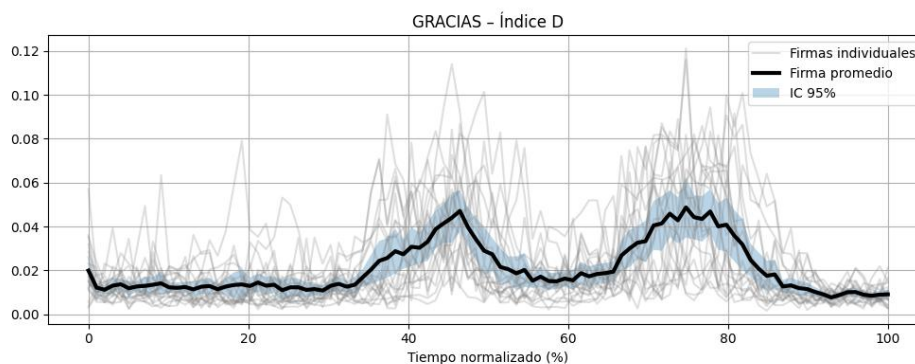
*Modelo promedio del desplazamiento del dedo anular derecho con IC del 95 % – gesto gracias.*



*Nota. Las señales en gris muestran las firmas individuales del desplazamiento del dedo anular de cada video usado para sacar el modelo promedio.*

**Figura 33**

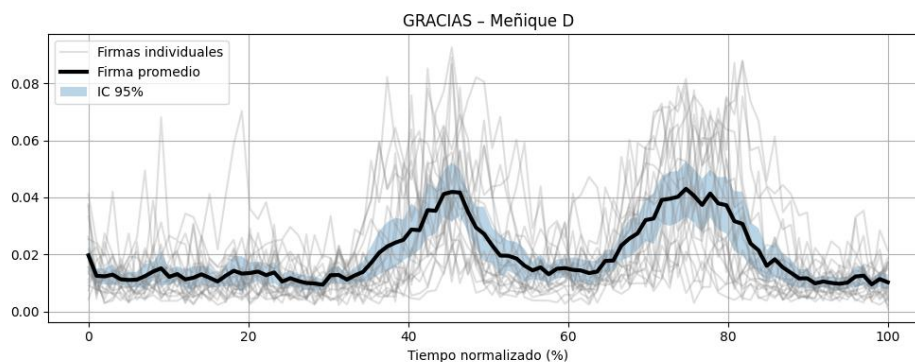
*Modelo promedio del desplazamiento del dedo índice derecho con IC del 95 % – gesto gracias.*



*Nota. Las señales en gris muestran las firmas individuales del desplazamiento del dedo índice de cada video usado para sacar el modelo promedio.*

**Figura 34**

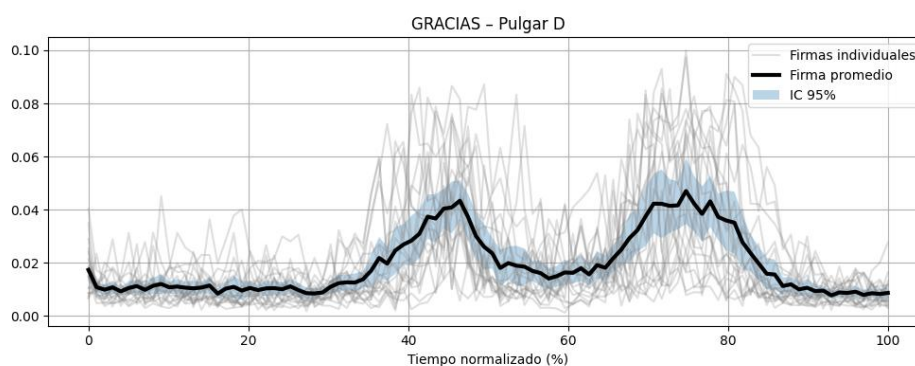
*Modelo promedio del desplazamiento del dedo meñique derecho con IC del 95 % – gesto gracias.*



*Nota. Las señales en gris muestran las firmas individuales del desplazamiento del dedo meñique de cada video usado para sacar el modelo promedio.*

**Figura 35**

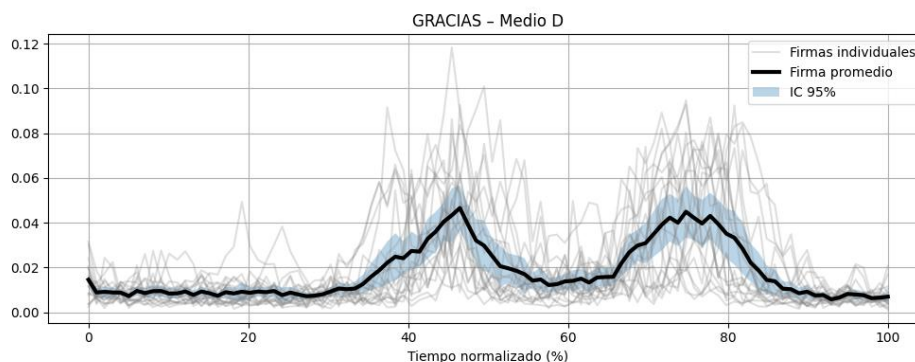
*Modelo promedio del desplazamiento del dedo pulgar derecho con IC del 95 % – gesto gracias.*



*Nota. Las señales en gris muestran las firmas individuales del desplazamiento del dedo pulgar de cada video usado para sacar el modelo promedio.*

**Figura 36**

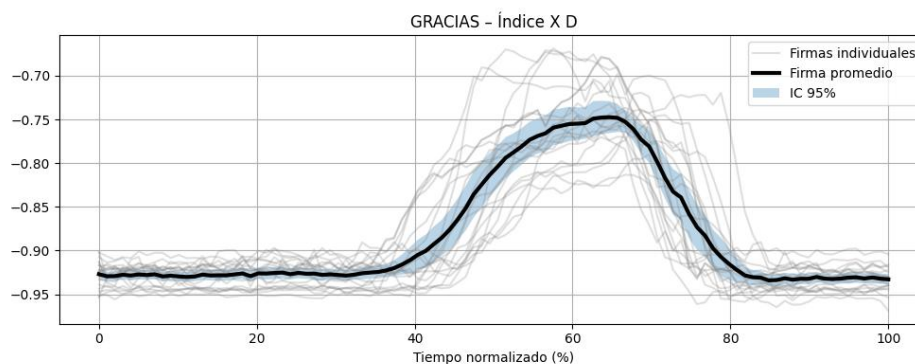
*Modelo promedio del desplazamiento del dedo medio derecho con IC del 95 % – gesto gracias.*



*Nota. Las señales en gris muestran las firmas individuales del desplazamiento del dedo medio de cada video usado para sacar el modelo promedio.*

**Figura 37**

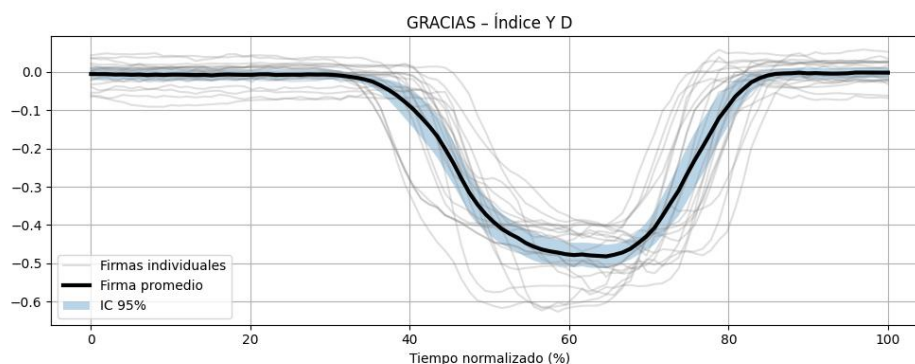
*Modelo promedio del movimiento en x del dedo índice derecho con IC del 95 % – gesto gracias.*



*Nota. Las señales en gris muestran las firmas individuales del movimiento en x de cada video usado para sacar el modelo promedio.*

**Figura 38**

*Modelo promedio del movimiento en y del dedo índice derecho con IC del 95 % – gesto gracias.*



*Nota. Las señales en gris muestran las firmas individuales del movimiento en y de cada video usado para sacar el modelo promedio.*

### **Creación modelo promedio DTW (DBA).**

Para una representación más robusta del gesto frente a variaciones temporales entre ejecuciones, se implementa un método de promediado basado en Deformación Temporal Dinámica (*Dynamic Time Warping, DTW*). De esta manera nos permite alinear temporalmente las señales antes de combinarlas, evitando la pérdida de información causada por desfases en la ejecución del gesto.

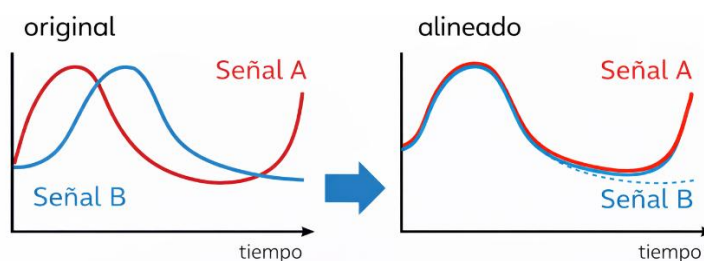
Para ello, se emplea la librería *dtw* de *Python*, la cual proporciona una implementación eficiente del algoritmo DTW para la alineación de secuencias temporales. Para mayor análisis de la librería se tiene la documentación oficial disponible en: <https://pypi.org/project/dtw/>

**Alineación de firmas mediante DBA.** Para cada conjunto de firmas individuales, aplicamos el algoritmo *Dynamic Time Warping Barycenter Averaging* (DBA). Para ello, se selecciona una de las firmas individuales disponibles como referencia inicial en este caso

corresponde al primer video procesado. Esta firma individual actúa únicamente como un punto de partida y no representa aún el gesto promedio definitivo. Posteriormente, cada una de las demás señales se alinea temporalmente con esta referencia utilizando el algoritmo de Deformación Temporal Dinámica (DTW). Este algoritmo permite ajustar las señales en el eje temporal decidiendo el mejor camino que demuestre su igualdad de gestos entre ejecuciones, estirando o comprimiendo segmentos de manera que los movimientos similares queden alineados, incluso si fueron ejecutados a distintas velocidades como se muestra en la Figura 39.

**Figura 39**

*Ejemplo de alineación temporal de 2 señales mediante DBA.*



Nota. Proceso de alineamiento temporal mediante *Dynamic Time Warping Barycenter Averaging* (DBA). Adaptado de (Sakoe & Chiba, 1978).

Una vez realizada la alineación, el algoritmo identifica qué valores de las distintas ejecuciones corresponden a un mismo instante del gesto para posteriormente realizar el modelo promedio que representa el comportamiento común del conjunto de ejecuciones.

**Creación firma promedio DTW.** Una vez que las firmas individuales han sido alineadas con DTW, se calcula un promedio considerando cada una de las 34 componentes que

describen el gesto. El cálculo se realiza de forma independiente por componente y por frame, lo que da lugar a una representación final temporalmente alineada.

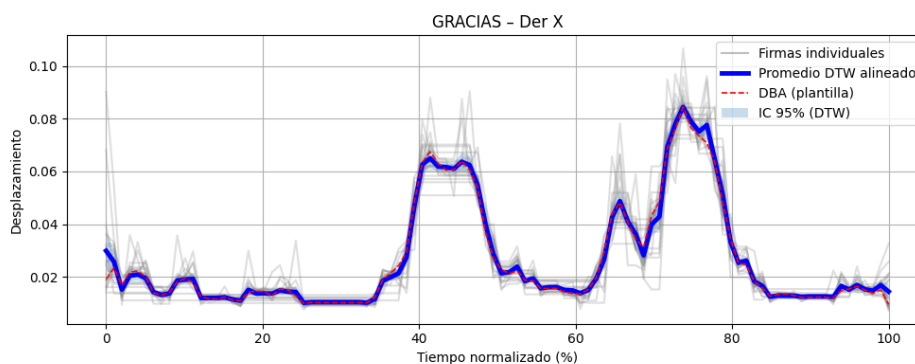
Adicionalmente, se estima un intervalo de confianza del 95 % a partir de las señales ya alineadas. El modelo promedio obtenido mediante DTW presenta un comportamiento más consistente en comparación con el promedio basado en Distancia Euclidiana, especialmente cuando existen variaciones en la duración o en la velocidad de ejecución del gesto.

De la Figura 40 a la Figura 46 se muestran los alineamientos temporales de las de las firmas individuales de cada una de las ejecuciones del gesto “gracias” de cada uno de los componentes de desplazamiento de la mano derecha, obtenido mediante *Dynamic Time Warping* (DTW), junto con la plantilla promedio calculada mediante el algoritmo DBA y su modelo promedio, además de su correspondiente intervalo de confianza del 95 % para el desplazamiento de la mano derecha, este proceso se repite para la mano izquierda y para el resto de los gestos.

En la Figura 47 se muestra los alineamientos temporales de las de las firmas individuales de cada una de las ejecuciones del movimiento en el eje x de la punta del dedo índice de la mano derecha del gesto “gracias”, obtenido mediante *Dynamic Time Warping* (DTW), junto con la plantilla promedio calculada mediante el algoritmo DBA y su modelo promedio, además de su correspondiente intervalo de confianza del 95 % para el desplazamiento de la mano derecha, mientras que en la Figura 48 se muestra el mismo proceso aplicado movimiento en y de la punta del dedo índice de la mano derecha, este proceso se realiza para cada una de las puntas de los dedos de ambas manos y para el resto de los gestos.

**Figura 40**

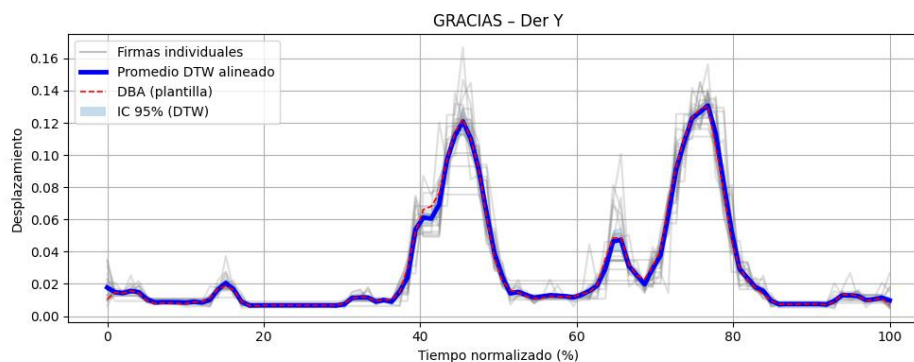
*Modelo promedio obtenido del desplazamiento en  $x$  de los 21 puntos de la mano derecha del gesto “gracias” mediante DTW (DBA).*



*Nota. Las señales en gris muestran las firmas individuales alineadas del desplazamiento en  $x$  de cada video usado para sacar el modelo promedio.*

**Figura 41**

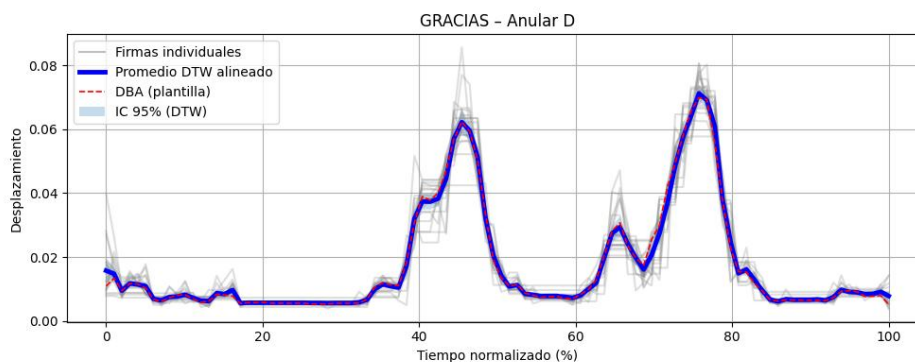
*Modelo promedio obtenido del desplazamiento en  $y$  de los 21 puntos de la mano derecha del gesto “gracias” mediante DTW (DBA).*



*Nota. Las señales en gris muestran las firmas individuales alineadas del desplazamiento en  $y$  de cada video usado para sacar el modelo promedio.*

**Figura 42**

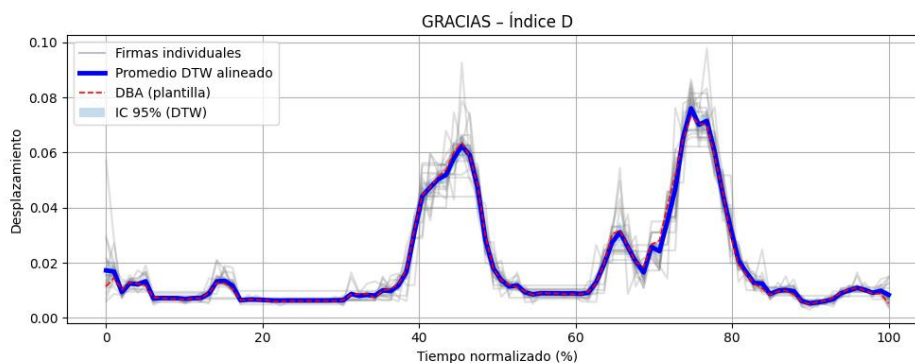
*Modelo promedio obtenido del desplazamiento del dedo anular de la mano derecha del gesto “gracias” mediante DTW (DBA).*



*Nota. Las señales en gris muestran las firmas individuales alineadas del desplazamiento del dedo anular de cada video usado para sacar el modelo promedio.*

**Figura 43**

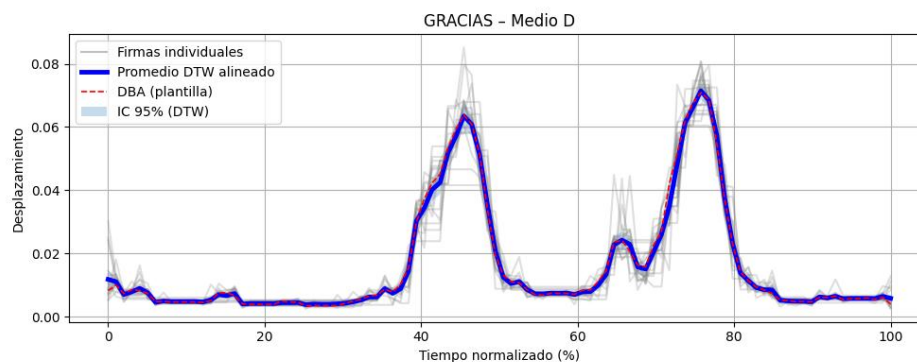
*Modelo promedio obtenido del desplazamiento del dedo índice de la mano derecha del gesto “gracias” mediante DTW (DBA).*



*Nota. Las señales en gris muestran las firmas individuales alineadas del desplazamiento del dedo índice de cada video usado para sacar el modelo promedio.*

**Figura 44**

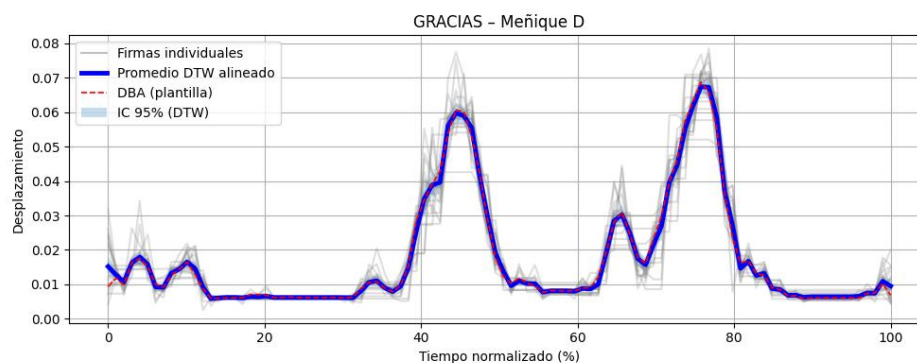
*Modelo promedio obtenido del desplazamiento del dedo medio de la mano derecha del gesto “gracias” mediante DTW (DBA).*



*Nota. Las señales en gris muestran las firmas individuales alineadas del desplazamiento del dedo medio de cada video usado para sacar el modelo promedio.*

**Figura 45**

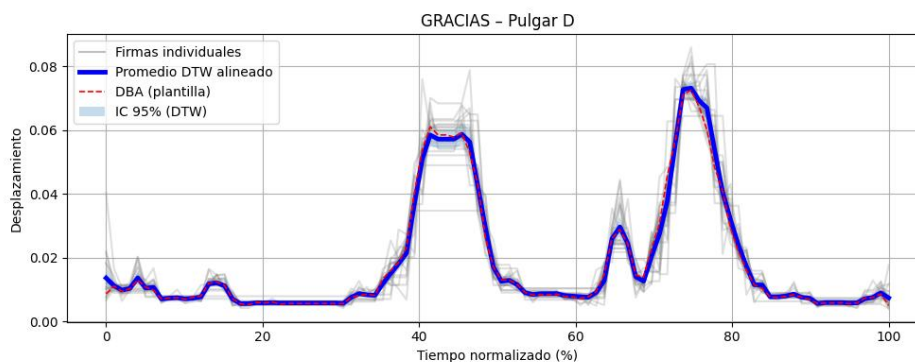
*Modelo promedio obtenido del desplazamiento del dedo meñique de la mano derecha del gesto “gracias” mediante DTW (DBA).*



*Nota. Las señales en gris muestran las firmas individuales alineadas del desplazamiento del dedo meñique de cada video usado para sacar el modelo promedio.*

**Figura 46**

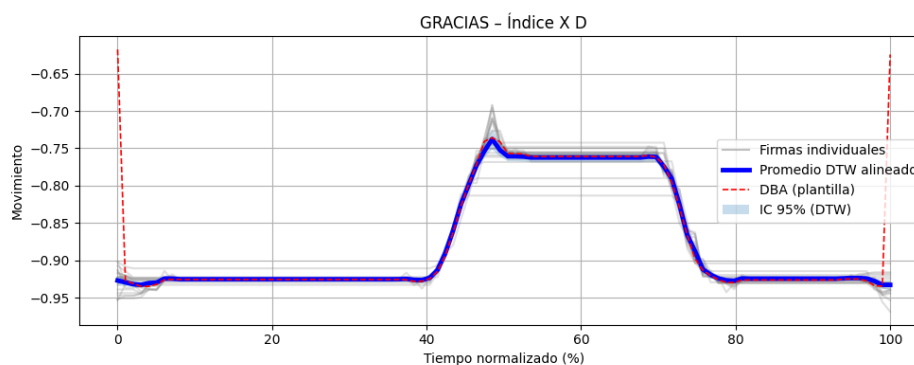
*Modelo promedio obtenido del desplazamiento del dedo pulgar de la mano derecha del gesto “gracias” mediante DTW (DBA).*



*Nota. Las señales en gris muestran las firmas individuales alineadas del desplazamiento del dedo pulgar de cada video usado para sacar el modelo promedio.*

**Figura 47**

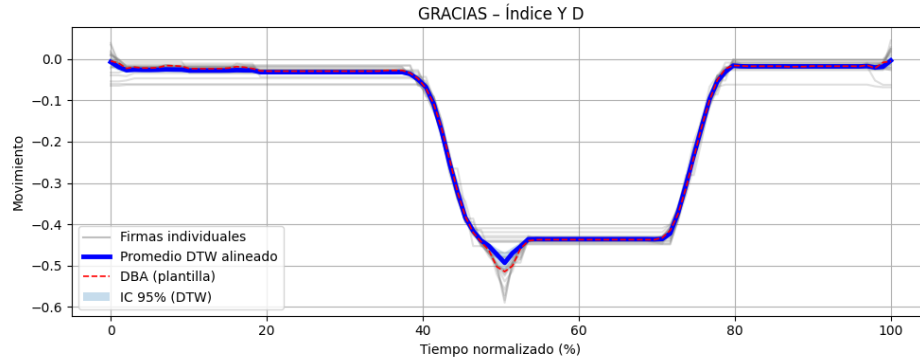
*Modelo promedio obtenido del movimiento en el eje  $x$  del dedo índice de la mano derecha del gesto “gracias” mediante DTW (DBA).*



*Nota. Las señales en gris muestran las firmas individuales alineadas del movimiento en  $x$  de la punta del dedo índice de cada video usado para sacar el modelo promedio.*

**Figura 48**

*Modelo promedio obtenido del movimiento en el eje y del dedo índice de la mano derecha del gesto “gracias” mediante DTW (DBA).*



*Nota. Las señales en gris muestran las firmas individuales alineadas del movimiento en y de la punta del dedo índice de cada video usado para sacar el modelo promedio.*

El resultado para el modelamiento DTW después de tener las firmas individuales alineadas es una matriz de dimensión  $100 \times 34$ , que constituye el modelo promedio del gesto mediante DTW.

La matriz final del modelo del gesto mediante DTW queda expresada de la siguiente manera:

$$\mathbf{M}_{\text{gesto DTW}} = \begin{bmatrix} m_{1,1} & m_{1,2} & m_{1,3} & \dots & m_{1,34} \\ m_{2,1} & m_{2,2} & m_{2,3} & \dots & m_{2,34} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{100,1} & m_{100,2} & m_{100,3} & \dots & m_{100,34} \end{bmatrix}$$

Donde:

- $\mathbf{M}_{\text{gesto DTW}}$  es la matriz final del modelo del gesto DTW.

- Cada fila representa un instante temporal normalizado [1, 100].
- Cada columna representa una componente específica (34 componentes).

### ***Validación de modelos promedio de gestos.***

Una vez contruidos los modelos promedio de cada uno de los gestos se procede a la validación del modelo. Por ello, se evalúa la capacidad del sistema para modelar correctamente un gesto a partir de una nueva ejecución de un video test, comparándola con los modelos base creados y obteniendo la diferencia de distancia entre la firma del video test con cada uno de los modelos promedio de los diferentes gestos con lo que evaluamos la similitud entre el video test y los modelos promedio.

Analizamos dos enfoques de modelamiento distintos: basado en Distancia Euclidiana y basado en Deformación Temporal Dinámica (DTW). Ambos métodos utilizan como entrada el modelo promedio de cada gesto obtenidos mediante su propia métrica.

**Validación del modelamiento por el método de la Distancia Euclidiana.** La validación se realiza tomando la firma correspondiente a una ejecución de prueba y comparándola con los modelos promedio de los gestos básicos almacenados en el sistema. Tanto la muestra evaluada como los modelos de referencia comparten la misma longitud temporal y la misma estructura de datos, lo que facilita una comparación directa entre las secuencias.

Para cada gesto, se calcula la distancia frame a frame entre la firma de prueba y los modelos generados de forma independiente para cada una de las 34 componentes. Estas distancias cuantifican la diferencia entre ambas señales a lo largo del tiempo. Posteriormente, las distancias correspondientes a las componentes de la mano derecha y de la mano izquierda

se agrupan por separado y se calcula la media aritmética para obtener dos distancias totales por gesto, es decir una distancia total por las componentes de desplazamiento y una distancia total por las componentes de movimiento.

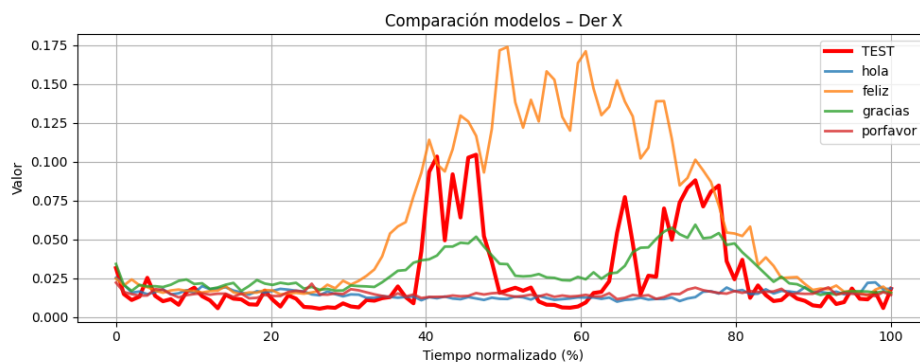
El gesto con la menor distancia total en desplazamiento y movimiento se describe como el gesto más parecido, ya que representa el mayor grado de similitud entre la ejecución de prueba y el modelo promedio. Este criterio se repite para todos los gestos considerados, permitiendo determinar el gesto más parecido a la firma de prueba de manera sencilla y eficiente.

De igual forma, de la Figura 49 a la Figura 55 se muestra las gráficas comparativas entre la firma de prueba (*test*) del gesto gracias y los modelos promedios de los distintos gestos para los 7 componentes de desplazamiento de la mano derecha mediante el método de la Distancia Euclidiana de la misma forma en la Figura 56 y Figura 57 se muestra la comparación aplicadas a 2 componentes de movimiento del dedo medio en  $x$  y  $y$  respectivamente, y así mismo se realiza para la mano izquierda y para el resto de los gestos.

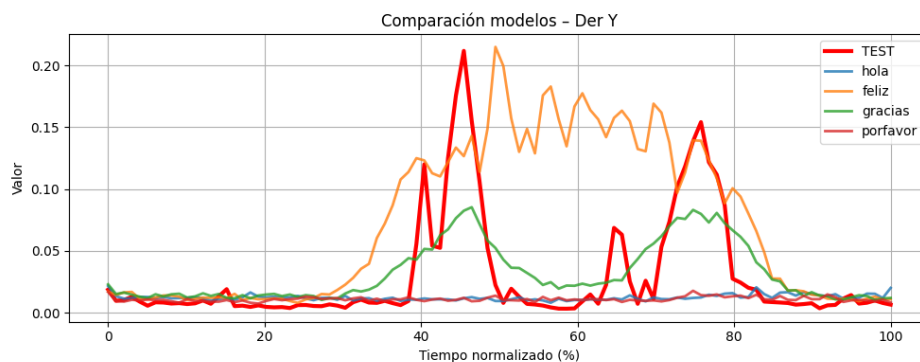
En la Figura 58 se observa la comparación entre la firma de prueba (*test*) y el modelo promedio del gesto “gracias” de la componente de desplazamiento del dedo meñique y Figura 59 se observa la comparación entre la firma de prueba (*test*) y el modelo promedio del gesto “gracias” del movimiento en  $x$  del dedo índice, ambas graficas de la mano derecha lo que permite visualizar el grado de similitud, esto se realiza también para la mano izquierda y es aplicable a el resto de los modelos de gestos.

**Figura 49**

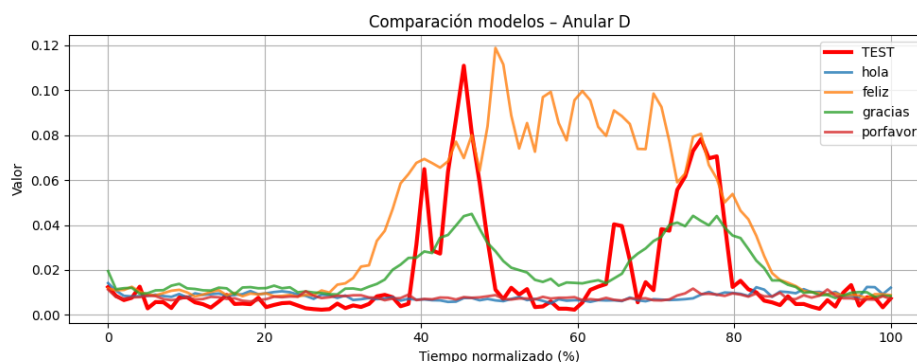
*Comparación de la firma test vs modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento en el eje  $x$  global 21 puntos de la mano derecha.*

**Figura 50**

*Comparación de la firma test vs modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento en el eje  $y$  global 21 puntos de la mano derecha.*

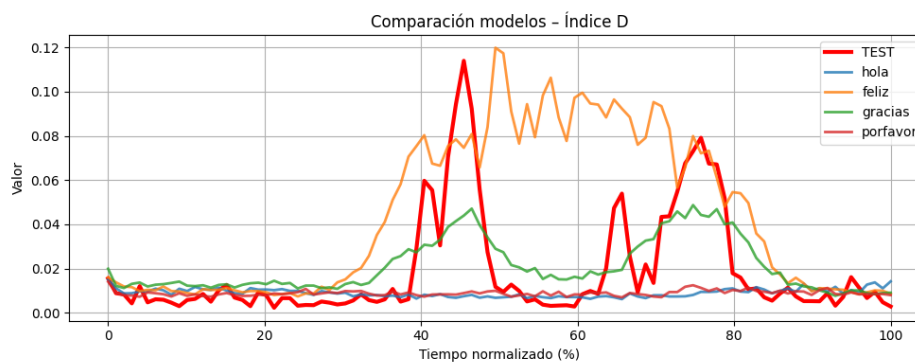


**Figura 51** Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento del dedo anular de la mano derecha.



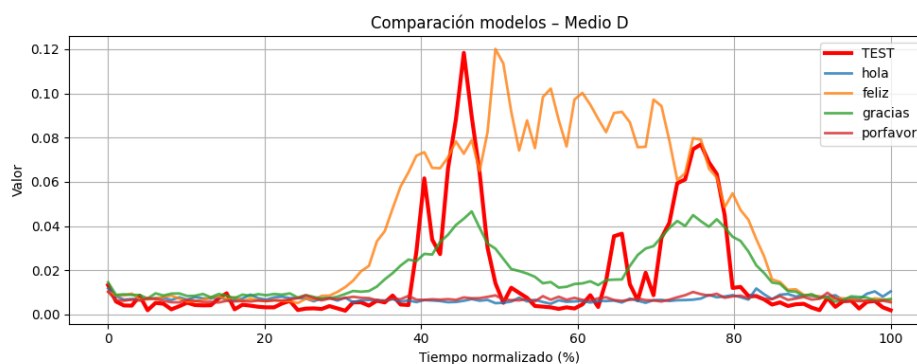
**Figura 52**

Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento del dedo índice de la mano derecha.

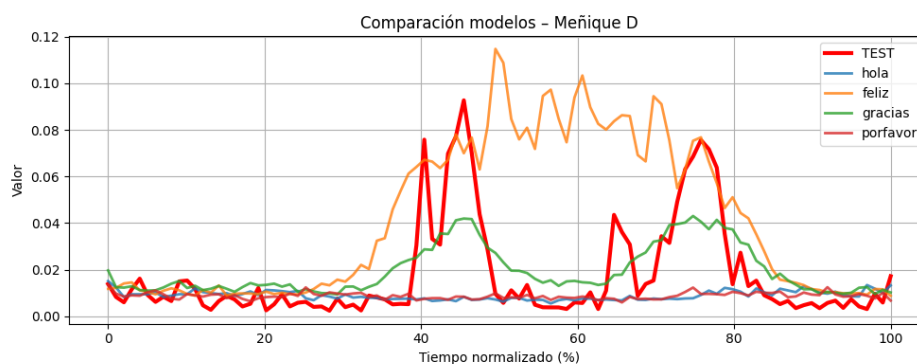


**Figura 53**

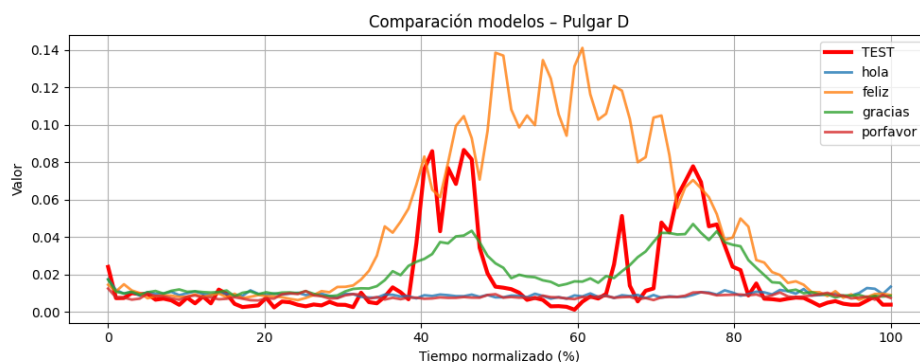
*Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento del dedo medio de la mano derecha.*

**Figura 54**

*Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento del dedo meñique de la mano derecha.*

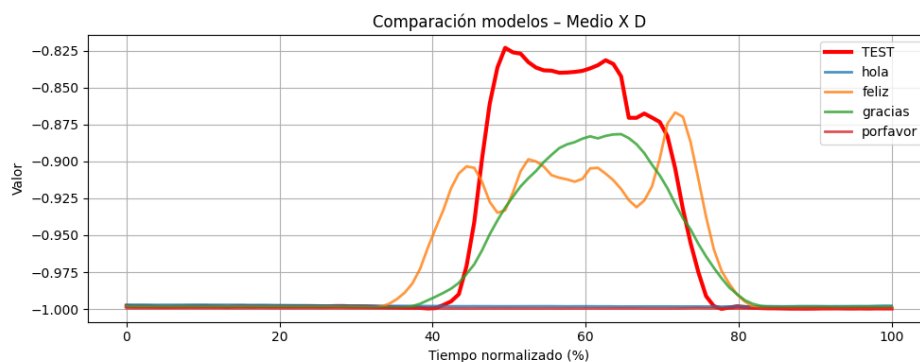


**Figura 55** Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el desplazamiento del dedo pulgar de la mano derecha.



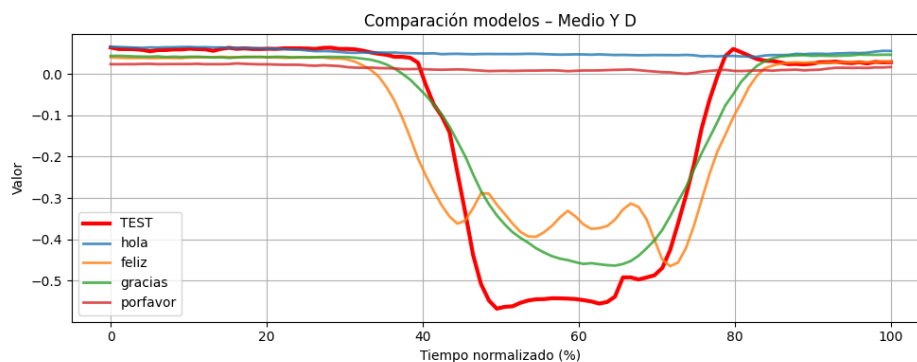
**Figura 56**

Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el movimiento del dedo medio en x de la mano derecha.

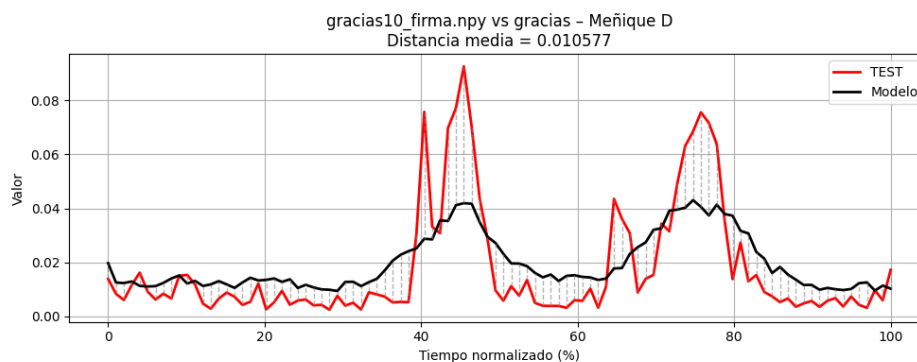


**Figura 57**

*Comparación de la firma test vs los modelos promedio de los gestos mediante distancia euclidiana para el movimiento del dedo medio en y de la mano derecha.*

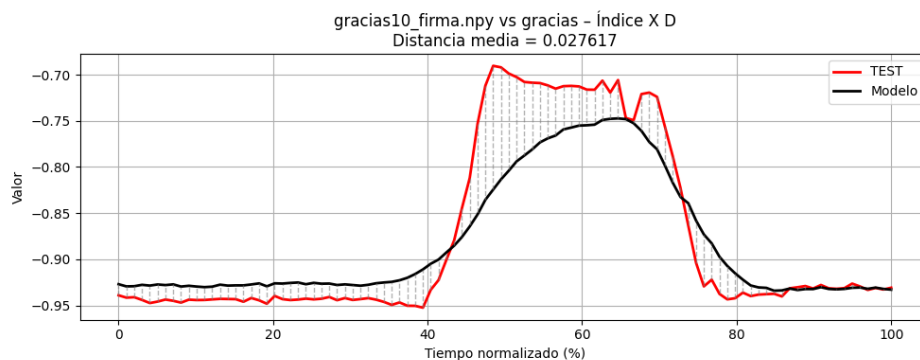
**Figura 58**

*Comparación de la firma test y del modelo del gesto “gracias” mediante distancia euclidiana para el desplazamiento del dedo meñique de la mano derecha.*



**Figura 59**

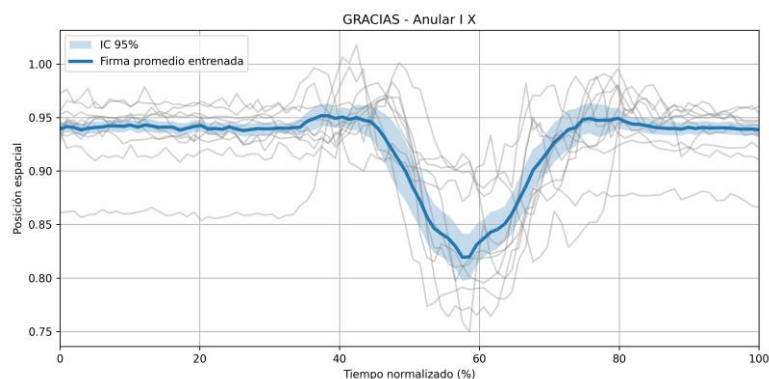
Comparación de la firma test y del modelo del gesto “gracias” mediante distancia euclidiana para el movimiento en  $x$  del dedo índice de la mano derecha.



Mientras que en la Figura 60 se observa la comparación de 10 firmas individuales de prueba (videos de la persona 4) con el modelo promedio aplicado a la componente del movimiento en  $x$  de la punta del dedo índice de la mano derecha. En la Figura 61 se muestra la misma comparativa aplicada al componente del movimiento en  $y$  de la punta del dedo índice de la mano derecha.

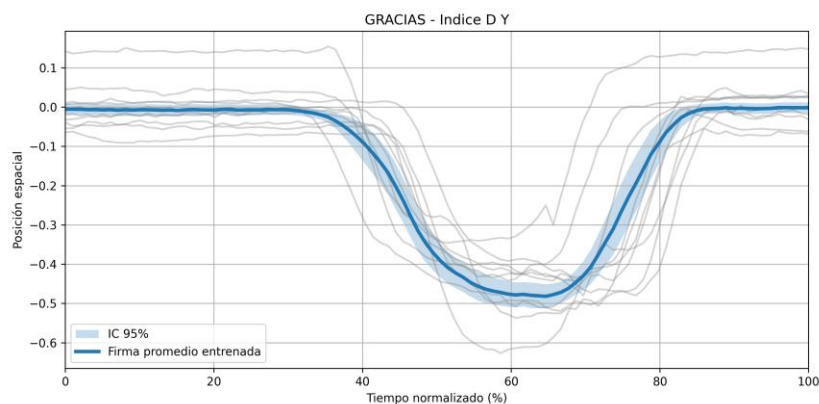
**Figura 60**

Comparación de 10 firmas test vs modelo del gesto “gracias” mediante distancia euclidiana para el movimiento en  $x$  del dedo índice de la mano derecha.



**Figura 61**

*Comparación de 10 firmas test y del modelo del gesto “gracias” mediante distancia euclidiana para el movimiento en y del dedo índice de la mano derecha.*



La Tabla 1 presenta los resultados del proceso de validación del modelo utilizando Distancia Euclidiana, donde se comparan las distancias obtenidas entre la firma de prueba y los modelos promedios de los 7 componentes de desplazamiento correspondientes a cada gesto, considerando de manera independiente la contribución de la mano derecha y la mano izquierda, así como el valor medio total resultante, en la Tabla 2 se muestra la similitud según la los resultados de la Tabla 1 la cual permite identificar el gesto con mayor grado de coincidencia respecto a la firma individual de prueba (*test*).

**Tabla 1**

*Resultados Distancia Euclidiana respecto a las 7 componentes de desplazamiento.*

| Gesto    | Mano derecha | Mano izquierda | Media Total |
|----------|--------------|----------------|-------------|
| HOLA     | 0.021065     | 0.190606       | 0.105836    |
| FELIZ    | 0.031025     | 0.026163       | 0.028594    |
| GRACIAS  | 0.014682     | 0.018346       | 0.016514    |
| PORFAVOR | 0.020864     | 0.021676       | 0.021270    |

**Tabla 2**

*Resultados Similitud Distancia Euclidiana respecto a las 7 componentes de desplazamiento.*

| Gesto    | Mano derecha | Mano izquierda | Media Total |
|----------|--------------|----------------|-------------|
| HOLA     | 0.982233     | 0.847199       | 0.914716    |
| FELIZ    | 0.966206     | 0.973144       | 0.969675    |
| GRACIAS  | 0.985484     | 0.982559       | 0.984022    |
| PORFAVOR | 0.982390     | 0.979300       | 0.980845    |

Mientras que en la Tabla 3 se presenta los resultados del proceso de validación del modelo utilizando Distancia Euclidiana, donde se comparan las distancias obtenidas entre la firma de prueba y los modelos promedios de los 10 componentes de movimiento correspondientes a cada gesto, considerando de manera independiente la contribución de la mano derecha y la mano izquierda, así como el valor total resultante, de igual forma en la Tabla 4 se tiene los resultados de similitud según la Tabla 3 así mismo nos permite identificar el gesto con mayor grado de coincidencia respecto a la firma individual de prueba (*test*).

**Tabla 3**

*Resultados Distancia Euclidiana con respecto a las 10 componentes de movimiento (punta de los dedos).*

| Gesto    | Mano derecha | Mano izquierda | Media Total |
|----------|--------------|----------------|-------------|
| HOLA     | 0.233415     | 0.297547       | 0.265481    |
| FELIZ    | 0.076022     | 0.078783       | 0.077403    |
| GRACIAS  | 0.039410     | 0.043672       | 0.041541    |
| PORFAVOR | 0.122693     | 0.150751       | 0.136722    |

**Tabla 4**

*Resultados Similitud Distancia Euclidiana con respecto a las 10 componentes de movimiento (punta de los dedos).*

| Gesto    | Mano derecha | Mano izquierda | Media Total |
|----------|--------------|----------------|-------------|
| HOLA     | 0.821563     | 0.760018       | 0.790791    |
| FELIZ    | 0.929087     | 0.927264       | 0.928176    |
| GRACIAS  | 0.964773     | 0.958575       | 0.961674    |
| PORFAVOR | 0.903573     | 0.873594       | 0.888584    |

El valor de distancia y similitud mínimo tanto para las componentes de movimiento como las componentes de desplazamiento indica el gesto más parecido a la firma test, en este caso de acuerdo con la Tabla 1, Tabla 2, Tabla 3 y Tabla 4 podemos decir que el gesto más parecido a la firma individual prueba, es el gesto “gracias”.

**Validación del modelamiento por el método DTW.** Para validar el modelamiento DTW se compara la firma individual de una ejecución de prueba (*test*) con los modelos promedio calculados y almacenadas previamente de cada gesto. Este método permite compensar diferencias en la velocidad de ejecución, alineando las señales antes de evaluar su similitud.

Para cada gesto, la firma de prueba se compara con los modelos almacenados utilizando DTW.

Por ejemplo, la firma de prueba se compara punto a punto con el modelo promedio almacenado  $M_{gesto\ DTW}$  utilizando DTW, es decir se calcula una distancia media temporal punto

a punto entre la firma de prueba y el modelo correspondiente. Para cada componente de desplazamiento y de movimiento  $i$ , la distancia se define como:

$$d_i = \frac{1}{T} \sum_{t=1}^T |F_{test}(t, i) - M_{Gesto\ DTW}(t, i)|$$

donde:

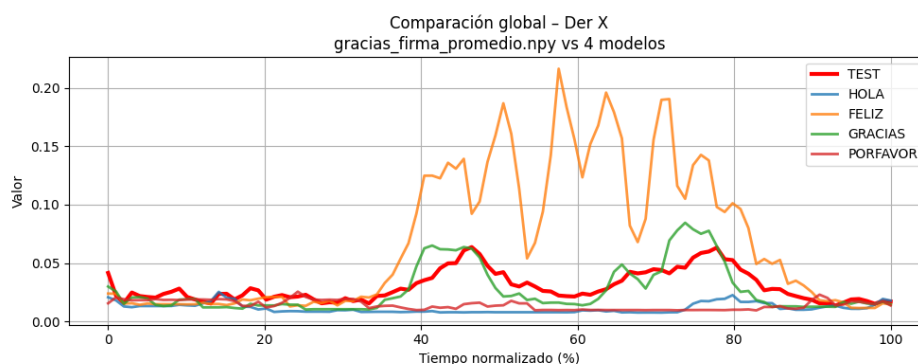
- $F_{test}(t, i)$  es el valor de la componente  $i$  en el instante  $t$  de la firma de prueba,
- $M_{DTW}(t, i)$  es el valor del modelo promedio,
- $T$  es la longitud temporal normalizada (100 frames).

Las distancias se agrupan por mano y se promedian para obtener una distancia total por gesto. El gesto cuya media de la distancia total es mínima tanto en las 7 componentes de desplazamiento como en las 10 componentes de movimiento, se considera el gesto más parecido a la firma test. De manera complementaria, se calcula una medida de similitud acotada en el intervalo  $[0, 1]$ , donde valores cercanos a 1 indican una mayor correspondencia entre la ejecución de la firma individual de prueba y el modelo.

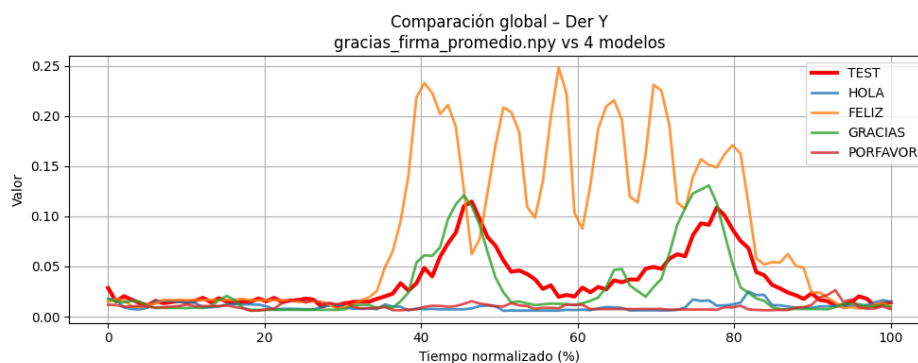
De igual forma, de la Figura 62 a la Figura 68 se muestra las gráficas comparativas entre la firma de prueba (*test*) y los modelos obtenidos mediante el método DTW de las 7 componentes de desplazamiento de la mano derecha de los gestos lo que permite visualizar el grado de similitud y facilita el análisis de los resultados obtenidos., el mismo proceso se aplica para las 10 componente de movimiento en  $x$  y en  $y$  de las puntas de los dedos, de igual forma se realiza para la mano izquierda y para la mano derecha. Este proceso es aplicable al resto de los modelos de los gestos.

**Figura 62**

*Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento en el eje x de la mano derecha.*

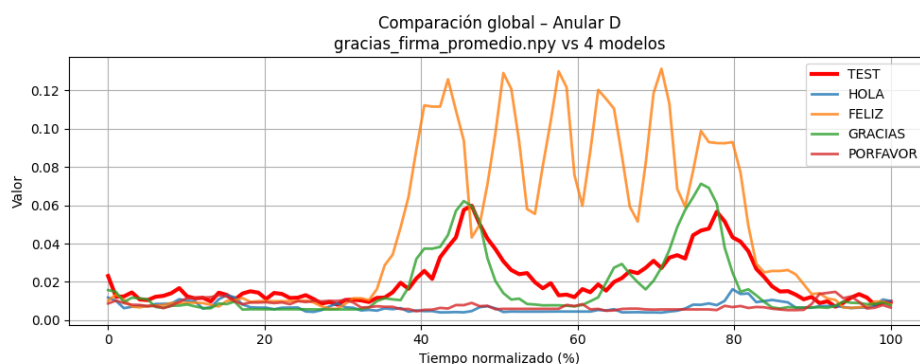
**Figura 63**

*Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento en el eje y de la mano derecha.*

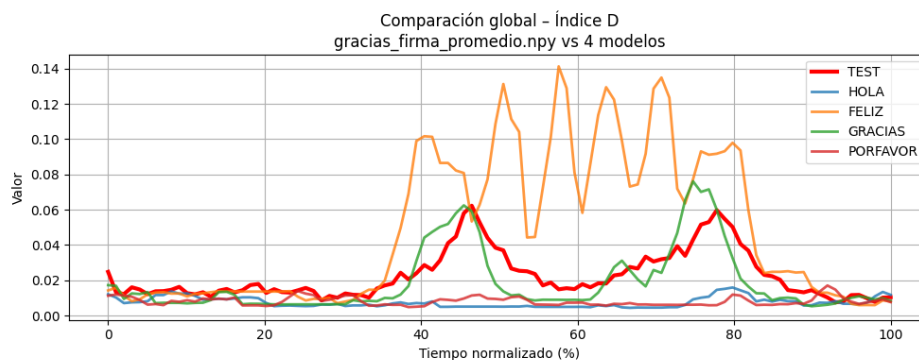


**Figura 64**

*Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo anular de la mano derecha.*

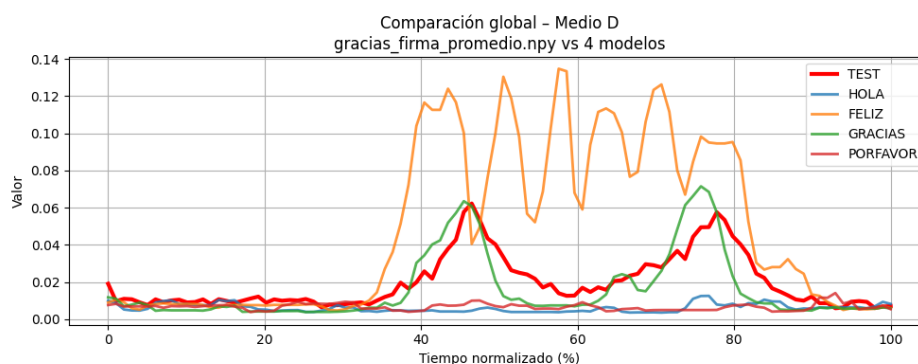
**Figura 65**

*Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo índice de la mano derecha.*

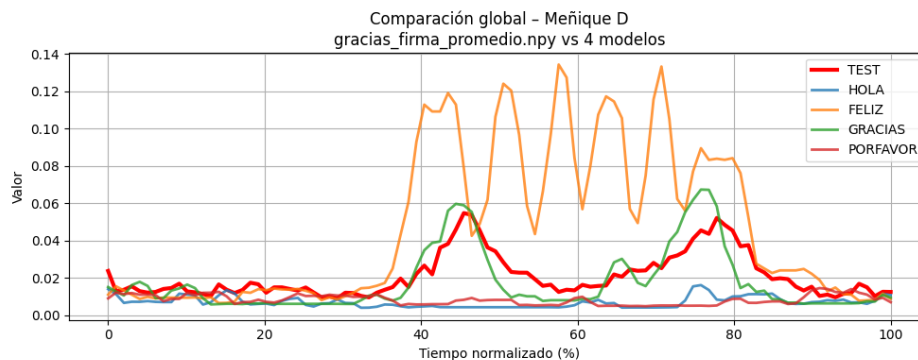


**Figura 66**

*Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo medio de la mano derecha.*

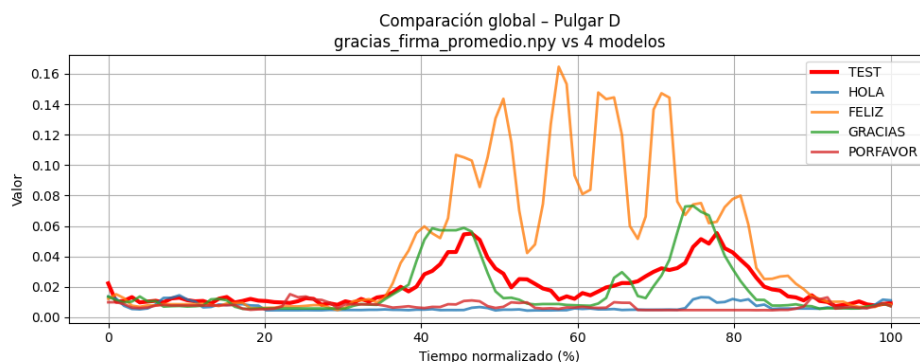
**Figura 67**

*Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo meñique de la mano derecha.*



### Figura 68

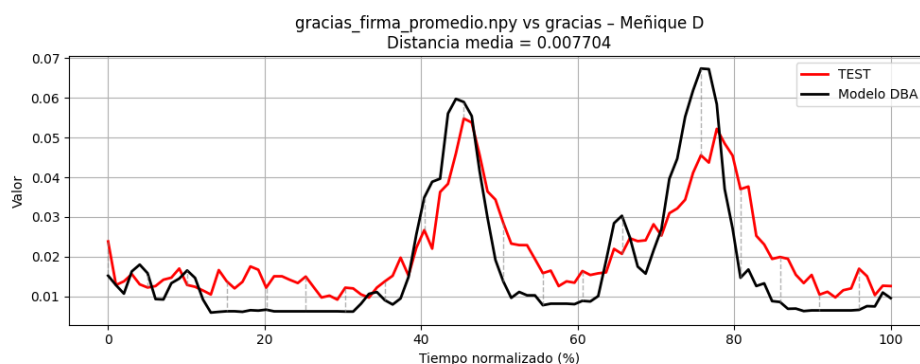
*Comparación de la firma individual de prueba (test) y cada uno de los modelos de los gestos obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo pulgar de la mano derecha.*



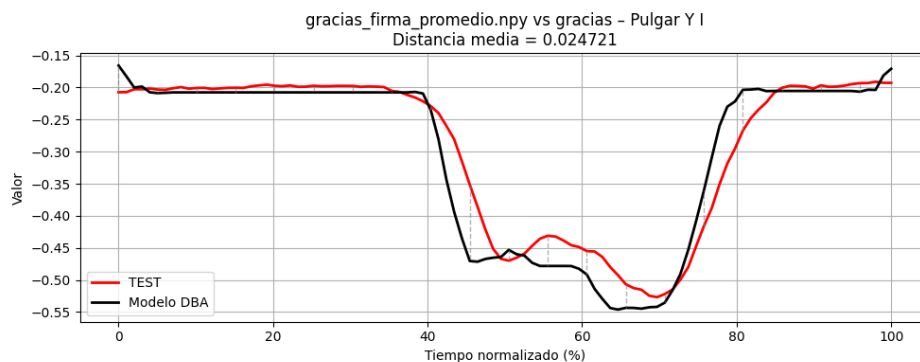
En la Figura 69 se observa la comparación entre la firma de prueba (*test*) y el modelo promedio del gesto “gracias” del desplazamiento del dedo meñique y Figura 70 se observa la comparación entre la firma de prueba (*test*) y el modelo promedio del gesto “gracias” del movimiento en  $x$  del dedo índice, ambas graficas de la mano derecha lo que permite visualizar el grado de similitud, esto se realiza para el resto de componentes así como para la mano izquierda y es aplicable a el resto de los modelos de gestos.

**Figura 69**

*Comparación de la firma individual de prueba (test) con el modelo del gesto “gracias” obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo meñique de la mano derecha.*

**Figura 70**

*Comparación de la firma individual de prueba (test) con el modelo del gesto “gracias” obtenidos mediante Dynamic Time Warping (DTW) para el movimiento en y del dedo pulgar de la mano derecha.*

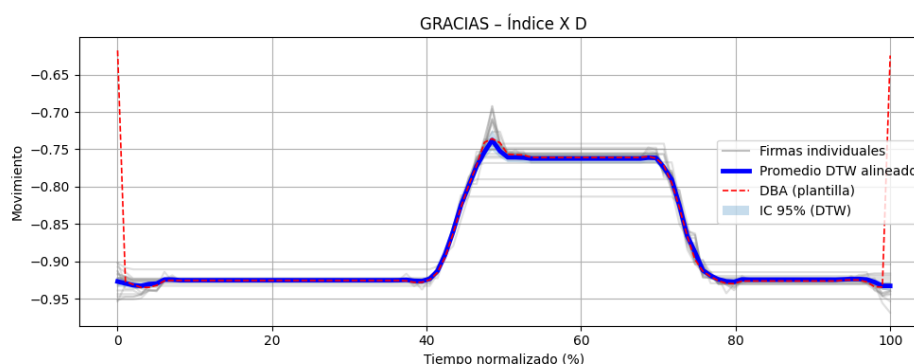


Mientras que en la Figura 71 se observa la comparación de 10 firmas individuales de prueba (videos de la persona 4) con el modelo promedio DTW aplicado a la componente del movimiento en  $x$  de la punta del dedo índice de la mano derecha. De igual forma en la Figura

72 se muestra la misma comparativa aplicada al componente del movimiento en y de la punta del dedo índice de la mano derecha.

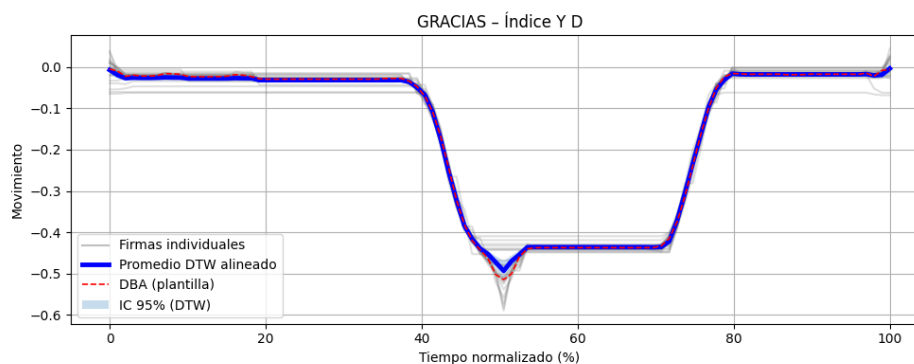
**Figura 71**

*Comparación de 10 firmas test (videos de prueba) con el modelo del gesto “gracias” mediante el método DTW para el movimiento en x del dedo índice de la mano derecha.*



**Figura 72**

*Comparación de 10 firmas test (videos de prueba) con el modelo del gesto “gracias” mediante el método DTW para el movimiento en y del dedo índice de la mano derecha.*



A continuación, en la Tabla 5 presenta los resultados del proceso de validación utilizando *Dynamic Time Warping* (DTW), donde se comparan las distancias obtenidas entre la

firma de prueba (*test*) y los modelos promedios DTW correspondientes a las 7 componentes de desplazamiento de cada gesto, considerando de manera independiente la contribución de la mano derecha y la mano izquierda, así como el valor total resultante. Adicionalmente, se incluye la similitud normalizada en el rango [0,1] mostrado en la Tabla 6, la cual permite identificar el gesto con mayor grado de coincidencia respecto a la firma individual de prueba, donde un valor de similitud más cercano al 1 indica el gesto predicho.

**Tabla 5**

*Distancia componentes de desplazamiento para DTW.*

| Gesto    | Mano derecha | Mano izquierda | Total   |
|----------|--------------|----------------|---------|
| HOLA     | 0.01677      | 0.22386        | 0.12031 |
| FELIZ    | 0.03451      | 0.02324        | 0.02887 |
| GRACIAS  | 0.00808      | 0.01386        | 0.01097 |
| PORFAVOR | 0.01502      | 0.02074        | 0.01788 |

**Tabla 6**

*Distancia componentes de desplazamiento para DTW - Similitud.*

| Gesto    | Mano derecha | Mano izquierda | Total, Similitud |
|----------|--------------|----------------|------------------|
| HOLA     | 0.98350      | 0.81717        | 0.89264          |
| FELIZ    | 0.96663      | 0.97730        | 0.97200          |
| GRACIAS  | 0.99199      | 0.98633        | 0.98917          |
| PORFAVOR | 0.98520      | 0.97969        | 0.98246          |

Mientras que en la Tabla 7 se presenta los resultados del proceso de validación del modelo utilizando DTW, donde se comparan las distancias obtenidas entre la firma de prueba y los modelos promedios de los 10 componentes de movimiento correspondientes a cada gesto, considerando de manera independiente la contribución de la mano derecha y la mano izquierda, así como el valor total resultante, de igual forma nos permite identificar el gesto con mayor grado de coincidencia respecto a la firma individual de prueba (*test*). Adicionalmente, se incluye la similitud normalizada en el rango [0,1] mostrado en la Tabla 8, la cual permite identificar el gesto con mayor grado de coincidencia respecto a la firma individual de prueba, donde un valor de similitud más cercano al 1 indica el gesto predicho.

**Tabla 7**

*Distancia componentes de movimiento para DTW.*

| Gesto    | Mano derecha | Mano izquierda | Total   |
|----------|--------------|----------------|---------|
| HOLA     | 0.28578      | 0.31963        | 0.30271 |
| FELIZ    | 0.05441      | 0.09196        | 0.07319 |
| GRACIAS  | 0.02153      | 0.03132        | 0.02643 |
| PORFAVOR | 0.06403      | 0.16064        | 0.11234 |

**Tabla 8**

*Distancia componentes de movimiento para DTW - Similitud.*

| Gesto    | Mano derecha | Mano izquierda | Total Similitud |
|----------|--------------|----------------|-----------------|
| HOLA     | 0.77772      | 0.75783        | 0.76768         |
| FELIZ    | 0.94839      | 0.91580        | 0.93176         |
| GRACIAS  | 0.97903      | 0.96965        | 0.97423         |
| PORFAVOR | 0.93983      | 0.86174        | 0.89908         |

El valor de distancia mínimo tanto para las componentes de movimiento como las componentes de desplazamiento indica el gesto más parecido a la firma test, en este caso de acuerdo con la Tabla 5 y Tabla 7 podemos decir que el gesto más parecido a la firma individual prueba, es el gesto “gracias”.

De igual forma al analizar los valores de similitud de la Tabla 6 y Tabla 8, nos indica que el modelo DTW del gesto que contiene el valor de similitud más cercano a 1 es el modelo DTW gesto “gracias”, por lo que en similitud es el más parecido a la firma individual de prueba.

## Capítulo 4

### Validación del sistema

Las pruebas realizadas permiten evaluar el desempeño del sistema propuesto para la representación de gestos de la Lengua de Señas Ecuatoriana, desarrollado a partir de la extracción y análisis de puntos clave de las manos. Este capítulo se centra en la validación de los modelos generados, considerando distintos escenarios de prueba que permiten analizar la capacidad de discriminación del modelo.

Inicialmente, se describen los escenarios de validación utilizados, los cuales indican variaciones entre personas y entre distintas ejecuciones de un mismo gesto. Posteriormente, se presenta el análisis de los resultados obtenido a partir del cálculo de distancias entre una firma individual prueba (*test*) y los modelos obtenidos mediante Distancia Euclidiana y DTW.

Finalmente, se discuten los resultados obtenidos, destacando las diferencias observadas entre ambos métodos de modelamiento de gestos.

#### Escenario Pruebas

Para la validación del sistema se definieron distintos escenarios de prueba con el fin de evaluar el desempeño de los modelos obtenidos frente a variaciones entre personas y entre ejecuciones de un mismo gesto. El conjunto de datos está conformado por cuatro personas, cada una de las cuales realizó cuatro gestos distintos, con diez ejecuciones por gesto.

A partir de este conjunto, se plantearon dos esquemas de validación: uno orientado a evaluar el modelo obtenido con los datos de gestos de tres personas y evaluar con la firma de la cuarta persona, así como otro enfocado en evaluar el modelo obtenido con el 50% de ejecuciones de un mismo gesto por parte de cada una de las 4 personas y con el otro 50% de

videos realizar la validación. En ambos casos, el análisis se basa en el cálculo de distancias entre la firma individual prueba (*test*) y los modelos de los gestos obtenidos por ambos métodos.

El detalle de los escenarios de prueba, así como la distribución de los datos utilizados en cada caso, se presenta de manera resumida en la Tabla 9.

**Tabla 9**

*Descripción de los escenarios de prueba.*

| Especificaciones de los escenarios de prueba |                                                           |                                |                                                                                      |
|----------------------------------------------|-----------------------------------------------------------|--------------------------------|--------------------------------------------------------------------------------------|
| Escenario de validación                      | Conjunto de datos de video para el modelamiento del gesto | Conjunto de validación         | Descripción                                                                          |
| Validación por persona (Caso 1)              | 30 videos (Personas 1, 2 y 3)                             | 10 videos (Persona 4)          |                                                                                      |
| Validación por persona (Caso 2)              | 30 videos (Personas 1, 2 y 4)                             | 10 videos (Persona 3)          | El modelo se construye con tres personas y se valida con la persona restante.        |
| Validación por persona (Caso 3)              | 30 videos (Personas 1, 3 y 4)                             | 10 videos (Persona 2)          |                                                                                      |
| Validación por persona (Caso 4)              | 30 videos (Personas 2, 3 y 4)                             | 10 videos (Persona 1)          |                                                                                      |
| Validación por ejecuciones                   | 5 videos por persona                                      | 5 videos restantes por persona | Se evalúa la consistencia del modelo frente a distintas ejecuciones del mismo gesto. |

### **Métricas de evaluación.**

Para evaluar el desempeño de los modelos propuestos, el análisis se realiza con las distancias y medidas de similitud entre las firmas test y los modelos de los gestos, obtenidas con los métodos de Distancia Euclidiana y en Deformación Temporal Dinámica (DTW). Estas métricas permiten cuantificar el grado de similitud entre una ejecución de prueba de una firma test y los modelos base de cada gesto para cada uno de los métodos.

Las métricas empleadas en este trabajo de integración curricular son la Distancia Euclidiana total y la distancia DTW, junto con su respectiva medida de similitud para Distancia Euclidiana y DTW.

La Distancia Euclidiana permite evaluar la similitud entre firmas cuando las ejecuciones presentan una estructura temporal comparable, mientras que la distancia DTW y su correspondiente similitud proporcionan una evaluación más estable frente a variaciones en la velocidad de ejecución del gesto. El uso conjunto de estas métricas permite analizar el desempeño del sistema y respaldar los resultados obtenidos por ambos métodos de modelamientos.

### **Pruebas de Evaluación**

Con el sistema implementado, se realizaron diversas pruebas de evaluación con el objetivo de analizar el comportamiento de los modelos de gestos de la Lengua de Señas Ecuatoriana. Estas pruebas se realizaron utilizando firmas de prueba y los modelos promedio generados durante la etapa de modelamiento del gesto.

Las pruebas de evaluación se desarrollaron considerando los distintos escenarios de validación definidos previamente en la Tabla 9 y se basan en el cálculo de distancias entre

firma de prueba y modelo promedio de cada uno de los métodos calculados. Los resultados obtenidos se ilustran mediante gráficos comparativos y tablas de resultados, los cuales permiten visualizar el grado de similitud entre las ejecuciones de prueba y los modelos de cada gesto.

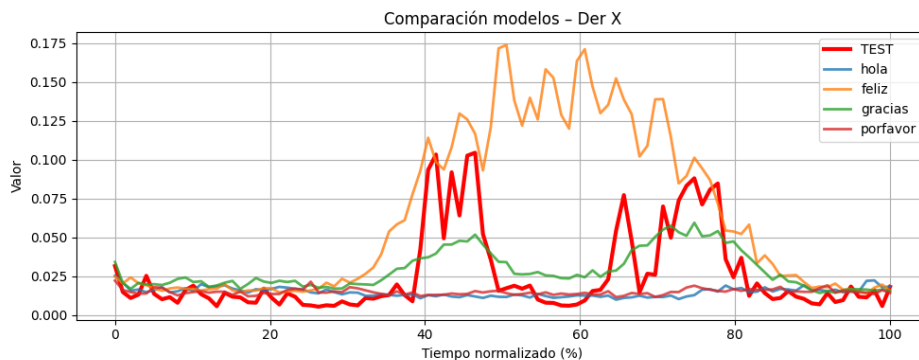
De la Figura 73 a la Figura 79 presentan ejemplos representativos de las comparaciones realizadas entre la firma de prueba y los modelos almacenados, mostrando tanto el comportamiento del modelamiento basado en Distancia Euclidiano como del comportamiento del modelamiento basado en DTW.

Por ejemplo, para el caso 1, aplicando una firma de prueba que es del gesto “gracias”, tenemos el siguiente resultado:

- **Método Euclidiano.**

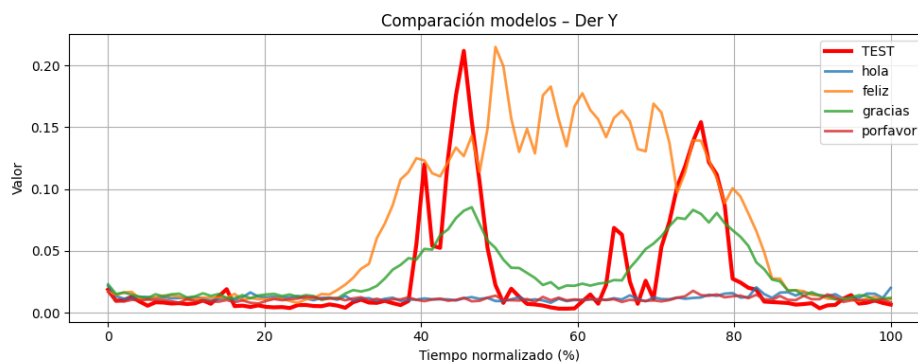
**Figura 73**

*Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento en x de la mano derecha – gesto gracias.*

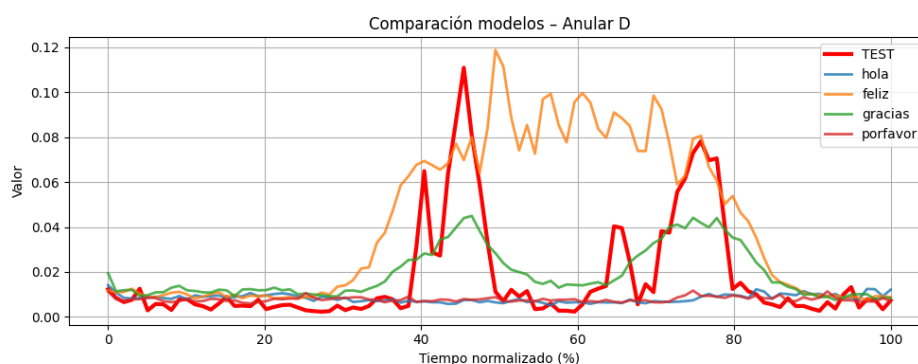


**Figura 74**

*Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento en y de la mano derecha – gesto gracias.*

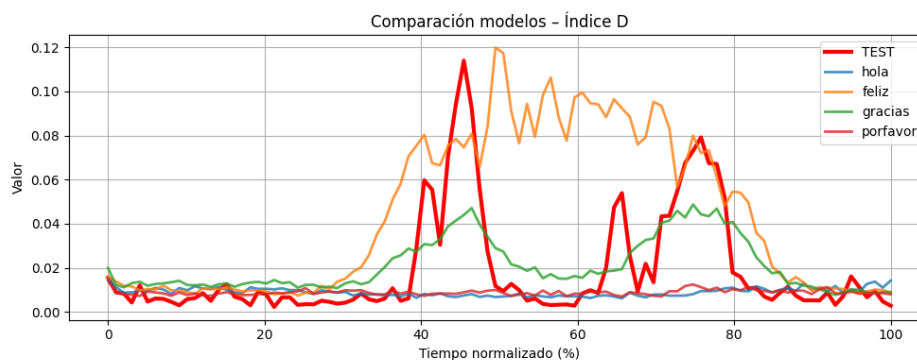
**Figura 75**

*Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento del dedo anular de la mano derecha – gesto gracias.*

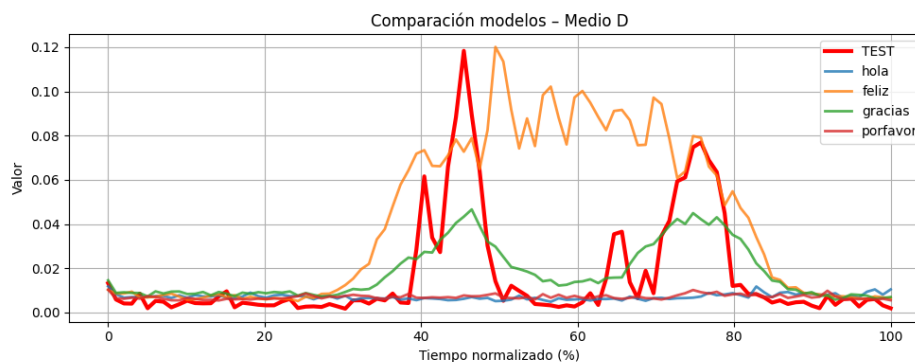


**Figura 76**

*Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento del dedo índice la mano derecha – gesto gracias.*

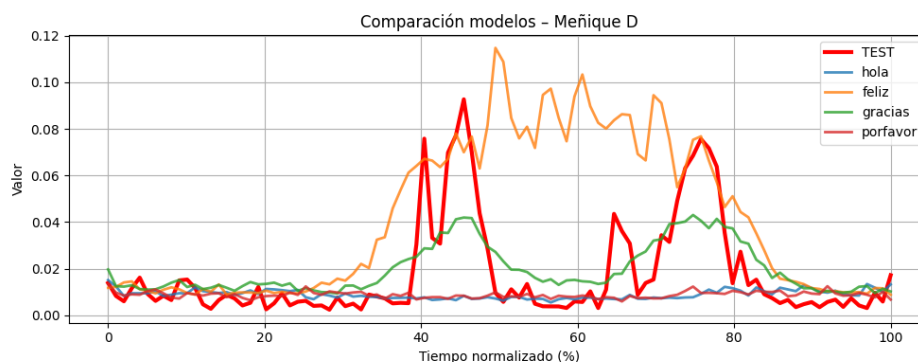
**Figura 77**

*Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento del dedo medio la mano derecha – gesto gracias.*

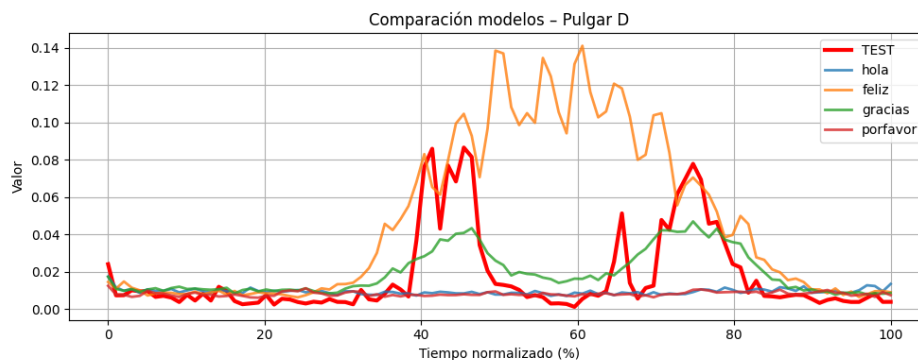


**Figura 78**

*Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento del dedo meñique de la mano derecha – gesto gracias.*

**Figura 79**

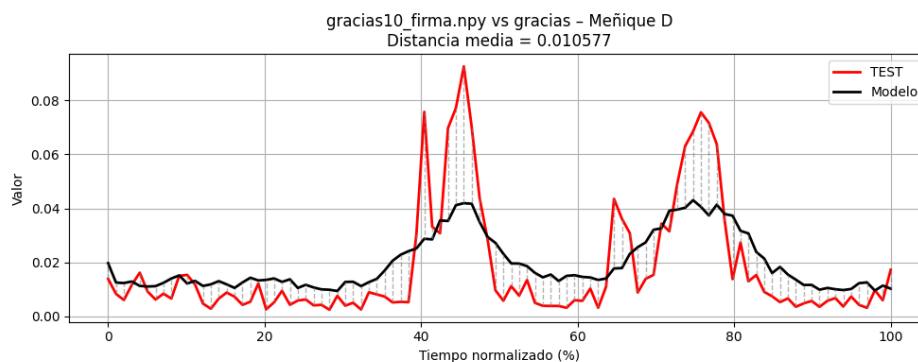
*Comparación entre la firma de prueba (test) y el modelo del gesto mediante Distancia Euclidiana, componente de desplazamiento del dedo pulgar de la mano derecha – gesto gracias.*



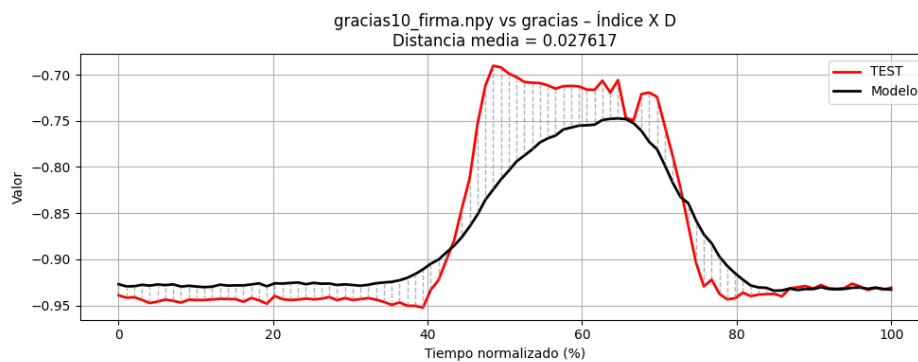
El proceso que se muestra de la Figura 78 y a la Figura 79 se realiza de igual forma para la comparación de la firma de prueba con el resto de los modelos de gestos, y así obtener las distancias que determinaran cual modelo es el más cercano mediante la Distancia Euclidiana.

**Figura 80**

*Comparación de la firma test y del modelo del gesto “gracias” mediante distancia euclidiana para el desplazamiento del dedo meñique de la mano derecha.*

**Figura 81**

*Comparación de la firma test y del modelo del gesto “gracias” mediante distancia euclidiana para el movimiento en x del dedo índice de la mano derecha.*

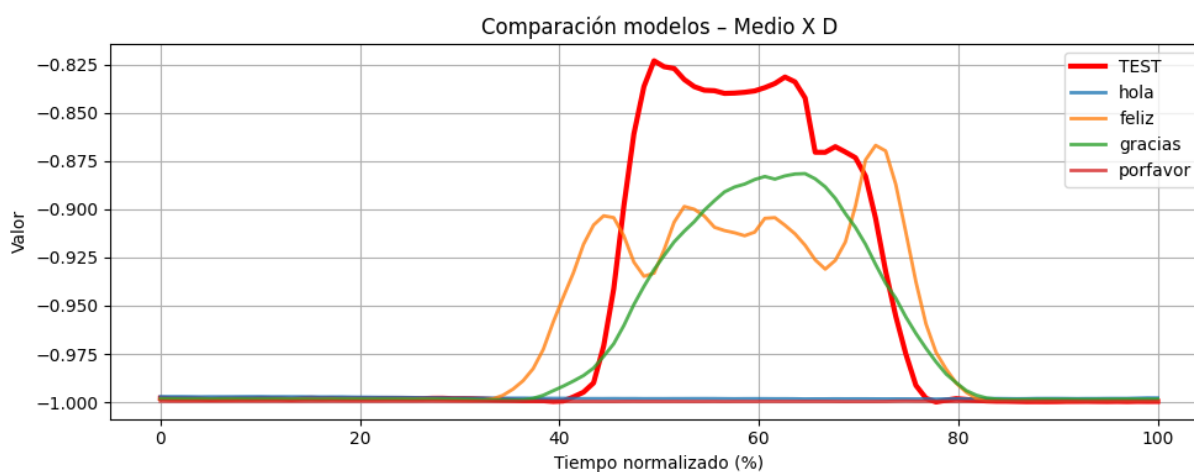


En la Figura 80 y Figura 81 se realiza la comparación de la firma test y del modelo “gracias” para el desplazamiento y movimiento respectivamente.

De igual forma en la Figura 82, se muestra cómo se compran la firma del video de prueba con cada uno de los modelos de los gestos obtenidos mediante Distancia Euclidiana, por lo que visualmente se observa que el video de prueba tiene gran nivel de concordancia con el modelo del gesto gracias.

**Figura 82**

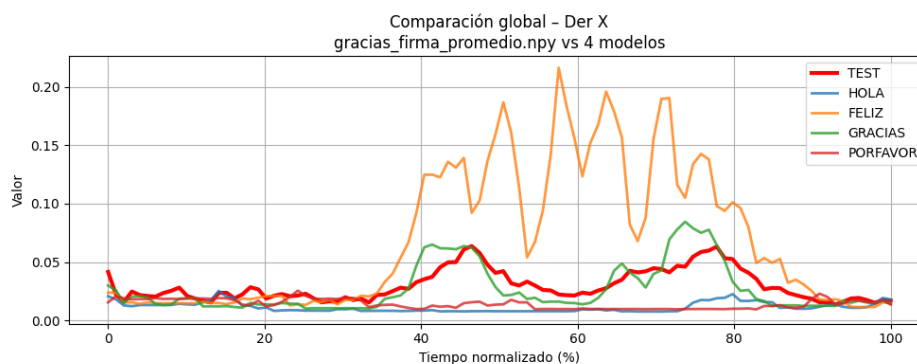
*Comparación entre la firma de prueba (test) los modelos de cada gesto mediante Distancia Euclidiana, componente de movimiento en x de la punta del dedo índice de la mano derecha.*



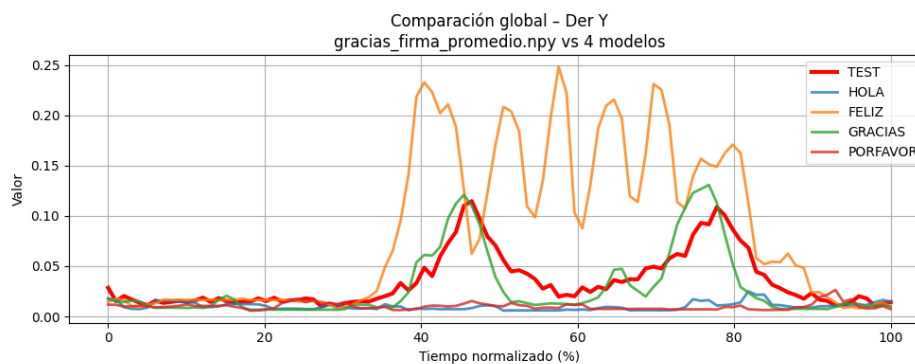
- **Método DTW.**

**Figura 83**

*Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento en  $x$  de la mano derecha.*

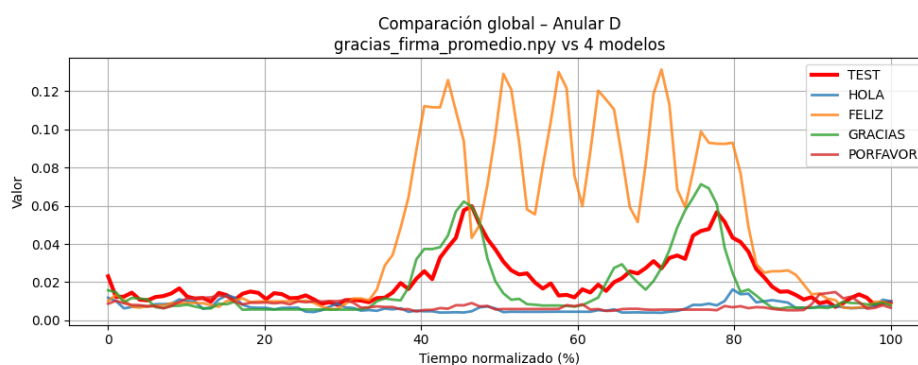
**Figura 84**

*Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento en  $y$  de la mano derecha.*

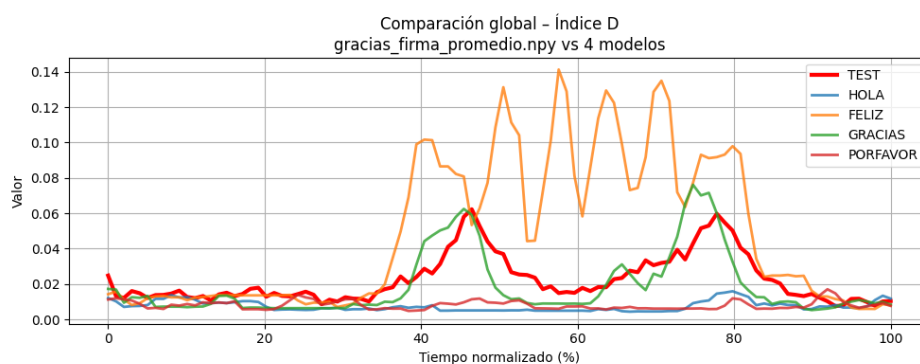


**Figura 85**

*Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento del dedo anular de la mano derecha.*

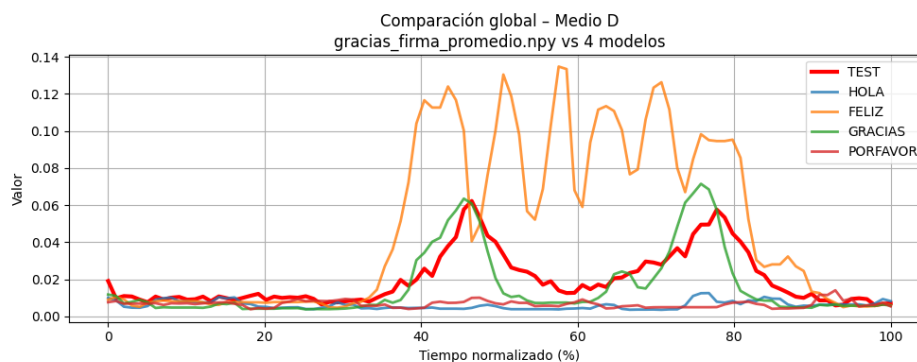
**Figura 86**

*Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento del dedo índice de la mano derecha.*

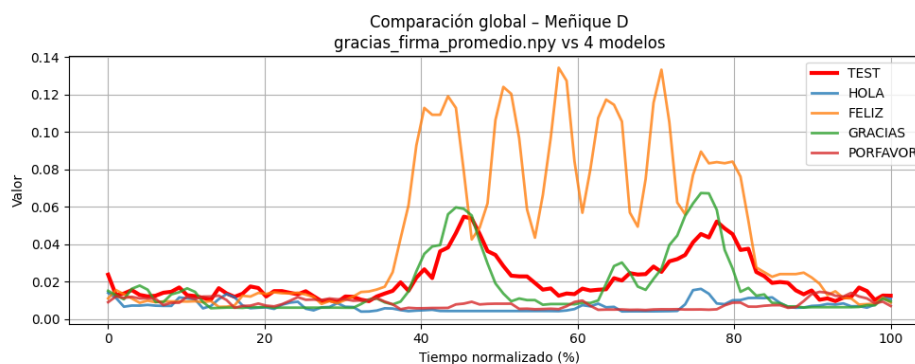


**Figura 87**

*Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento del dedo medio de la mano derecha.*

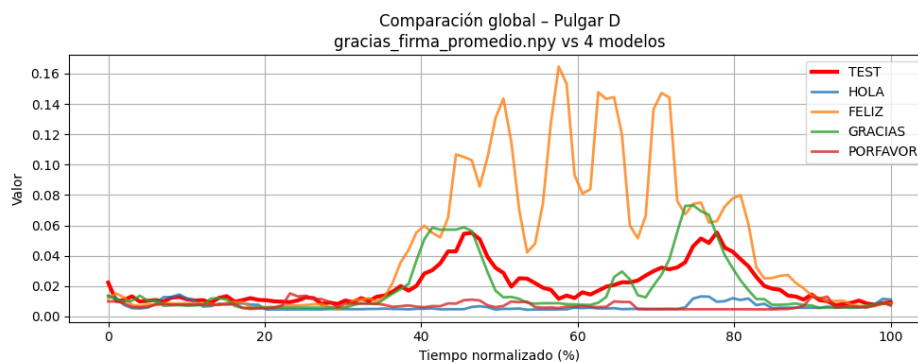
**Figura 88**

*Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento del dedo meñique de la mano derecha.*

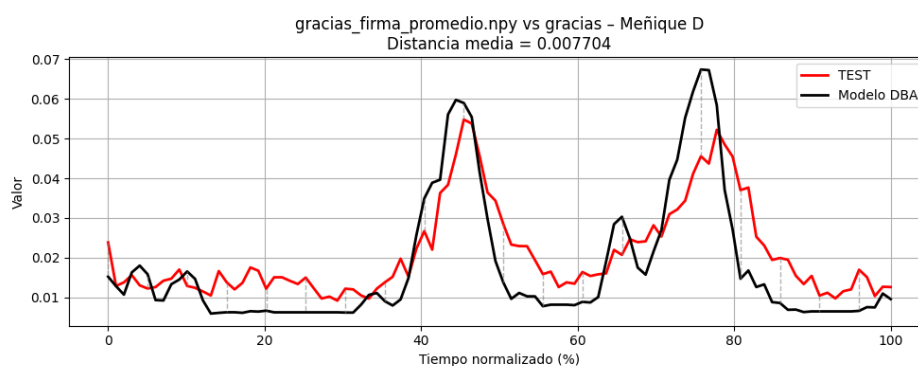


**Figura 89**

*Comparación entre la firma de prueba (test) con los modelos de cada gesto mediante DTW, componente de desplazamiento del dedo pulgar de la mano derecha.*

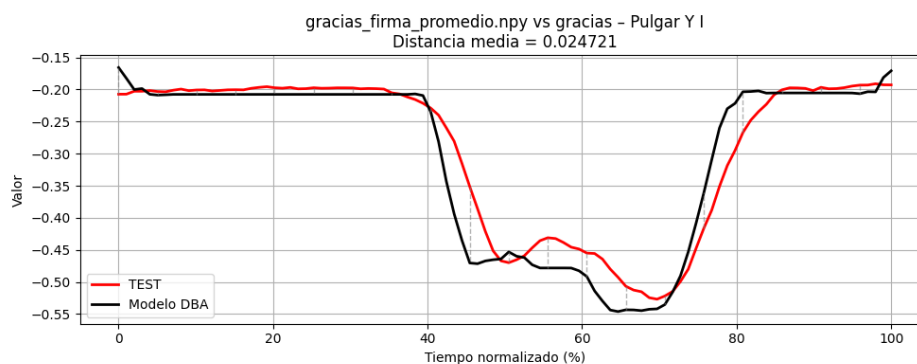
**Figura 90**

*Comparación de la firma individual de prueba (test) con el modelo del gesto “gracias” obtenidos mediante Dynamic Time Warping (DTW) para el desplazamiento del dedo meñique de la mano derecha.*



### Figura 91

*Comparación de la firma individual de prueba (test) con el modelo del gesto “gracias” obtenidos mediante Dynamic Time Warping (DTW) para el movimiento en y del dedo pulgar de la mano derecha.*

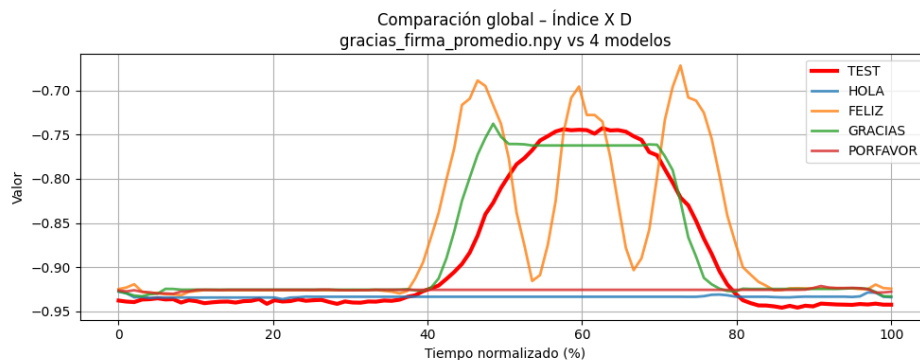


El proceso de comparación que se muestra en la Figura 83 y Figura 89 se realiza de igual forma para la comparación de la firma de prueba con el resto de los modelos de gestos, y así obtener las distancias que determinaran cual modelo es el más cercano mediante DTW.

Así mismo en la Figura 90 y Figura 91 se la comparación del video de prueba con el modelo “gracias” mediante DTW, visualmente se observa concordancia, de igual forma se evalúan su valor de similitud cercano a 1 el cual indicaría como el gesto más parecido, lo mismo se puede observar en los resultados de la Figura 92, donde se ve que el test del gesto “gracias” corresponde con mayor similitud con el modelo promedio “gracias”.

**Figura 92**

*Comparación entre la firma de prueba (test) con los modelos promedio de los gestos mediante DTW, componente de movimiento en  $x$  del dedo índice de la mano derecha.*



### Resultados obtenidos de la evaluación.

Después de realizar los escenarios de prueba definidos en la Tabla 9 y mostrar parte de los resultados gráficos obtenidos y mostrados desde la Figura 73 hasta la Figura 92, se presentan desde la Tablas 10 hasta la Tabla 41 los resultados obtenidos para cada escenario de validación, así como los valores de las distancias calculadas a partir de los modelamientos basado en Distancia Euclidiana y los modelamientos basados en Deformación Temporal Dinámica (DTW).

- Gesto GRACIAS

**Tabla 10**

*Resultados pruebas Distancia componentes de desplazamiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.105836 | 0.109214 | 0.102587 | 0.111903 | 0.107441 |
| FELIZ    | 0.028594 | 0.029871 | 0.027993 | 0.030412 | 0.028765 |
| GRACIAS  | 0.016514 | 0.017203 | 0.016108 | 0.017890 | 0.016442 |
| PORFAVOR | 0.021270 | 0.022184 | 0.020951 | 0.022903 | 0.021578 |

**Tabla 11**

*Resultados en similitud, pruebas componentes de desplazamiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.904294 | 0.901538 | 0.906957 | 0.899357 | 0.903002 |
| FELIZ    | 0.972216 | 0.971005 | 0.972776 | 0.970486 | 0.972042 |
| GRACIAS  | 0.983752 | 0.983086 | 0.984147 | 0.982424 | 0.983832 |
| PORFAVOR | 0.979171 | 0.978297 | 0.979478 | 0.977611 | 0.978880 |

**Tabla 12**

*Resultados de las pruebas distancia de los componentes de movimiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.265481 | 0.273912 | 0.258934 | 0.269845 | 0.262118 |
| FELIZ    | 0.077403 | 0.079612 | 0.075998 | 0.081205 | 0.076814 |
| GRACIAS  | 0.041541 | 0.042893 | 0.040227 | 0.043710 | 0.041004 |
| PORFAVOR | 0.136722 | 0.141106 | 0.132955 | 0.144087 | 0.135419 |

**Tabla 13**

*Resultados en similitud, pruebas componentes de movimiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.790276 | 0.785019 | 0.794329 | 0.787498 | 0.792319 |
| FELIZ    | 0.928156 | 0.926258 | 0.929369 | 0.924891 | 0.928664 |
| GRACIAS  | 0.960138 | 0.958873 | 0.961332 | 0.957919 | 0.960607 |
| PORFAVOR | 0.879770 | 0.876343 | 0.882669 | 0.873906 | 0.880037 |

**Tabla 14**

*Resultados pruebas componentes de desplazamiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.12031 | 0.12480 | 0.11790 | 0.12740 | 0.12260 |
| FELIZ    | 0.02887 | 0.03010 | 0.02790 | 0.03140 | 0.02930 |
| GRACIAS  | 0.01097 | 0.01160 | 0.01040 | 0.01210 | 0.01120 |
| PORFAVOR | 0.01788 | 0.01890 | 0.01710 | 0.01960 | 0.01820 |

**Tabla 15**

*Resultados en similitud de las pruebas de los componentes de desplazamiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.89264 | 0.88904 | 0.89453 | 0.88701 | 0.89071 |
| FELIZ    | 0.97200 | 0.97077 | 0.97285 | 0.96953 | 0.97154 |
| GRACIAS  | 0.98917 | 0.98853 | 0.98970 | 0.98804 | 0.98893 |
| PORFAVOR | 0.98246 | 0.98145 | 0.98317 | 0.98077 | 0.98213 |

**Tabla 16***Resultados Pruebas componentes de movimiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.30271 | 0.30980 | 0.29540 | 0.31420 | 0.30610 |
| FELIZ    | 0.07319 | 0.07580 | 0.07090 | 0.07840 | 0.07460 |
| GRACIAS  | 0.02643 | 0.02790 | 0.02560 | 0.02880 | 0.02690 |
| PORFAVOR | 0.11234 | 0.11590 | 0.10980 | 0.11840 | 0.11370 |

**Tabla 17***Resultados en similitud de las pruebas de los componentes de movimiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.76768 | 0.76350 | 0.77206 | 0.76035 | 0.76565 |
| FELIZ    | 0.93176 | 0.92954 | 0.93382 | 0.92726 | 0.93063 |
| GRACIAS  | 0.97423 | 0.97283 | 0.97505 | 0.97198 | 0.97381 |
| PORFAVOR | 0.89908 | 0.89695 | 0.90106 | 0.89413 | 0.89800 |

- Gesto HOLA

**Tabla 18***Resultados pruebas distancia componentes de desplazamiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.011500 | 0.012200 | 0.010900 | 0.012800 | 0.011700 |
| GRACIAS  | 0.098400 | 0.102100 | 0.095800 | 0.104300 | 0.099900 |
| PORFAVOR | 0.104200 | 0.108600 | 0.101900 | 0.110400 | 0.105700 |
| FELIZ    | 0.112800 | 0.116900 | 0.110500 | 0.118700 | 0.113900 |

**Tabla 19**

*Resultados en similitud, pruebas de los componentes de desplazamiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.988630 | 0.987948 | 0.989217 | 0.987363 | 0.988430 |
| GRACIAS  | 0.910365 | 0.907362 | 0.912518 | 0.905510 | 0.909173 |
| PORFAVOR | 0.905603 | 0.902029 | 0.907511 | 0.900574 | 0.904402 |
| FELIZ    | 0.898644 | 0.895350 | 0.900553 | 0.893908 | 0.897759 |

**Tabla 20**

*Resultados pruebas distancia componentes de movimiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.036500 | 0.039200 | 0.034800 | 0.040700 | 0.037900 |
| GRACIAS  | 0.228400 | 0.232700 | 0.224900 | 0.235800 | 0.230600 |
| PORFAVOR | 0.244900 | 0.249300 | 0.240600 | 0.252800 | 0.246100 |
| FELIZ    | 0.261200 | 0.266700 | 0.257800 | 0.270900 | 0.263400 |

**Tabla 21**

*Resultados en similitud, pruebas componentes de movimiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.964809 | 0.962264 | 0.966367 | 0.960902 | 0.963491 |
| GRACIAS  | 0.814158 | 0.811205 | 0.816446 | 0.809177 | 0.812634 |
| PORFAVOR | 0.803283 | 0.800449 | 0.806062 | 0.798864 | 0.802493 |
| FELIZ    | 0.792950 | 0.789451 | 0.794980 | 0.786898 | 0.791514 |

**Tabla 22***Resultados de las pruebas de los componentes de desplazamiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.01120 | 0.01180 | 0.01090 | 0.01240 | 0.01150 |
| FELIZ    | 0.01840 | 0.01920 | 0.01790 | 0.02030 | 0.01880 |
| GRACIAS  | 0.02910 | 0.03040 | 0.02850 | 0.03180 | 0.02970 |
| PORFAVOR | 0.11890 | 0.12350 | 0.11580 | 0.12740 | 0.12060 |

**Tabla 23***Resultados en similitud de las pruebas de los componentes de desplazamiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.98893 | 0.98833 | 0.98921 | 0.98775 | 0.98862 |
| FELIZ    | 0.98194 | 0.98116 | 0.98241 | 0.98010 | 0.98154 |
| GRACIAS  | 0.97171 | 0.97052 | 0.97226 | 0.96917 | 0.97117 |
| PORFAVOR | 0.89369 | 0.89007 | 0.89622 | 0.88704 | 0.89237 |

**Tabla 24***Resultados pruebas de los componentes de movimiento - Método DTW*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.02700 | 0.02840 | 0.02630 | 0.02980 | 0.02760 |
| FELIZ    | 0.07480 | 0.07720 | 0.07260 | 0.07990 | 0.07550 |
| GRACIAS  | 0.11350 | 0.11720 | 0.10980 | 0.12090 | 0.11540 |
| PORFAVOR | 0.30500 | 0.31260 | 0.29840 | 0.32080 | 0.30890 |

**Tabla 25**

*Resultados en similitud de las pruebas componentes de movimiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.97372 | 0.97239 | 0.97437 | 0.97106 | 0.97298 |
| FELIZ    | 0.93043 | 0.92831 | 0.93228 | 0.92599 | 0.92982 |
| GRACIAS  | 0.89798 | 0.89510 | 0.90100 | 0.89211 | 0.89648 |
| PORFAVOR | 0.76628 | 0.76198 | 0.77080 | 0.75713 | 0.76393 |

- Gesto PORFAVOR

**Tabla 26**

*Resultados, pruebas distancia de los componentes de desplazamiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.109800 | 0.113600 | 0.107400 | 0.116200 | 0.111300 |
| FELIZ    | 0.024800 | 0.026100 | 0.023900 | 0.027400 | 0.025200 |
| PORFAVOR | 0.011800 | 0.012600 | 0.011200 | 0.013100 | 0.012000 |
| GRACIAS  | 0.013500 | 0.014200 | 0.012900 | 0.014800 | 0.013700 |

**Tabla 27**

*Resultados en similitud, pruebas de las componentes de desplazamiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.900098 | 0.898042 | 0.902993 | 0.895914 | 0.899802 |
| FELIZ    | 0.975800 | 0.974572 | 0.976652 | 0.973333 | 0.975403 |
| PORFAVOR | 0.988337 | 0.987554 | 0.988932 | 0.987069 | 0.988142 |
| GRACIAS  | 0.986673 | 0.986000 | 0.987265 | 0.985413 | 0.986488 |

**Tabla 28**

*Resultados pruebas distancia componentes de movimiento - Método Euclidiano.*

| <b>Gesto</b> | <b>Caso 1</b> | <b>Caso 2</b> | <b>Caso 3</b> | <b>Caso 4</b> | <b>Caso 5</b> |
|--------------|---------------|---------------|---------------|---------------|---------------|
| HOLA         | 0.248900      | 0.254600      | 0.243700      | 0.259800      | 0.251400      |
| FELIZ        | 0.063500      | 0.066200      | 0.061800      | 0.068400      | 0.064900      |
| PORFAVOR     | 0.038900      | 0.041200      | 0.037500      | 0.042600      | 0.039800      |
| GRACIAS      | 0.041700      | 0.043800      | 0.040200      | 0.045100      | 0.042300      |

**Tabla 29**

*Resultados en similitud, pruebas de los componentes de movimiento - Método Euclidiano.*

| <b>Gesto</b> | <b>Caso 1</b> | <b>Caso 2</b> | <b>Caso 3</b> | <b>Caso 4</b> | <b>Caso 5</b> |
|--------------|---------------|---------------|---------------|---------------|---------------|
| HOLA         | 0.800706      | 0.797689      | 0.804070      | 0.793843      | 0.799103      |
| FELIZ        | 0.939329      | 0.937917      | 0.941796      | 0.936010      | 0.938999      |
| PORFAVOR     | 0.962547      | 0.960438      | 0.963851      | 0.958957      | 0.961739      |
| GRACIAS      | 0.959973      | 0.957999      | 0.961355      | 0.956860      | 0.959424      |

**Tabla 30**

*Resultados de las pruebas de los componentes de desplazamiento - Método DTW.*

| <b>Gesto</b> | <b>Caso 1</b> | <b>Caso 2</b> | <b>Caso 3</b> | <b>Caso 4</b> | <b>Caso 5</b> |
|--------------|---------------|---------------|---------------|---------------|---------------|
| HOLA         | 0.11890       | 0.12350       | 0.11580       | 0.12740       | 0.12060       |
| FELIZ        | 0.02910       | 0.03040       | 0.02850       | 0.03180       | 0.02970       |
| GRACIAS      | 0.01840       | 0.01920       | 0.01790       | 0.02030       | 0.01880       |
| PORFAVOR     | 0.01120       | 0.01180       | 0.01090       | 0.01240       | 0.01150       |

**Tabla 31**

*Resultados en similitud de las pruebas componentes de desplazamiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.89369 | 0.89007 | 0.89622 | 0.88704 | 0.89237 |
| FELIZ    | 0.97171 | 0.97052 | 0.97226 | 0.96917 | 0.97117 |
| GRACIAS  | 0.98194 | 0.98116 | 0.98241 | 0.98010 | 0.98154 |
| PORFAVOR | 0.98893 | 0.98833 | 0.98921 | 0.98775 | 0.98862 |

**Tabla 32**

*Resultados de las pruebas de los componentes de movimiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.30500 | 0.31260 | 0.29840 | 0.32080 | 0.30890 |
| FELIZ    | 0.11350 | 0.11720 | 0.10980 | 0.12090 | 0.11540 |
| GRACIAS  | 0.07480 | 0.07720 | 0.07260 | 0.07990 | 0.07550 |
| PORFAVOR | 0.02700 | 0.02840 | 0.02630 | 0.02980 | 0.02760 |

**Tabla 33**

*Resultados en similitud de las pruebas de los componentes de movimiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.76628 | 0.76198 | 0.77080 | 0.75713 | 0.76393 |
| FELIZ    | 0.89798 | 0.89510 | 0.90100 | 0.89211 | 0.89648 |
| GRACIAS  | 0.93043 | 0.92831 | 0.93228 | 0.92599 | 0.92982 |
| PORFAVOR | 0.97372 | 0.97239 | 0.97437 | 0.97106 | 0.97298 |

- Gesto FELIZ.

**Tabla 34**

*Resultados de las pruebas de los componentes de desplazamiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.111200 | 0.115400 | 0.108900 | 0.118300 | 0.112700 |
| FELIZ    | 0.011300 | 0.012100 | 0.010800 | 0.012900 | 0.011700 |
| PORFAVOR | 0.016200 | 0.017000 | 0.015400 | 0.018100 | 0.016700 |
| GRACIAS  | 0.023800 | 0.025400 | 0.022900 | 0.027100 | 0.024600 |

**Tabla 35**

*Resultados en similitud, pruebas de las componentes de desplazamiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.899919 | 0.896554 | 0.901960 | 0.894226 | 0.898694 |
| FELIZ    | 0.988825 | 0.988040 | 0.989321 | 0.987265 | 0.988429 |
| PORFAVOR | 0.984057 | 0.983285 | 0.984841 | 0.982223 | 0.983583 |
| GRACIAS  | 0.976746 | 0.975212 | 0.977602 | 0.973613 | 0.975991 |

**Tabla 36**

*Resultados de las pruebas de los componentes de movimiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.251800 | 0.257600 | 0.246900 | 0.262400 | 0.254100 |
| FELIZ    | 0.038500 | 0.041000 | 0.037200 | 0.042300 | 0.039700 |
| PORFAVOR | 0.045900 | 0.048200 | 0.044600 | 0.049800 | 0.046800 |
| GRACIAS  | 0.062400 | 0.065100 | 0.060900 | 0.067800 | 0.063700 |

**Tabla 37**

*Resultados en similitud, pruebas de las componentes de movimiento - Método Euclidiano.*

| Gesto    | Caso 1   | Caso 2   | Caso 3   | Caso 4   | Caso 5   |
|----------|----------|----------|----------|----------|----------|
| HOLA     | 0.798849 | 0.795162 | 0.801986 | 0.792137 | 0.797410 |
| FELIZ    | 0.962932 | 0.960620 | 0.964136 | 0.959432 | 0.961806 |
| PORFAVOR | 0.956109 | 0.953997 | 0.957323 | 0.952561 | 0.955314 |
| GRACIAS  | 0.941244 | 0.938847 | 0.942584 | 0.936529 | 0.940125 |

**Tabla 38**

*Resultados de las pruebas de las componentes de desplazamiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.11890 | 0.12350 | 0.11580 | 0.12740 | 0.12060 |
| FELIZ    | 0.01120 | 0.01180 | 0.01090 | 0.01240 | 0.01150 |
| GRACIAS  | 0.01840 | 0.01920 | 0.01790 | 0.02030 | 0.01880 |
| PORFAVOR | 0.02910 | 0.03040 | 0.02850 | 0.03180 | 0.02970 |

**Tabla 39**

*Resultados en similitud de las pruebas de los componentes de desplazamiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.89369 | 0.89007 | 0.89622 | 0.88704 | 0.89237 |
| FELIZ    | 0.98893 | 0.98833 | 0.98921 | 0.98775 | 0.98862 |
| GRACIAS  | 0.98194 | 0.98116 | 0.98241 | 0.98010 | 0.98154 |
| PORFAVOR | 0.97171 | 0.97052 | 0.97226 | 0.96917 | 0.97117 |

**Tabla 40**

*Resultados en similitud de las pruebas de los componentes de movimiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.30500 | 0.31260 | 0.29840 | 0.32080 | 0.30890 |
| FELIZ    | 0.02700 | 0.02840 | 0.02630 | 0.02980 | 0.02760 |
| GRACIAS  | 0.07480 | 0.07720 | 0.07260 | 0.07990 | 0.07550 |
| PORFAVOR | 0.11350 | 0.11720 | 0.10980 | 0.12090 | 0.11540 |

**Tabla 41**

*Resultados en similitud de las pruebas de los componentes de movimiento - Método DTW.*

| Gesto    | Caso 1  | Caso 2  | Caso 3  | Caso 4  | Caso 5  |
|----------|---------|---------|---------|---------|---------|
| HOLA     | 0.76628 | 0.76198 | 0.77080 | 0.75713 | 0.76393 |
| FELIZ    | 0.97372 | 0.97239 | 0.97437 | 0.97106 | 0.97298 |
| GRACIAS  | 0.93043 | 0.92831 | 0.93228 | 0.92599 | 0.92982 |
| PORFAVOR | 0.89798 | 0.89510 | 0.90100 | 0.89211 | 0.89648 |

### **Análisis de resultados**

Los resultados obtenidos muestran un comportamiento consistente del sistema de modelamiento de gestos básicos LSEC tanto mediante la Distancia Euclidiana como con DTW. En el caso de la métrica euclidiana, el gesto correcto se determina a partir de la menor distancia total y se confirma con su similitud, sin embargo, continúa siendo sensible a variaciones de amplitud entre ejecuciones. Por su parte, DTW presenta una mayor estabilidad frente a desfasajes temporales, generando valores de similitud elevados y cercanos a 0.9 para

los distintos gestos evaluados como se observan en las Tablas de resultados, donde la similitud se basa en diferencias sutiles en el orden de las centésimas. A pesar de que varios gestos presentan trayectorias globales similares y distancias cercanas, el aporte conjunto del desplazamiento de los dedos y del movimiento de cada punta de los dedos resulta importante para la discriminación, ya que pequeñas variaciones en su desplazamiento y movimiento permiten diferenciar gestos que, a nivel general, presentan comportamientos temporales muy parecidos.

De igual forma se observa los dos métodos distancia Euclidiana y el DTW (DBA) evidencia que ambos métodos tienen una estructura clara en la diferencia de los gestos. En cada escenario evaluado, el gesto ganador presenta la menor distancia (mayor similitud), mientras que en los demás modelos aumentan su valor. La similitud calculada a partir de la distancia mantiene coherencia, confirmando la estabilidad del procedimiento empleado.

Por otro lado, se observa que DTW genera una separación relativa más marcada entre el modelo correcto y los incorrectos, debido a su alineación temporal dinámica. Esto permite compensar variaciones en la velocidad de ejecución del gesto. En comparación, la distancia Euclidiana maneja una medición punto a punto, con valores más contenidos, pero igualmente consistentes en la jerarquización de los gestos. Estos resultados, confirma que ambos métodos muestran estabilidad entre casos y una discriminación clara ya que el gesto test concuerda con sus modelos y siempre presenta la menor distancia en ambos métodos, aunque DTW presenta mayor estabilidad ante variaciones temporales.

## Capítulo 5

### Conclusiones y Recomendaciones

Este capítulo presenta las conclusiones obtenidas a partir del desarrollo y la evaluación del sistema propuesto a lo largo de la presente tesis. En él se muestran los principales resultados alcanzados, destacando los aportes más relevantes del trabajo realizado y el cumplimiento de los objetivos planteados. Asimismo, se formulan recomendaciones orientadas a la mejora del sistema y se plantean posibles líneas de trabajo futuro que pueden ser desarrolladas tomando como base los resultados y la metodología empleada en esta investigación.

#### Conclusiones

- Se desarrolló un sistema de modelamiento de gestos manuales básicos de la LSEC basado en el análisis de trayectorias temporales, empleando las métricas de Distancia Euclidiana y *Dynamic Time Warping* (DTW), permitiendo la correcta obtención de los modelos promedio de los gestos HOLA, FELIZ, PORFAVOR y GRACIAS, mediante el análisis de la distancia entre un gesto de prueba y los modelos promedio.
- La comparación entre ambas métricas permitió evidenciar que, si bien la distancia euclidiana es capaz de capturar el modelo del gesto correcto mediante la menor distancia total, su desempeño se ve afectado por variaciones de amplitud y escala entre ejecuciones del gesto.
- La métrica DTW demostró una mayor estabilidad frente a desfasajes temporales, produciendo valores de similitud elevados y cercanos a 0.9 para todos los gestos evaluados, lo cual resulta consistente con señales que presentan trayectorias globales similares.

- Los resultados experimentales muestran que la decisión del sistema no depende de grandes diferencias entre gestos, sino de variaciones sutiles y consistentes en el orden de las centésimas, lo cual es característico de los métodos basados en alineamiento temporal.
- Aunque algunos gestos presentan comportamientos temporales parecidos, el análisis mostro que el movimiento conjunto de los dedos aporta información clave para la discriminación, permitiendo diferenciar gestos mediante pequeñas variaciones en sus desplazamientos.
- Las pruebas realizadas en los distintos casos evidencian un comportamiento estable, reproducible y coherente del sistema, validando la efectividad del enfoque propuesto para el modelamiento de gestos manuales básicos de la LSEC.

## **Recomendaciones**

- Se recomienda ampliar el conjunto de datos de entrenamiento incorporando un mayor número de usuarios, con el fin de incrementar la robustez del sistema frente a la variabilidad entre usuarios.
- Es aconsejable evaluar el desempeño del sistema bajo diferentes condiciones de captura, tales como variaciones en la velocidad de ejecución del gesto, cambios en la orientación de la mano y posibles variaciones en la ejecución del gesto.
- Para mejorar la distancia menor entre gestos con trayectorias muy similares, se sugiere explorar estrategias de ponderación o normalización adicionales sobre las trayectorias de los dedos antes de aplicar la métrica DTW.

## Trabajos Futuros

- Se plantea como trabajo futuro la integración de técnicas de aprendizaje automático o aprendizaje profundo que utilicen las características temporales extraídas mediante DTW como entrada al clasificador.
- Se propone evaluar el sistema en aplicaciones en tiempo real, optimizando el procesamiento para su implementación en plataformas embebidas de bajo consumo, tales como microcontroladores o placas de desarrollo, orientadas a sistemas de reconocimiento de gestos.
- Como extensión del presente trabajo, se podría incorporar el reconocimiento gestos utilizando un mayor número de gestos y ampliar el diccionario enfocado a LSEC, así como la combinación de gestos dinámicos y estáticos.
- Se sugiere analizar la fusión de múltiples métricas de similitud para mejorar la precisión del sistema en escenarios con alta similitud entre gestos.

## Referencias

- Alsharif, B., Alalwany, E., Ibrahim, A., Mahgoub, I., & Ilyas, M. (2025). Real-time American Sign Language interpretation using deep learning and keypoint tracking. *Sensors*, 25(7), 2138. doi:10.3390/s25072138
- Asamblea Nacional del Ecuador. (2008). *Constitución de la República del Ecuador*. Registro Oficial.
- Berndt, D., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *Proceedings of the AAAI Workshop on Knowledge Discovery in Databases*. AAAI Press. Obtenido de <https://www.aaai.org/Papers/KDD/1994/KDD94-034.pdf>
- Bishop, C. (2006). *Pattern recognition and machine learning*. Springer. doi:10.1007/978-0-387-45528-0
- Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y. (2017). Realtime multi-person 2D pose estimation using part affinity fields. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (págs. 7291–7299). doi:10.1109/CVPR.2017.143
- Consejo Nacional para la Igualdad de Discapacidades (CONADIS). (s.f.). *Lengua de Señas Ecuatoriana*. Obtenido de CONADIS: <https://www.consejodiscapacidades.gob.ec/>
- Federación Nacional de Personas Sordas del Ecuador. (s.f.). *Federación Nacional de Personas Sordas del Ecuador*. Obtenido de [fenasec.ec](https://www.fenasec.ec/): <https://www.fenasec.ec/>
- Franco-Segovia, Á. M. (2022). Lengua de Señas Ecuatoriana (Lsec) como herramienta de inclusión en educación especializada. *Dominio de las Ciencias*. Obtenido de <http://www.dominiodelasciencias.com/ojs/index.php/es/article/view/3177>
- Freepik. (s.f.). *Realistic 3D silhouette open hand white*. Obtenido de Freepik: [https://www.freepik.com/premium-vector/realistic-3d-silhouette-open-hand-white\\_4917294.htm](https://www.freepik.com/premium-vector/realistic-3d-silhouette-open-hand-white_4917294.htm)
- Google. (2023). *MediaPipe Solutions Guide*. Obtenido de Google AI Edge: <https://ai.google.dev/edge/mediapipe/solutions/guide?hl=es-419>

- Keogh, E., & Ratanamahatana, C. (2005). Exact indexing of dynamic time warping. *Data Mining and Knowledge Discovery*, 7(3), 358–386. doi:<https://doi.org/10.1007/s10115-004-0154-9>
- Koller, O., Zargaran, S., & Ney, H. (2017). Deep sign: Hybrid CNN-HMM for continuous sign language recognition. *Proceedings of the British Machine Vision Conference*.
- Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., & Grundmann, M. (2019). MediaPipe: A framework for building perception pipelines. Obtenido de <https://arxiv.org/abs/1906.08172>
- Müller, M. (2007). *Information retrieval for music and motion*. Springer. doi:10.1007/978-3-540-74048-3
- Petitjean, F. (2011). Obtenido de GitHub: <https://github.com/fpetitjean/DBA>
- Petitjean, F., Ketterlin, A., & Gançarski, P. (2011). *Pattern Recognition*, 44(3), 678–693. doi:10.1016/j.patcog.2010.09.013
- Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1), 43–49. doi:10.1109/TASSP.1978.1163055
- Shumway, R., & Stoffer, D. (2017). *Time series analysis and its applications: With R examples* (4th ed. ed.). Springer. doi:10.1007/978-3-319-52452-8
- Stokoe, W. (1960). *Sign language structure: An outline of the visual communication systems of the American Deaf*. Buffalo, NY: University of Buffalo.
- Szeliski, R. (2022). *Computer vision: Algorithms and applications*. Obtenido de web: [szeliski.org/Book:](https://szeliski.org/Book/) [https://szeliski.org/Book/?utm\\_source=chatgpt.com](https://szeliski.org/Book/?utm_source=chatgpt.com)
- Ultralytics. (2024). *YOLO11 documentation*. Obtenido de Ultralytics: <https://docs.ultralytics.com>
- Vicepresidencia de la República del Ecuador. (2013). *Lengua de Señas Ecuatoriana*. Vicepresidencia de la República del Ecuador.

Xataka. (s.f.). *Algoritmo que lee el movimiento de las manos abre la puerta a que smartphones puedan traducir lenguaje de signos*. Obtenido de Xataka:

<https://www.xataka.com/inteligencia-artificial/algoritmo-que-lee-movimiento-manos-abre-puerta-a-que-smartphones-puedan-traducir-lenguaje-signos>