Revista Facultad de Ingeniería (Rev. Fac. Ing.) Vol. 34 No. 71.
L-ISSN: 0121-1129, e-ISSN: 2357-5328.

DOI: 10.19053/01211129.v34.n71.2025.18059

RECONOCIMIENTO DE LA LENGUA DE SEÑAS COLOMBIANA MEDIANTE REDES NEURONALES CON MEMORIA A LARGO Y CORTO PLAZO

Recognition of Colombian Sign Language using Neural Networks with Long- and Short-Term Memory

Diego-Fernando Rivera-Vásquez
Universidad del Cauca, Popayán, Cauca.
diegoferivera@unicauca.edu.co

Carolina González-Serrano
Universidad del Cauca, Popayan, Cauca.
cgonzals@unicauca.edu.co

Fecha de recibido: 08-09-2024
Fecha de aceptado: 12-01-2025

RESUMEN

Este estudio explora el uso de redes neuronales de memoria larga a corto plazo para el reconocimiento del la lengua de señas colombiana. Abarca tanto señas estáticas (letras) como dinámicas (palabras). Los resultados muestran que el modelo alcanzó una precisión del 90 % en el reconocimiento de letras y del 82 % en palabras, y se logró identificar en promedio 27 señas independientes. Se analizaron, además, distintas estrategias de extracción de características espaciotemporales por medio de MediaPipe y se encontró que para detectar señas estáticas solo bastan los puntos de control de manos y para señas dinámicas se necesitan los puntos de control de manos y postura. Sin embargo, los resultados no superaron el 90 % de precisión alcanzado en estudios internacionales, lo que sugiere que la calidad y cantidad del conjunto de datos utilizado podría mejorarse. Como trabajo futuro, se plantea evaluar el desempeño del modelo en tiempo real, con el fin de facilitar la comunicación entre personas sordas y oyentes. También se recomienda explorar arquitecturas de aprendizaje profundo más avanzadas, como redes convolucionales de gráficos, redes neuronales *transformer* o combinaciones de red neuronal convolucional con memoria larga a corto plazo; estas han mostrado buenos resultados en el reconocimiento de signos dinámicos.

Palabras clave: aprendizaje profundo; interprete de lengua de señas; lengua de señas; reconocimiento de la lengua de señas; red neuronal artificial; visión artificial.

ABSTRACT

This study explores the use of long short-term memory neural networks for the recognition of Colombian sign language. It covers both static (letters) and dynamic (word) signs. The results show that the model achieved 90% accuracy in letter recognition and 82% in words, identifying an average of 27 independent signs. In addition, different strategies for extracting spatiotemporal features using MediaPipe were analyzed, and it was found that to detect static signs only the hand control points are sufficient, and for dynamic signs the hand and posture control points are needed. However, the results did not exceed the 90% accuracy achieved in international studies, suggesting that the quality and quantity of the data set used could be improved. As future work, it is proposed that the model's performance be evaluated in real time to facilitate communication between deaf and hearing people. It is also recommended to explore more advanced deep learning architectures, such as

Cómo citar este artículo: D. F. Rivera-Vásquez, y C. González-Serrano, "Reconocimiento de la lengua de señas colombiana mediante redes neuronales con memoria a largo y corto plazo". *Revista Facultad de Ingeniería*, vol. 34, no. 71, e18059, 2025. https://doi.org/10.19053/01211129.v34.n71.2025.18059

graph convolutional networks, transformer neural networks, or combinations of convolutional neural networks with long short-term memory, as these have shown good results in dynamic sign recognition.

Keywords: artificial neural network; artificial vision; deep learning; sign language; sign language interpreter; sign language recognition.

RECONHECIMENTO DA LÍNGUA DE SINAIS COLOMBIANA POR MEIO DE REDES NEURAIS COM MEMÓRIA DE LONGO E CURTO PRAZO

RESUMO

Este estudo explora o uso de redes neurais com memória de longo e curto prazo (LSTM) para o reconhecimento da Língua de Sinais Colombiana. Abrange tanto sinais estáticos (letras) quanto sinais dinâmicos (palavras). Os resultados mostram que o modelo atingiu uma precisão de 90% no reconhecimento de letras e 82% no reconhecimento de palavras, identificando em média 27 sinais distintos. Foram analisadas diferentes estratégias de extração de características espaço-temporais por meio do MediaPipe, observando-se que, para detectar sinais estáticos, bastam os pontos de controle das mãos, enquanto que para sinais dinâmicos são necessários os pontos de controle das mãos e da postura corporal. No entanto, os resultados não superaram a precisão de 90% alcançada em estudos internacionais, sugerindo que a qualidade e a quantidade do conjunto de dados utilizado podem ser aprimoradas. Como trabalho futuro, propõe-se avaliar o desempenho do modelo em tempo real, a fim de facilitar a comunicação entre pessoas surdas e ouvintes. Recomenda-se também explorar arquiteturas de aprendizado profundo mais avançadas, como redes convolucionais de grafos, redes neurais do tipo transformer ou combinações entre redes convolucionais e LSTM, que têm demonstrado bons resultados no reconhecimento de sinais dinâmicos.

P**alavras-chave:** aprendizado profundo; intérprete de linguagem de sinais; linguagem de sinais; reconhecimento de linguagem de sinais; rede neural artificial; visão computacional.

1. INTRODUCCIÓN

La sordera o discapacidad auditiva es una de las condiciones más prevalentes a nivel mundial; afecta alrededor de 430 millones de personas, lo que corresponde al 5 % de la población mundial. La sordera se puede presentar en tres niveles en función de la cantidad de pérdida auditiva que una persona experimenta: leve, moderado y grave o profundo [1]. Las personas que se encuentran en nivel leve o moderado pueden apoyarse en herramientas como audífonos y amplificadores para aumentar su umbral de decibeles y discriminar los sonidos de su entorno, situación que si es detectada a tiempo les permitirá aprender un idioma nativo para comunicarse [2]. Por su parte, la condición grave o profunda limita a las personas en cuanto a la interacción y comunicación con su entorno, al no poder desarrollar el habla como medio de comunicación, por lo cual adoptan la lengua de señas y logran comunicarse únicamente con personas que lo manejen. Esto dificulta en gran medida el acceso a educación, trabajo y participación ciudadana, escenarios en los que la comunicación directa con los demás es fundamental [3].

Para mitigar lo anterior, se han realizado investigaciones desde el área de la computación sobre reconocimiento automático de señas que permitan disminuir brechas de comunicación entre personas sordas y oyentes [4]. En el trabajo de Morillas-Espejo y Martínez-Martin [5], se describe el desarrollo de un sistema que facilita la comunicación entre una persona sorda y una oyente, por medio de una red neuronal convolucional (CNN) que reconoce el alfabeto de España, es decir, letras o caracteres independientes del castellano.

En Colombia, diferentes estudios, como el de Flórez et al. [6], presentan la implementación de redes neuronales *transformer* (TNN, por sus siglas del inglés *transformer neural network*) para reconocimiento

de señas colombiana que representan palabras. Otras estrategias estudiadas se enfocan en la memoria larga a corto plazo (LSTM, por sus siglas del inglés *long short-term memory*) para el reconocimiento del alfabeto colombiano y proponen un modelo que identifica diez palabras básicas de interacción del la lengua de señas colombiana: {*"Hola"*, *"Yo"*, *"Nombre"*, *"Buenos"*, *"Años"*, *"Gustar"*, *"Tardes"*, *"Noches"*, *"Licor"*, *"Días"*} [7, 8]. Este modelo aporta de manera significativa a la problemática; sin embargo, es una solución limitada en cuanto a la cantidad de señas. En él se estudian los tipos de señas por separado, es decir, alfabeto o palabras, y por lo general cada estudio requiere la creación de conjuntos de datos propios que se ajusten a su contexto y necesidades para entrenar las estrategias de inteligencia artificial que aporten en el desarrollo de intérpretes de la lengua de señas.

Con base en lo anterior, se evidencia la necesidad de indagar y proponer marcos experimentales, estrategias, métodos y modelos que faciliten el reconocimiento de señas colombianas —estáticas (letras) y dinámicas (palabras)—, con el fin de evaluar la capacidad de estas para identificar cada tipo de seña.

Miah et al. [7] proponen una estrategia de inteligencia artificial para extraer información contextual espaciotemporal. El proceso de experimentación se realizó con conjuntos de datos de gran escala, como el WLASL [8] junto a uno propio, el cual registró 3000 videos de 30 señas diferentes. Para todos los conjuntos de datos, se aplicó la estrategia de extracción de puntos de control por medio de la herramienta MediaPipe, con la cual se extrajeron 67 puntos de control por fotograma y se creó una secuencia de 20 fotogramas por video, con el fin de entrenar su estrategia y calcular la precisión en pruebas de laboratorio. Para el WLASL, se obtuvieron resultados de precisión entre el 34,41 y el 63,25 %, lo que supera a investigaciones previas. Es importante resaltar que para su contexto los autores evidenciaron una precisión del 99,75 %. Por su parte, Ihsan et al. [9] describen MediSing, un modelo híbrido CNN-LSTM bidireccional (CNN-BiLSTM) para la clasificación de señas del contexto médico. Este utiliza capas convolucionales para la extracción de características y BiLSTM para procesar secuencias de fotogramas, lo que permite reconocer 30 señas dinámicas del la lengua de señas americana. Los resultados experimentales evidenciaron una precisión del 95,83 %. Otros estudios, como los de Shin et al. [10], se enfocan en extraer características espaciotemporales y de pixeles, implementando técnicas de transferencia de conocimiento junto a una ResNet101 para la detección del alfabeto y detección de 77 señas dinámicas por medio de la estrategia de redes convolucionales de gráficos (GCN, por sus siglas del inglés *graph convolutional network*). Los estudios evidenciaron una precisión entre el 99,87 y el 100 %. Por su parte, estudios como el de Shanableh [11] no solo reconocen letras y palabras, sino también oraciones. En ese trabajo, se implementa una estrategia llamada *imagen en movimiento*, la cual permite representar cada seña en una sola imagen, que se concatena luego con cada palabra de la oración para así construir la secuencia de palabras en señas que posteriormente es procesada por una arquitectura BiLSTM. Esta última permite hacer el reconocimiento de las señas a lo largo de una oración y logra reconocer 40 oraciones diferentes compuestas por 80 señas de palabras arábicas, con una precisión de hasta el 97,3 %.

A nivel nacional, se encuentran estudios en los que se implementan redes CNN [12], para procesar y extraer información de fotografías de señas estáticas del lenguaje colombiano, con una precisión del 93,3 %. Otro estudio relevante es el de Barrero [13], en el que se prueban estrategias de extracción de características espaciotemporales con MediaPipe para la detección del alfabeto colombiano, incluidas las letras S y Z, que son dinámicas, por medio de una red neuronal LSTM; la precisión es del 80 %. Flórez et al. [6] aplican estrategias de tipo transformadores o TNN para la detección de señas dinámicas (palabras) colombianas, con una precisión del 90 %. Se tiene en cuenta que su innovación principal es considerar diferentes enfoques para cada seña, y que su conjunto de datos está basado en la captura de cuatro videos simultáneos de una persona realizando un signo en diferentes ángulos, con el fin de

evaluar si la extracción de características multimodales aporta a la precisión en la detección de señas. Por último, se encuentra la implementación de una arquitectura híbrida entre una red VGG16+LSTM para la detección de diez señas dinámicas (palabras) y se evidencia una precisión del 76 % [14].

Es importante destacar que los estudios a nivel nacional procesan y entrenan sus estrategias utilizando únicamente un tipo de señas, ya sean estáticos o dinámicos. Como resultado, no se tiene certeza de si los modelos propuestos pueden reconocer ambos tipos de señas con la misma efectividad.

2. METODOLOGÍA

Para el desarrollo del presente estudio, se siguió un proceso metodológico de validación de arquitecturas para detección de señas en contextos específicos, similar a la utilizada en investigaciones previas [7], que consideran diferentes etapas:

- **Obtención del conjunto de datos.** Se busca un conjunto de datos adecuado o construirlo por medio de videos, teniendo en cuenta que se requiere trabajar con señas dinámicas y estáticas.
- **Extracción de características.** Se procesan los videos de los señas de interés y se extraen las características necesarias para el modelo o estrategia seleccionada. Para este estudio, se utiliza la extracción de características espaciotemporales con MediaPipe.
- **Entrenamiento de la estrategia.** Se entrena la arquitectura o estrategia de interés junto a los datos de estudio.
- **Validación de la estrategia.** Se valida el desempeño del modelo entrenado utilizando un conjunto de datos de prueba y se calcula la métrica de "precisión", para determinar si el modelo se ajusta o no al contexto definido.

En cada una de las etapas descritas se estudia si un modelo basado en arquitectura LSM permite la detección de señas colombianas tanto estáticas como dinámicas y cómo afecta la cantidad puntos de control de MediaPipe en la detección de señas estáticas y dinámicas.

2.1 Conjunto de datos

El conjunto de datos utilizado correspondió al conjunto de señas básicos de la lengua de señas colombiana disponible en el curso online [15]. Se manejaron los enfoques estáticos y dinámicos, y se obtuvieron dos conjuntos de datos. En la Tabla 1 se presentan las señas de cada conjunto de datos.

Tabla 1. Signos colombianos de estudio

Conjuntos señas estáticas (letras)	Conjunto señas dinámicas (palabras)
A, B, C, D, E, F, G, H, I, J, K, L, M, N, Ñ, O, P, Q, R, RR, S, T, U, V, W, X, Y, Z	Sordo, Hola, Bien, Mal, Adiós, Bienvenido, Gracias, Perdón, Permiso, Yo, Tu, El, Ella, Nosotros, Usted, Ustedes, Que, Cuando, Donde, Como, Quien, Cuento, Cual, Buenos días, Buenas tardes, Buenas noches, Por favor, Como estas.

Al no contar con conjuntos de datos públicos, se construyó uno propio similar al del estudio de Miah et al. [7], con base en la captura del video hasta la extracción de características, como se describe a continuación:

- Captura de señas por medio de videos de 5 segundos.
- Extracción de 30 fotogramas secuenciales por cada video.
- Extracción de puntos de control con MediaPipe para cada fotograma.
- Etiquetado de las secuencias de puntos de control.

Cabe resaltar que MediaPipe permite capturar puntos de control de manos, postura y rasgos faciales. Para el presente estudio, los puntos de control seleccionados fueron los de manos y postura, como se ven en la Figura 1.

Con el fin de validar la estrategia de construcción del conjunto de datos, se procesó diferente cantidad de videos y extracción de puntos de control de manos, solo en conjunto de datos de letras y palabras, además de un conjunto de datos con puntos de control de manos-postura. Esto con el fin de determinar cómo afecta la cantidad de puntos de control al modelo propuesto; así, se obtuvieron los siguientes conjuntos de datos (Tabla 2).

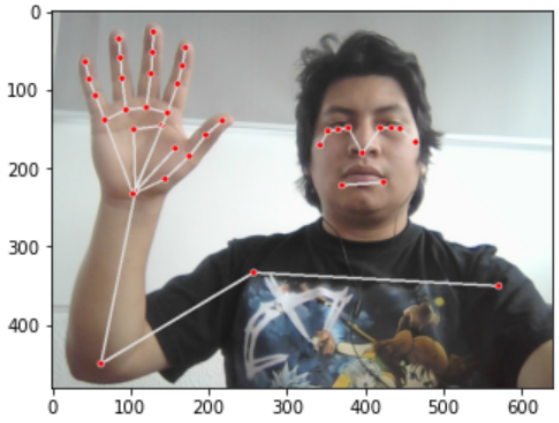

Figura 1. *Puntos de control de manos y postura MediaPipe.*

Tabla 2. Características del conjunto de datos

Nombre conjunto de datos	N.º de señas	N.º de videos por seña	N.º de puntos MediaPipe
Letras	27	30	126
Palabras V1	19	30	126
Palabras V2	28	20	258

2.2 Modelo

El modelo propuesto está basado en la integración de tres capas LSTM y tres capas densas distribuidas de la siguiente manera (Tabla 3).

Tabla 3. Arquitectura propuesta

Capa	Tipo de capa	N.º de neuronas	Función de activación
Entrada	LSTM	64	ReLu
Intermedia 1	LSTM	128	ReLu
Intermedia 2	LSTM	64	ReLu
Intermedia 3	Densa	64	ReLu
Intermedia 4	Densa	32	ReLu
Salida	Densa	N.º de signos dataset	Sotfmax

Se configuró desde el entorno de Python con los siguientes hiperparametros: 1) el optimizador "Adam", 2) la función de pérdida "categorical crossentropy", 3) la métrica "accuracy" y 4) el tamaño de datos de validación del 20 % del total de datos de cada conjunto de datos.

3. RESULTADOS

El modelo propuesto fue entrenado con cada uno de los conjuntos de datos presente en la Tabla 2. A continuación, se presentan los resultados de cada entrenamiento. En la Figura 2, se puede ver cómo se comportan el entrenamiento y la validación para el conjunto de datos de letras compuesto por el abecedario de la lengua de señas colombiana, el cual maneja los puntos de control de las manos. Este entrenamiento obtiene una precisión del 90 %, con 500 épocas de entrenamiento, pero logra estabilizarse a partir de las 100 épocas.

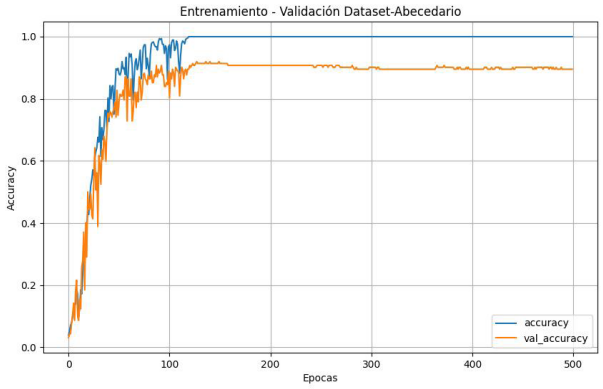

Figura 2. *Resultado del conjunto de datos (letras).*

En la Figura 3, se puede ver cómo se comporta el entrenamiento del modelo utilizando el conjunto de datos compuesto por 19 palabras, llamado Palabras V1. Este también cuenta solo con puntos de control de MediaPipe de manos. Como se puede observar, tanto en la fase de entrenamiento como en la de validación, el modelo no es capaz de superar un 82 % de precisión, incluso cuando procesa un menor número de señas que el anterior; además, no se observa estabilidad de optimización durante el proceso.

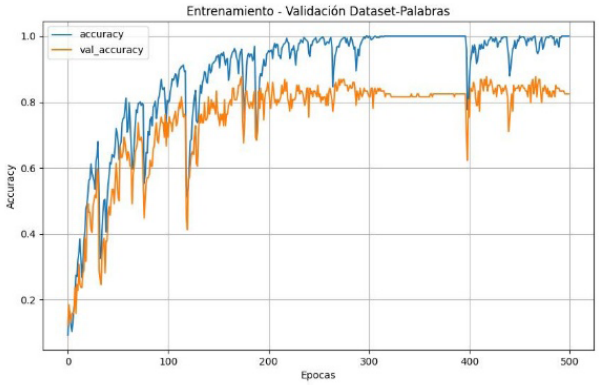

Figura 3. *Resultado del conjunto de datos (Palabras V1).*

En la Figura 4, se presentan los resultados del entrenamiento con un conjunto de datos compuesto por 28 señas de palabras, pero en este caso solo se procesan 20 videos por signo, a diferencia de los conjuntos anteriores, en los que son 30 videos por seña. En este caso, se extrajeron puntos de control de la mano y la postura corporal. Este entrenamiento obtuvo mejores resultados que el anterior, ya que con solo 300 épocas se alcanzó un 82 % de precisión. Además, su optimización y capacidad de generalización se estabilizaron desde la época 160. Se logró, por tanto, generalizar una mayor cantidad de señas que en el entrenamiento anterior.

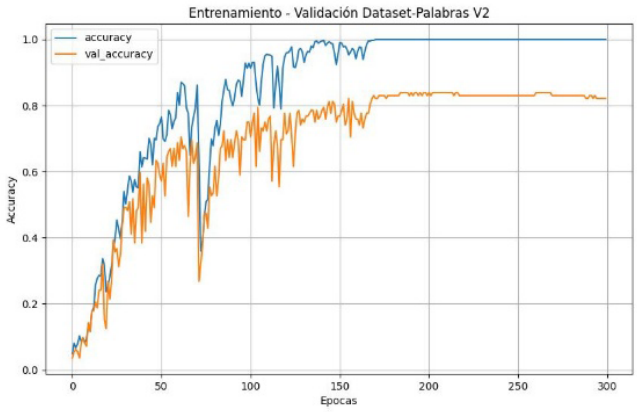

Figura 4. *Resultado del conjunto de datos (Palabras V2).*

En la Tabla 4, se presenta un resumen de los resultados mencionados y se observa que, para las señas dinámicas, el modelo propuesto supera el 80 % de precisión. Esto sugiere la necesidad de contar con una mayor cantidad de datos de prueba o explorar la integración de arquitecturas con mayor profundidad y su combinación con CNN. Asimismo, se podría considerar la integración de estrategias como las reportadas en el trabajo de Miah et al. [7], las cuales han demostrado ser eficaces en el procesamiento

de grandes conjuntos de datos de señas, y logran un desempeño más sólido en contextos similares. Todo lo relacionado con el entrenamiento y creación del conjunto de datos de este estudio se encuentra disponible en la página de GitHub [16].

Tabla 4. *Resultados del entrenamiento del modelo LSTM*

Conjunto de datos	N.º de señas	Épocas	Precisión del entrenamiento	Precisión de validación
Letras	27	500	1	0,9
Palabras V1	19	500	1	0,82
Palabras V2	28	300	1	0,82

4. DISCUSIÓN Y CONCLUSIONES

Este estudio permitió validar que las redes neuronales LSTM pueden ser usadas para proponer modelos de aprendizaje profundo enfocados en el reconocimiento de la lengua de señas colombiana, tanto en señas estáticas (letras) como dinámicas (palabras).

Los resultados evidencian que una misma arquitectura logró una precisión del 90 % en el reconocimiento de letras y del 82 % en palabras; se alcanzó en promedio la identificación de 27 señas independientes.

El análisis de aproximaciones previas permitió identificar la estrategia más efectiva para la extracción de características espaciotemporales, y se destaca el uso de MediaPipe para la extracción de puntos de control de manos y postura en videos. Sin embargo, el modelo no logró superar el 90 % de precisión, como sí lo han reportado estudios internacionales en contextos similares. Esta diferencia se atribuye a las limitaciones del conjunto de datos utilizado, lo que evidencia la necesidad de considerar estrategias más robustas que permitan obtener un conjunto de datos de calidad. Disponer de un conjunto de datos más amplio y diverso permitirá al modelo identificar un mayor número de señas independientes y mejorar su rendimiento.

Como trabajo futuro, se considera fundamental evaluar el desempeño del modelo en tiempo real, con el objetivo de facilitar la identificación y el reconocimiento de señas colombianas. Esto contribuiría a mejorar la comunicación entre personas sordas y oyentes, y reducirá la brecha existente en la transmisión de la información.

Adicionalmente, se hace necesario explorar estrategias avanzadas de modelado en aprendizaje profundo, como las GCN, TNN o las combinaciones de CNN-LSTM. Estas arquitecturas han demostrado ser altamente eficaces en la detección de señas dinámicas, con la obtención de resultados prometedores en investigaciones recientes.

CONTRIBUCIÓN DE AUTORÍA

Diego-Fernando Rivera-Vásquez: Investigación, conceptualización, análisis formal, experimentación, escritura y edición.

Carolina González-Serrano: Conceptualización, proceso metodológico, investigación, escritura, supervisión, revisión y edición.

AGRADECIMIENTOS

Los autores agradecen al Grupo de Investigación en Inteligencia Computacional (GICO) y al programa de Maestría en Computación de la Universidad del Cauca por su apoyo en la orientación y supervisión de este trabajo.

REFERENCIAS

[1] WHO, *Sordera y pérdida de la audición*, 2024. https://www.who.int/es/news-room/fact-sheets/detail/deafness-and-hearing-loss

[2] M. R. Paradinas, J. Alexander, S. Suárez, T. Rivera Rodríguez, *Libro Virtual de Formación en Otorrinolaringología*, 2021. https://www.udocz.com/apuntes/978613/libro-virtual-de-formacion-en-otorrinolaringologia

[3] Y. Tatiana *et al.*, *Análisis sobre los procesos de enseñanza - aprendizaje, implementados para Personas con Discapacidad Auditiva y Visual en la Educación Superior que contribuyan a la creación de una estrategia educativa incluyente: Una revisión sistemática de literatura con la ventana temporal del 2017 al 2022*, Tesis de Grado, Universidad Industrial de Santander. 2023. https://noesis.uis.edu.co/handle/20.500.14071/14808

[4] B. Joksimoski *et al.*, "Technological Solutions for Sign Language Recognition: A Scoping Review of Research Trends, Challenges, and Opportunities," *IEEE Access*, vol. 10, pp. 40979-40998, 2022. https://doi.org/10.1109/ACCESS.2022.3161440

[5] F. Morillas-Espejo E. Martinez-Martin, "Sign4all: A Low-Cost Application for Deaf People Communication," *IEEE Access*, vol. 11, pp. 98776–98786, 2023. https://doi.org/10.1109/ACCESS.2023.3312636 z

[6] A. Flórez, *Colombian sign language analysis and recognition*, Tesis de Grado, Universidad de los Andes, 2022. http://hdl.handle.net/1992/64165

[7] A. S. M. Miah, M. A. M. Hasan, S. Nishimura, J. Shin, "Sign Language Recognition Using Graph and General Deep Neural Network Based on Large Scale Dataset," *IEEE Access*, vol. 12, pp. 34553-34569, 2024. https://doi.org/10.1109/ACCESS.2024.3372425

[8] D. Li, C. Rodriguez Opazo, X. Yu, H. Li, *Word-level Deep Sign Language Recognition from Video: A New Large-scale Dataset and Methods Comparison*, 2020. https://dxli94.github.io/

[9] M. A. Ihsan, A. F. Eram, L. Nahar, M. A. Kadir, "MediSign: An Attention-Based CNN-BiLSTM Approach of Classifying Word Level Signs for Patient-Doctor Interaction in Hearing Impaired Community," *IEEE Access*, vol. 12, pp. 33803-33815, 2024. https://doi.org/10.1109/ACCESS.2024.3370684

[10] J. Shin, A. S. M. Miah, Y. Akiba, K. Hirooka, N. Hassan, Y. S. Hwang, "Korean Sign Language Alphabet Recognition through the Integration of Handcrafted and Deep Learning-Based Two-Stream Feature Extraction Approach," *IEEE Access*, vol. 12, pp. 68303-68318, 2024. https://doi.org/10.1109/ACCESS.2024.3399839

[11] T. Shanableh, "Two-Stage Deep Learning Solution for Continuous Arabic Sign Language Recognition Using Word Count Prediction and Motion Images," *IEEE Access*, vol. 11, pp. 126823-126833, 2023. https://doi.org/10.1109/ACCESS.2023.3332250.

[12] C. J. Da *et al.*, *Aprendizaje automático de lengua de señas colombiana CIS2210CP03*, 2022.

[13] J. G. Barrero, *Sistema Experto para la Identificación de Gestos del Lenguaje de Señas Colombiano*, Tesis de Grado, Universidad Industrial de Santander, 2022. https://noesis.uis.edu.co/handle/20.500.14071/11302

[14] J. A. Muñoz-Galindez, R. Vargas-Cañas, "Modelo de interpretación de lengua de señas colombiano usando inteligencia artificial," *Revista de Investigación, Desarrollo e Innovación*, vol. 13, no. 2, pp. 357-366, Aug. 2023. https://doi.org/10.19053/20278306.V13.N2.2023.16840

[15] INSOR, *Diccionario*, 2025. https://educativo.insor.gov.co/diccionario/

[16] GitHub, *FerchoRV/LSTM-Reconocimiento-de-signos-colombianos: Experimentaciones de reconocimiento de lenguaje de señas colombianos implementando una red neuronal LSTM*. 2024. https://github.com/FerchoRV/LSTM-Reconocimiento-de-signos-colombianos