



# Reconocimiento del lenguaje de señas usando técnicas de aprendizaje profundo

*Sign language recognition using Deep Learning techniques*

ALEJANDRO MEDINA REYES<sup>a</sup> , BORIS JESÚS MEDEROS MADRAZO<sup>b\*</sup> , JOSÉ MANUEL MEJÍA MUÑOZ<sup>a</sup> 

<sup>a</sup>{Maestría en Ingeniería Eléctrica, Departamento de Ingeniería Eléctrica y Computación, ...}, <sup>b</sup>{Departamento de Física y Matemáticas, ...}, ... Instituto de Ingeniería y Tecnología, Universidad Autónoma de Ciudad Juárez, México

\*Autor de correspondencia. Correo electrónico: boris.mederos@uacj.mx

<b>N.º de resumen</b> 9CP25-5	<b>Formato</b> Resumen
<b>Tema</b> Diseño y rediseño del producto	<b>Presentador</b> Alejandro Medina Reyes
<b>Fecha de la presentación</b> Mayo 20, 2025	<b>Estatus</b> Estudio terminado

## Resumen

En este trabajo se explora el uso de diversas técnicas de aprendizaje profundo para la tarea del reconocimiento del lenguaje de señas. Durante esta investigación se trabajó con el conjunto de datos AUTSL, se probaron arquitecturas de redes neuronales convolucionales 2D+1 para la detección de señas, con imágenes con y sin fondo. También se exploró el uso de diversas arquitecturas de redes neuronales con los datos de articulaciones, que fueron extraídos de los vídeos usando el modelo PoseNet. Tras entrenar diversas redes, se probaron ensambles de modelos, combinando 20 redes del mismo tipo. Los resultados mostraron un incremento en la exactitud con técnicas de ensamble; con imágenes como entrada de una red se logró una precisión del 88.9 % y con datos de articulaciones con arquitecturas simples compuestas de convoluciones, mecanismos de atención y redes densas se alcanzó una precisión del 88.32 %. El mejor resultado, de 95.45 %, se logró al utilizar los datos como grafos a lo largo del tiempo con redes convolucionales de grafos. Los resultados muestran que el uso de datos de articulaciones son una entrada valiosa para la tarea del reconocimiento del lenguaje de señas, ya que proveen una mejor generalización al descartar detalles que no son relevantes para la predicción, como puede ser el fondo del vídeo, color de piel del sujeto, color de cabello, etc. Otra ventaja de esta entrada es la reducción de dimensionalidad, lo que permite un tiempo de entrenamiento e inferencia más cortos.

**Palabras clave:** prueba eléctrica *in-circuit*; circuitos de interfaz directa; aprendizaje automático.

## Abstract

This work explores the use of various deep learning techniques for the task of sign language recognition. Working with the AUTSL dataset, this research tested 2D+1 convolutional neural network architectures for sign detection, using images with and without backgrounds. The use of various neural network architectures was also explored with joint data, which were extracted from videos using the PoseNet model. After training various networks, ensemble models were tested, combining 20 networks of the same type. The results showed an increase in accuracy with ensemble techniques; with images as input to a network, an accuracy of 88.9% was achieved, and with joint data, simple architectures composed of convolutions, attention mechanisms, and dense networks achieved an accuracy of 88.32%. The best result, 95.45%, was achieved by using the data as graphs over time with graph convolutional networks. The results show that articulation data is a valuable input for sign language recognition, as it provides improved generalization by discarding details that are not relevant to pre-



diction, such as the video background, the subject's skin color, hair color, etc. Another advantage of this input is dimensionality reduction, which allows for shorter training and inference times.

**Keywords:** in-circuit test, direct interface circuits; machine learning.

**Entidad legal responsable del estudio**

Universidad Autónoma de Ciudad Juárez

**Financiamiento**

Beca SECIHTI, CVU núm. 886626.

**Conflictos de interés**

Los autores declaran que no hay conflicto de intereses.