Universidad Iberoamericana Puebla

Repositorio Institucional **http://repositorio.iberopuebla.mx**

ExpoIbero Primavera 2025

Desarrollo de un sistema de inteligencia artificial para la traducción de la lengua de señas mexicana (LSM)

Islas Guzmán, Luis Humberto

2025

https://hdl.handle.net/20.500.11777/6194

http://repositorio.iberopuebla.mx/licencia.pdf

Desarrollo de un Sistema de Inteligencia Artificial para la traducción de la Lengua de Señas Mexicana (LSM)

Islas Guzmán Luis Humberto (Décimo semestre en Ingeniería Biomédica)[1], Pedro Jaramillo Andrea (séptimo semestre en Ingeniería Biomédica) [1, *], Moreno Hernández Ana (profesor responsable) [1]
Universidad Iberoamericana Puebla, San Andrés Cholula, Puebla, México

Palabras clave: Azure ML., discapacidad auditiva, Lengua de Señas Mexicana,
[*]**Autor Corresponsal:** andrea.pedro@iberopuebla.mx

Introducción

La comunicación es una necesidad fundamental en la sociedad, y para las personas con discapacidad auditiva, la lengua de señas representa una solución visual-gestual eficiente. Sin embargo, la falta de conocimiento generalizado sobre este lenguaje limita seriamente la inclusión. Ante esta barrera, los avances tecnológicos han propuesto soluciones innovadoras como los guantes traductores que permiten convertir señas en texto o voz en tiempo real.
En investigaciones recientes, se han desarrollado diversos prototipos de guantes inteligentes capaces de traducir el alfabeto de la lengua de señas mediante sensores de flexión y unidades de medición inercial (IMU) [1], [2]. Estos proyectos suelen combinar hardware portátil con algoritmos de aprendizaje automático que pueden ejecutarse localmente o a través de plataformas en la nube. En este contexto, servicios como Microsoft Azure Machine Learning ofrecen herramientas robustas para entrenar, evaluar e implementar modelos de clasificación de señales sensorizadas [3].
El presente proyecto propone un dispositivo portátil basado en un guante inteligente que traduce gestos de la Lengua de Señas Mexicana (LSM) a texto. A través del uso de sensores de movimiento, microcontroladores, y un sistema de inteligencia artificial alojado en la nube, se busca crear una herramienta ligera, eficiente y accesible, capaz de facilitar la comunicación para personas con discapacidad auditiva.
El propósito principal es desarrollar una herramienta accesible, ligera y eficaz que ayude a las personas con discapacidad auditiva, facilitando la conversión de sus gestos en texto entendible para cualquier individuo, y en etapas futuras, incluso transformarlo en voz a través de tecnología de texto a voz.

Metodología

El avance del guante inteligente se organizó en diversas etapas que van desde la captura física de los gestos hasta su interpretación a través de inteligencia artificial, incluyendo la transmisión en la nube y la representación de los resultados en una aplicación final. Se creó un sistema de captura que utilizó sensores flexibles ZD10-100 (flexores de resistencias) instalados en cada dedo de la mano, junto con un sensor de movimiento MPU6050 que brindó datos adicionales acerca de la dirección y aceleración de la mano. Estos sensores se vincularon a un microcontrolador Xiao ESP32S3, como se observa en la figura 1 (ver anexo), seleccionado por su tamaño reducido y su habilidad para comunicarse vía WiFi integrada [4].

El guante fue diseñado con el objetivo de que los sensores registraran simultáneamente la flexión de los dedos y la inclinación de la mano cada vez que el usuario ejecutaba un gesto de una letra. Una vez implementado el sistema de captura, la etapa subsiguiente implicó el manejo local de la información en el microcontrolador, el cual fue programado para interpretar las señales analógicas y digitales de los sensores, generando vectores de atributos que representaban cada gesto. El uso de sensores flexores y unidades de medición inercial (IMU) para capturar gestos manuales ha sido ampliamente explorado en aplicaciones de traducción de lenguas de señas. Estudios como los de Makkar et al. [1] y Salmankhah et al. [2] han mostrado que estos datos pueden convertirse en vectores temporales, interpretados con precisión por modelos de aprendizaje automático.

Posteriormente, una etapa crítica del sistema consistió en la comunicación con la nube. El ESP32S3 se conecta a Internet vía Wi-Fi y transmite los vectores a un punto de enlace HTTP POST en Microsoft Azure Machine Learning, previamente configurado con un modelo entrenado con ejemplos etiquetados del alfabeto mexicano. En este caso, se utilizó una red neuronal multicategoría (Multiclass Neural Network), elegida por su capacidad para manejar tareas de clasificación con múltiples clases (27 letras) y su buen desempeño en la identificación de patrones complejos no lineales. Cada vez que se recibe un nuevo vector, el modelo realiza una inferencia y devuelve como resultado la letra correspondiente, alineada con el gesto registrado.

Este tipo de arquitectura basada en la nube ha sido propuesta en investigaciones como la de Singh et al. [4], que destacan las ventajas de combinar microcontroladores livianos con plataformas como Azure, permitiendo un sistema distribuido y eficiente. Azure facilita la escalabilidad, mantenimiento y supervisión del modelo de machine learning.

Finalmente, para permitir una interacción fluida, se desarrolló una interfaz gráfica ligera que muestra en pantalla la letra interpretada en tiempo real y utiliza tecnología de texto a voz para emitir el sonido correspondiente. Esta aplicación acumula cada letra reconocida para formar palabras completas, mejorando la comprensión por parte del interlocutor oyente. Sistemas similares de retroalimentación se han explorado en propuestas como DeepASL [3], donde la experiencia del usuario es fundamental para la aceptación de la herramienta.

Resultados y Discusión

Los hallazgos del desarrollo del guante inteligente evidenciaron un rendimiento altamente satisfactorio en la interpretación de gestos asociados a las letras del alfabeto mexicano. El sistema, que utiliza sensores de flexión ZD10-100 en cada dedo y un sensor de movimiento MPU6050, logró una captura precisa de los movimientos, transmitiendo los datos al microcontrolador Xiao ESP32S3 para su procesamiento y posterior envío a la nube.

En las pruebas controladas, el guante alcanzó una precisión media del 87% con una desviación estándar de ±2%. Se observó un mejor desempeño en la ejecución de gestos a velocidad moderada, mientras que, en gestos realizados de manera precipitada, la precisión disminuyó hasta un 83%, debido a la captura parcial o incompleta de los vectores de flexión y movimiento. Esto indica que la velocidad de ejecución es un factor crítico en el rendimiento del sistema. El tamaño de la mano de los diferentes usuarios no mostró un impacto significativo en el desempeño, ya que al aplicar la formula de desviación estandar, las variaciones en precisión se mantuvieron dentro de un rango menor al 3%, lo cual sugiere una buena capacidad de adaptación individual del guante.

A continuación, en la Tabla 1 se muestra la precisión promedio alcanzada bajo diferentes condiciones de ejecución, destacándose una mayor exactitud en gestos moderados y una ligera disminución en gestos rápidos.

Tabla 1. Precisión alcanzada dependiendo de la condición de ejecución

Condición de ejecución	Precisión alcanzada
Gestos moderados	87%
Gestos rápidos	83%
Variación entre usuarios	±2.5%

De manera complementaria, en la Tabla 2 se presentan las principales métricas de validación del modelo, donde se observa que el modelo mantuvo un Recall (qué tan bien el modelo identifica las letras) de 86% y un F1-Score (promedio de Precisión y Recall) de 86.5% sobre un total de 27 clases correspondientes a las letras del alfabeto mexicano. Cada clase fue representada por 110 ejemplos para su entrenamiento.

Tabla 2. Valores porcentuales de cada métrica

Métrica	Valor obtenido
Precisión general	87%
Recall	86%
F1-Score	86.5%
Número de clases	27
Ejemplos por clase	110

Conclusiones

El presente trabajo desarrolló un sistema biomédico basado en un guante inteligente para el reconocimiento de letras en Lengua de Señas Mexicana (LSM). El prototipo alcanzó una precisión promedio del 87% mediante el uso de sensores físicos y un modelo de red neuronal multicategoría entrenado y desplegado en Microsoft Azure Machine Learning. Su rendimiento fue consistente entre distintos usuarios y condiciones de ejecución, lo que evidencia su fiabilidad como dispositivo de asistencia. Se identifican como mejoras futuras la ampliación hacia reconocimiento de frases completas y la implementación de procesamiento local para mayor autonomía. La integración de sensorial, aprendizaje automático y conectividad en un dispositivo portátil representa una propuesta innovadora dentro del campo de la ingeniería biomédica.

Referencias

[1] C. E. Cruz Mendoza, “Translator Glove for Deaf-Mute Individuals,” Fab Academy 2024, Fab Lab Puebla, Puebla, México, 2024. [En línea]. Disponible en: https://fabacademy.org/2024/labs/puebla/students/carlos-cruz/Project.html

[2] A. Makkar, "SignSpeak: Open-Source Time Series Classification for ASL Translation," arXiv preprint arXiv:2407.12020, 2024. [En línea]. Disponible en: https://arxiv.org/abs/2407.12020

[3] S. Salmankhah y M. Rajabi, "PenSLR: Persian End-to-End Sign Language Recognition Using Ensembling," arXiv preprint arXiv:2406.16388, 2024. [En línea]. Disponible en: https://arxiv.org/abs/2406.16388

[4] B. Fang, J. Co y M. Zhang, "DeepASL: Enabling Ubiquitous and Non-Intrusive Word and Sentence-Level Sign Language Translation," en Proceedings of the 15th ACM Conference on Embedded Networked Sensor Systems, 2017, pp. 1–14. [En línea]. Disponible en: https://dl.acm.org/doi/10.1145/3131672.3131693

[5] A. P. Singh y S. K. Singh, "A New Architecture for Hand-Worn Sign Language to Speech Translator," arXiv preprint arXiv:2009.03988, 2020. [En línea]. Disponible en: https://arxiv.org/abs/2009.03988

[6] M. C. Aldrete, Gramática de la lengua de señas mexicana. México: El Colegio de México, 2008.

Anexo

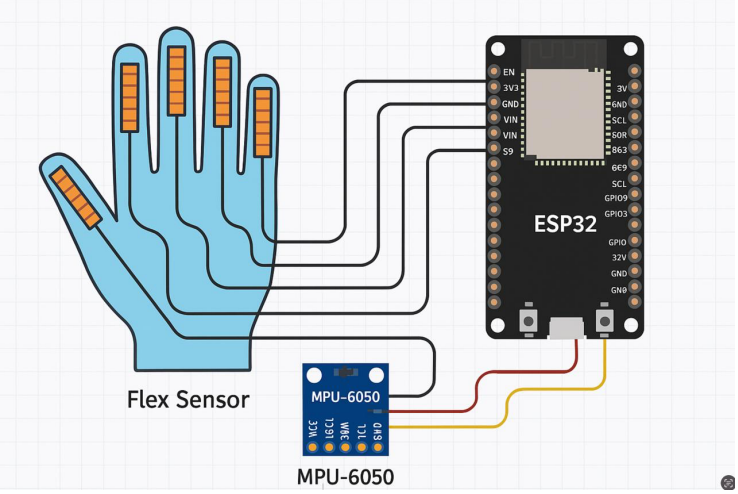

Fig. 1. Diagrama de conexiones