

Representatio Horarum



UNIVERSIDAD
SERGIO ARBOLEDA

Daniel Camilo Bernal Ternera

Escuela de ciencias exactas e ingeniería, Universidad Sergio Arboleda, Bogotá, Colombia

daniel.bernal01@usa.edu.co

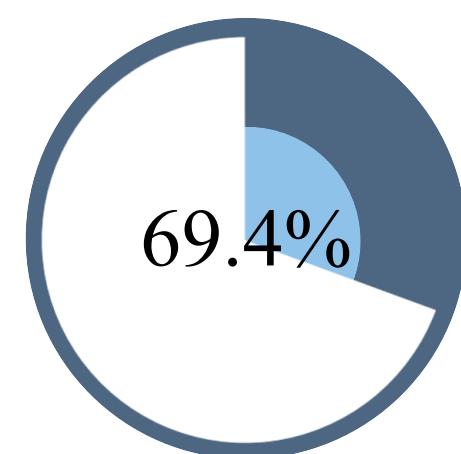
El lenguaje de señas es un sistema lingüístico complejo con gramática, vocabulario y sintaxis propios. Donde cada uno depende del contexto donde se desarrolla, permeándose de la cultura y el contexto social donde se origina.^[1]

Dando lugar a más de 200 lenguajes de señas alrededor del mundo, cada uno con sus variaciones.

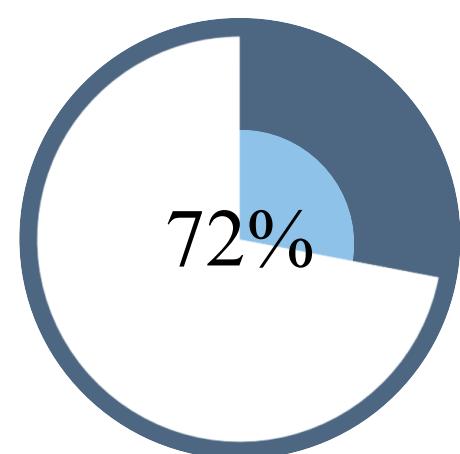
Durante los años se han estudiado múltiples soluciones, pero ninguna definitiva. Y con el auge de tecnologías emergentes, la IA se ve como parte de la respuesta a innovar y derribar las barreras.^[3]

Justificación

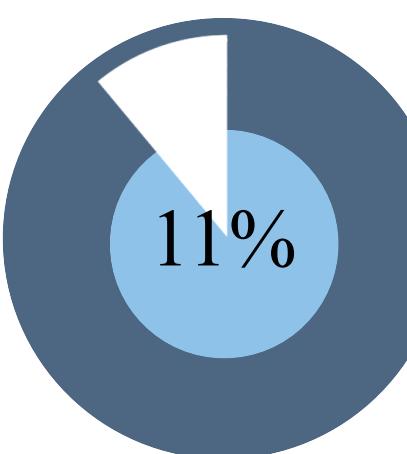
Royal National Institute for Deaf



Pierden más de la mitad de la información^[2]



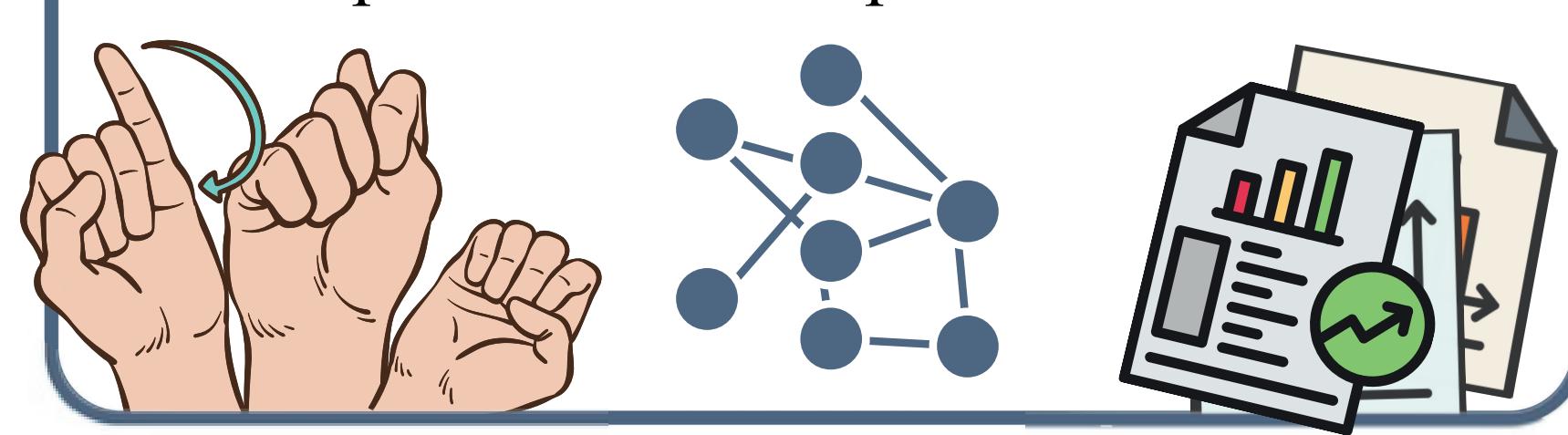
Detecta hostilidad de parte del personal^[2]



Tuvo un acceso equitativo al sistema^[2]

Problema de Investigación

¿En qué medida puede un autoencoder, a través de determinadas técnicas, entrenado con tres lenguajes de señas y evaluado con recursos determinados, ayudar a identificar patrones útiles para representar palabras en un espacio latente?

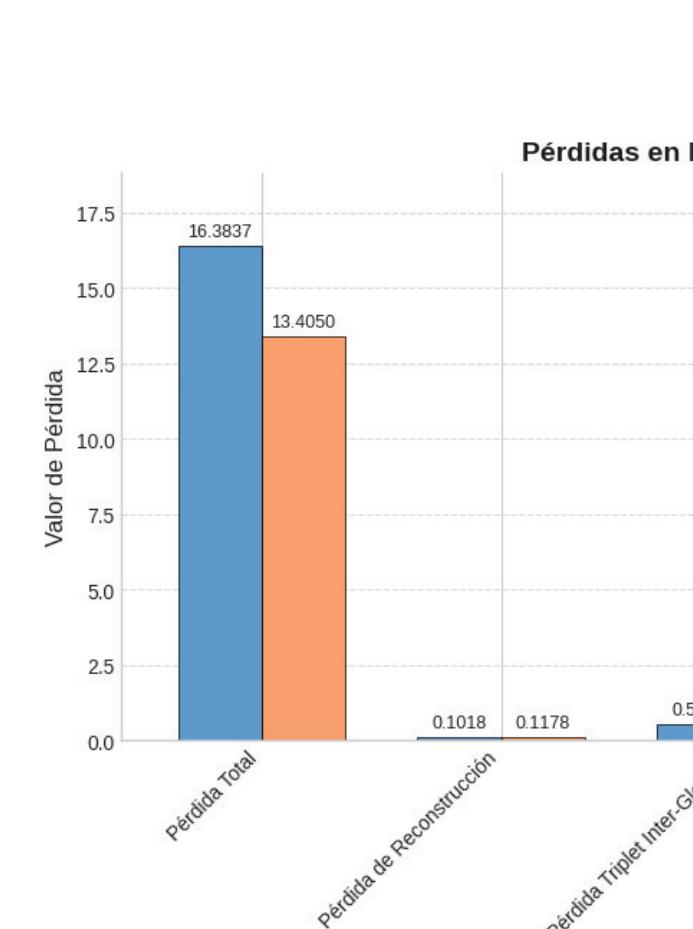


WLSL - 2Gloss



Ambos experimentos tienen un muy buen desempeño en todas las métricas, las pérdidas son un poco altas debido al margen establecido.

Se comparan las diferentes pérdidas de entrenamiento y validación con 3 baselines distintos, con el fin de evaluar de manera uniforme el rendimiento de cada uno de los experimentos.

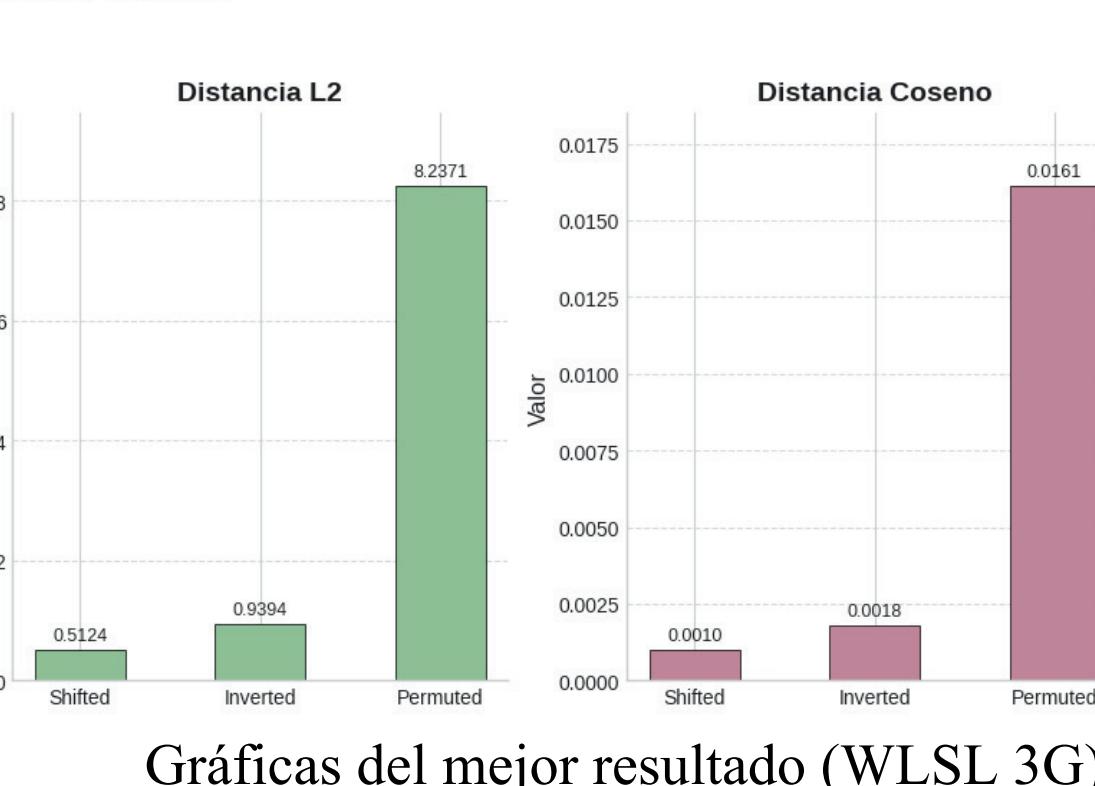


WLSL - 3Gloss



Las pérdidas empiezan a desmejorar.

Análisis de Métricas y Distancias del Modelo



ISL - 2Gloss



El margen se vuelve cada vez más difícil de superar.

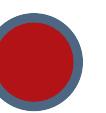
Gráficas del mejor resultado (WLSL 3G)

SLOVO - 2Gloss



El margen se vuelve cada vez más difícil de superar.

ISL - 3Gloss



No es capaz de separar las secuencias INV y SHIFT.

LOS3 - 2Gloss



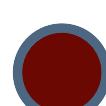
El modelo encuentra un atajo donde empieza a separar todas las secuencias.

LOS3 - 3Gloss



Utiliza el atajo separando todas las secuencias sin generalizar ningún patrón en los datos.

SLOVO - 3Gloss



Rompe la jerarquía temporal entre las secuencias, sacrifica pérdidas en pro de las que tengan mayor peso.

Conclusiones

- Supera consistentemente todos los baselines del modelo simple y de la separación PCA.
- También supera algunos baselines del modelo no entrenado.
- WLSL al tener datos más limpios, llevó a los mejores resultados. Logró una separación semántica muy buena. Aprendizaje temporal exitoso.
- ISL y SLOVO, al tener datos ruidosos con alta variabilidad, impidió una correcta separación semántica, el modelo falló en aprender la jerarquía temporal correcta al evaluarlo con sus 3 etiquetas.

- Con los tres al tiempo, el modelo falló en la agrupación semántica, confundiendo señas entre idiomas. Pero aprendiendo la estructura temporal.
- La arquitectura propuesta es buena para codificar el orden temporal.
- La calidad y uniformidad de los datos es más importante para la separación semántica que la cantidad de los mismos.
- El proyecto sienta una base sólida para futuras investigaciones sobre la dinámica temporal en el CSLR con componentes que se exploraron en esta técnica.

Referencias

- B. Probierz et al. «Sign language interpreting-relationships between research in different areas overview». *Annals of Computer Science and Information Systems* 35 (2023)
- B. Parmar et al. «I always feel like I'm the first deaf person they have ever met: "Deaf Awareness, Accessibility and Communication in the United Kingdom's National Health Service (NHS): How can we do better?"» *PLoS ONE* 20.5 (2025)
- X. Jiang et al. «A survey on Chinese sign language recognition: From traditional methods to Artificial Intelligence». *Computer Modeling in Engineering & Sciences* 140.1 (2024)