

# Deep Learning para visión por computadora

Facundo Quiroga, Franco Ronchetti, Gastón Rios, Pedro Dal Bianco, Santiago Ponte Ahon, Oscar Stanchi, Juan Seery, Tatiana Badaracco, Federico Rabinovich, Javier Vazquez, Aurelio Fernandez-Bariviera, Saif Khalid, Hatem Rashwan, Domenec Puig Valls, Yael Aidelman, Roberto Gamen, Laura Lanzarini, Waldo Hasperué.

fquiroga@lidi.info.unlp.edu.ar

## Contexto

El Instituto de Investigación en Informática LIDI (III-LIDI) tiene una larga trayectoria en el estudio, investigación y desarrollo de Sistemas Inteligentes basados en distintos métodos de Aprendizaje Automático y Redes Neuronales. Esta presentación corresponde a algunas de las tareas de investigación que se llevan a cabo en el III-LIDI en el marco del proyecto "Inteligencia de Datos. Técnicas y Modelos de Machine Learning" perteneciente al Programa de Incentivos (2023-2026).

## Líneas de Investigación y Desarrollo

### Reconocimiento de Lengua de Señas

En esta línea de investigación, el objetivo es la traducción de videos donde las personas se comunican utilizando la lengua de señas a una lengua escrita.

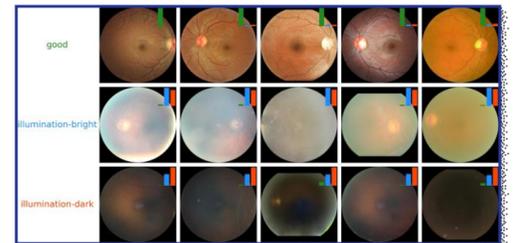
Se desarrolló y se validó la base de datos "LSA-T", que contiene 14.880 videos a nivel de oración de LSA extraídos del canal de YouTube CN Sordos. Esta es la base más grande al momento con señas de la LSA y una de las más grandes del mundo en general. Además, se puso en marcha la plataforma *app.seni.ar* para permitir que la comunidad sorda contribuya con nuevos videos de LSA, y agregue información sobre los existentes



Por otro lado, se desarrollaron modelos basados en Redes Generativas Adversarias (GAN) para la generación de imágenes de forma de mano que puedan asistir a entrenar o preentrenar modelos de TLS para la LSA. Los modelos obtenidos permiten condicionar al generador utilizando la pose mediante un módulo SPADE. De esta forma es posible generar imágenes de formas de mano arbitrarias que mejoran el desempeño de los modelos.

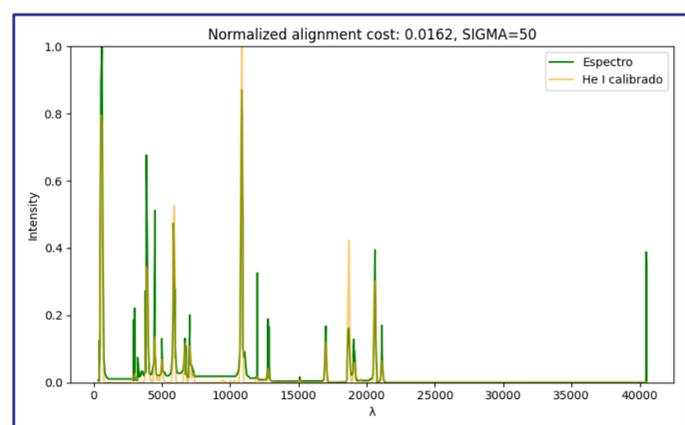


Por este motivo, se presentó un modelo basado en la CNN UNet de predicción de calidad de imagen que se denominó VISTA. El modelo contiene un módulo de Concept Whitening, una técnica de interpretabilidad intrínseca, que indica en base a qué conceptos la red realiza sus predicciones.



### Sistema de análisis de placas espectrográficas antiguas

Continuando el desarrollo de un sistema de análisis automatizado de placas espectrográficas del proyecto ReTroH[2], se está desarrollando el módulo de alineamiento de lámparas de comparación empíricas y teóricas. Se experimentó con varios métodos de base para ser comparados con los métodos a desarrollar basados en CNNs



## Formación de Recursos Humanos

El grupo de trabajo de la línea de I/D aquí presentada está formado por: 3 profesores con dedicación exclusiva, un JTP dedicación exclusiva, 2 investigadores CIC-PBA, 4 becarios de posgrado de la UNLP con dedicación docente, 3 tesis de maestría de la UBA, y 4 profesores extranjeros.

Actualmente se están desarrollando 4 tesis de doctorado y 3 de magister. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.