

Deep Learning para visión por computadora

Franco Ronchetti, Facundo Quiroga, Gastón Ríos, Pedro Dal Bianco, Santiago Ponte Ahon, Oscar Stanchi, Laura Lanzarini, Alejandro Rosete, Waldo Hasperué

fronchetti@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 las tareas de investigación que se llevan a cabo en el III LIDI en el marco del proyecto F025 "Sistemas inteligentes. Aplicaciones en reconocimiento de patrones, minería de datos y big data" perteneciente al Programa de Incentivos (2018-2023).

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 la base de datos "LSA-T", que contiene 14.880 videos a nivel de oración de LSA extraídos del canal de YouTube CN Sordos. Contiene, además, para cada video, su traducción al español e información posicional correspondiente a cada señante. También se presentó un método para inferir al señante entre varias personas que pudieran aparecer en un mismo video, un análisis detallado de las características del conjunto de datos, una herramienta de visualización para explorar el conjunto de datos y un modelo de SLT neuronal para servir como línea de base para futuros experimentos.

	LSA-T	PHO ENIX*	SIGN UM	CSL	GSL	KSL
País	Arg	Ale	Ale	Chi	Gre	Korea
Real	Si	Si	No	No	No	No
Señantes	103	9	25	50	7	14
Frases	14,254	5672	780	100	331	105



También se ha desarrollado una aplicación web con Tensorflow-JS con el doble propósito de facilitar el acceso a los modelos de traducción de lengua de señas y, mediante crowdsourcing, recolectar nuevos videos de la misma para mejorar los modelos.

¡Traduce en Tiempo Real! (BETA)
Utiliza el sistema de reconocimiento automático de lengua de señas argentina enfocándolo a la persona que realizará las señas, el sistema captará sus movimientos y obtendrá los subtítulos en tiempo real.



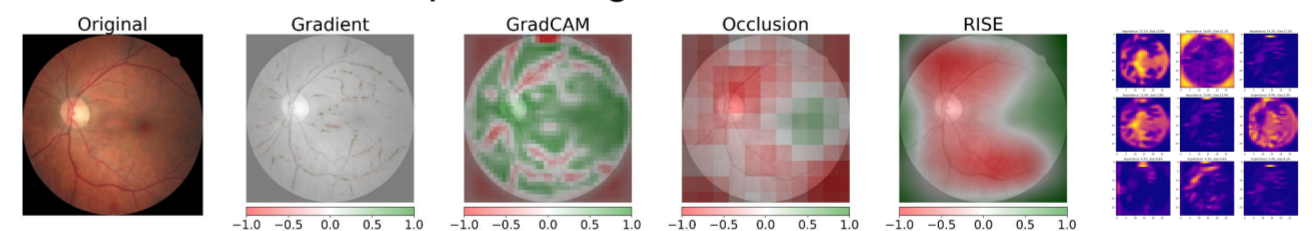
Finalmente, se están utilizando Redes Generativas Adversarias (GANs) y Modelos de Difusión para generar imágenes y luego videos artificiales relacionadas con la lengua de señas. Este tipo de redes permitirá, por un lado, generar videos de LSA a partir de texto, a modo de traducción y por otro, aumentar las bases de datos de formas de mano.

En particular, se está haciendo foco en el uso complementario de tanto imágenes como información obtenida con modelos de localización de articulaciones para mejorar la calidad de las imágenes generadas por una GAN y poder controlar las poses.



Interpretabilidad de modelos de Deep Learning

En el marco de esta línea de investigación se analizaron los conjuntos de datos EyePACS y EyeQ cuyas imágenes sirven para evaluar retinopatía diabética y la calidad de las imágenes para esta patología, respectivamente. Para ello, se realizó una implementación de RISE (Randomized Input Sampling for Explanation) para Captum. También se aplicaron métodos no supervisados sobre los resultados de técnicas de interpretabilidad post-hoc locales para ofrecer una visión global del funcionamiento de un modelo de Deep Learning.



Formación de Recursos Humanos

El grupo de trabajo de la línea de I/D aquí presentada está formado por: 3 profesores con dedicación exclusiva, 1 investigador CIC-PBA, 3 becarios de posgrado de la UNLP con dedicación docente, 2 becarios CIN, y un profesor extranjero.

Dentro de los temas involucrados en esta línea de investigación, en los últimos dos años se han finalizado diversas tesinas de grado, como así también trabajos de final de carrera y de especialización.

Actualmente se están desarrollando 2 tesis de doctorado, 2 tesis de especialista, 1 de magister, 3 tesinas de grado de Licenciatura y 2 trabajos finales de Ingeniería en Computación. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.