

## APRENDIZAJE AUTOMÁTICO. APLICACIONES EN RECONOCIMIENTO DE GESTOS, ACCIONES Y SEÑAS

Quiroga Facundo

Lanzarini Laura Cristina (Dir.)

Instituto de Investigación en Informática LIDI (III-LIDI), Facultad de Informática, UNLP-CIC.

[facundoq@gmail.com](mailto:facundoq@gmail.com)

**PALABRAS CLAVE:** Redes Neuronales Artificiales, Lengua de Señas, Visión por computadora.

El Aprendizaje Automático es una rama de la Inteligencia Artificial que estudia sistemas capaces de aprender a realizar una tarea a partir de datos de ejemplo. Es de naturaleza inductiva, a diferencia de la inteligencia artificial clásica, y comprende técnicas y métodos para realizar clasificación, optimización y predicción, mayormente en dominios en donde los problemas no pueden definirse de forma explícita o no existen soluciones analíticas aplicables. Por estos motivos, las técnicas que presenta resultan adecuadas para el procesamiento de imágenes y otras señales.

En los últimos años el procesamiento de texto, sonido, video y otras señales ha experimentado grandes progresos mediante el uso de una técnica de Aprendizaje Automático denominada Redes Neuronales Profundas o Aprendizaje Profundo (Deep Learning), que extiende los modelos previos de redes neuronales artificiales con arquitecturas y algoritmos de optimización que permiten entrenar redes de varias capas con grandes cantidades de datos de entrenamiento.

El objetivo de éstos métodos es aprender u obtener representaciones ideales de los datos sin intervención del usuario, mediante el uso de modelos apropiados y bases de datos con grandes cantidades de ejemplos. Esto las diferencia de las herramientas tradicionales de Aprendizaje Automático, en donde los expertos usualmente deben invertir una cantidad significativa de tiempo en desarrollar representaciones adecuadas de la información.

Las redes neuronales profundas están siendo utilizadas en todos los campos en donde se utiliza aprendizaje automático. En particular, las Redes Convolucionales Profundas, es decir, redes neuronales profundas

con capas convolucionales, se han aplicado al reconocimiento de secuencias de acciones, gestos y señas en videos con resultados que superan al estado del arte.

Como aplicación, en este plan se propone estudiar métodos para entender que sucede en una escena a partir de una imagen o video de la misma. Se hará énfasis en tres subproblemas: reconocer señas, acciones y gestos realizados por personas. Cada uno de estos problemas tiene sus particularidades, pero se encuentran estrechamente interrelacionados. El reconocimiento de señas busca convertir a texto un video donde una persona realiza señas en alguna lengua de señas existente, como la Lengua de Señas Argentina (LSA), con sus reglas particulares pre-establecidas. En el reconocimiento de acciones, se busca entender y categorizar una acción arbitraria que realiza una persona. Por último, en el reconocimiento de gestos usualmente se busca reconocer un conjunto de gestos previamente establecidos y de uso particular para un dominio específico.

Si bien en el último tiempo han habido avances en este área, impulsados principalmente por el desarrollo de nuevas tecnologías, aún queda un largo camino por recorrer para construir aplicaciones precisas y robustas que permitan, por ejemplo, la traducción e interpretación de las señas realizadas por un intérprete.

## RECONOCIMIENTO DE GESTOS DINÁMICOS Y SU APLICACIÓN A LA LENGUA DE SEÑAS

Ronchetti Franco

Lanzarini Laura Cristina (Dir.)

Instituto de Investigación en Informática LIDI (III-LIDI), Facultad de Informática, UNLP-CIC.

[francoronchetti@gmail.com](mailto:francoronchetti@gmail.com)

**PALABRAS CLAVE:** Machine Learning, Visión por Computador, Reconocimiento de gestos.

Esta investigación se plantea como una continuación y transferencia concreta del plan de trabajo de tesis de quien suscribe. El objetivo general de este plan consiste en la extensión del clasificador de gestos dinámicos desarrollado en la tesis doctoral con el fin de realizar una transferencia tecnológica concreta al lenguaje de señas. Con esto se espera ayudar a la comprensión de dicho lenguaje por parte del público en general haciendo

innecesaria cualquier tipo de instrucción previa. Para lograr este fin, los objetivos específicos son los siguientes:

- Focalizarse en la etapa de detección de manos para poder realizar una segmentación sin necesidad de marcadores de color. Esto permitiría realizar pruebas en otras bases de datos existentes donde no existe