

Desafíos del Aprendizaje Profundo en la Visión por Computador

Victor Hugo Ayma Quirita (Moderador)

vayma@ulima.edu.pe

<https://orcid.org/0000-0002-0284-2610>

Carrera de Ingeniería de Sistemas, Universidad de Lima, Perú

Pedro Marco Achanccaray Diaz

p.diaz@tu-braunschweig.de

<https://orcid.org/0000-0002-7324-9611>

Institute of Geodesy and Photogrammetry

Technical University of Braunschweig, Alemania

Smith Washington Arauco Canchumuni

saraucoc@puc-rio.br

<https://orcid.org/0000-0003-0812-0085>

Pontifical Catholic University of Rio de Janeiro, Brasil

<https://doi.org/10.26439/ciis2022.6070>

Resumen

La Visión por Computador es un área de estudio en la Inteligencia Artificial que se enfoca en el desarrollo de técnicas computacionales para percibir el mundo a través de entradas visuales, como videos o imágenes. El Aprendizaje Profundo ha demostrado ser una técnica eficiente para el análisis e interpretación de datos visuales. Sin embargo, afronta inúmeros desafíos según su aplicación en las diferentes tareas de la Visión por Computador. Este panel reúne un grupo de expertos en Aprendizaje Profundo, quienes ofrecerán información sobre su aplicación y desafíos a ser superados en sus respectivas áreas de investigación con relación a la Visión por Computador.

Palabras clave: visión por computador, aprendizaje profundo.

Challenges of Deep Learning in Computer Vision

Abstract

Computer Vision is a field of study within Artificial Intelligence that focuses on developing computational techniques to perceive the world through visual data, such as video or images. Deep Learning has proven to be efficient in visual data analysis and interpretation. Nevertheless, it faces countless challenges given its application in several Computer Vision tasks. This panel brings together Deep Learning experts, who will share information about Deep Learning applications and challenges to overcome in their research fields regarding Computer Vision.

Keywords: computer vision, deep learning.

Cómo citar

Ayma, V., Achanccaray, P., & Arauco, S. (2022). Desafíos del Aprendizaje Profundo en la Visión por Computador. *Actas del Congreso Internacional de Ingeniería de Sistemas 2022: Entornos híbridos en la pospandemia: posibilidades para las nuevas tecnologías*, e6070. <https://doi.org/10.26439/ciis2022.6070>

1. Introducción

La Visión por Computador es una subárea de la Inteligencia Artificial que se enfoca en el desarrollo de técnicas computacionales para el análisis e interpretación de datos visuales, como videos o imágenes (Klette, 2013; Prince, 2012). La Visión por Computador se ha convertido en una tecnología fundamental para muchos campos de la industria como la seguridad, el cuidado de la salud, la agricultura, el entretenimiento, así como la industria textil y automotriz. El Aprendizaje Profundo es un conjunto de técnicas basadas en Redes Neuronales (RN) que han contribuido al desarrollo de la Visión por Computador. Desde el sorprendente rendimiento alcanzado por las Redes Neuronales Convolucionales (CNN) en la competencia ImageNet¹, estas se han convertido en los modelos de Aprendizaje Profundo de referencia para la clasificación de imágenes, detección de objetos segmentación semántica, y estimación de la postura humana (Guo et al., 2016). No obstante, a la popularidad de las CNNs, otros modelos de Aprendizaje Profundo, como los Autoencoders (AE), las Redes Neuronales Recurrentes (RNN), y las Redes Generativas Adversariales (GAN), han demostrado ser eficientes en la realización de diferentes tareas de la Visión por Computador. Por ejemplo, las RNNs se han aplicado en el reconocimiento de acciones, el reconocimiento de escritura y seguimiento de objetos en video; los AEs se han aprovechado de forma eficiente para eliminar el ruido de imágenes y realizar búsquedas en la web a base de imágenes; las GANs se han aplicado en la generación de imágenes realistas a partir de texto y bocetos, así como en la generación de vistas frontales de rostros para sistemas de reconocimiento facial.

A pesar de los grandes avances que ha experimentado el Aprendizaje Profundo en la última década, y los impresionantes resultados alcanzados en la realización de tareas de Visión por Computador, los diferentes modelos de Redes Neuronales que componen la tecnología del Aprendizaje Profundo aún enfrentan desafíos que necesitan ser atendidos.

2. Presentación

Este panel reúne un grupo de tres expertos en Aprendizaje Profundo y Visión por Computador con reconocida trayectoria en investigación, proyectos de desarrollo e innovación, quienes ofrecerán sus perspectivas sobre los avances del Aprendizaje Profundo y los desafíos que esta tecnología enfrenta en sus áreas de investigación relacionadas a la Visión por Computador.

¹ <https://image-net.org/>

2.1 Ph.D. Victor Hugo Ayma Quirita (moderador)

Doctor y Máster en Ingeniería Eléctrica de la Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio - Brasil). Egresado de la Universidad Nacional de San Antonio Abad del Cusco (UNSAAC - Perú). Tiene experiencia en investigación usando aprendizaje automático y aprendizaje profundo. Actualmente se desempeña como investigador y docente en el Instituto de Investigación Científica (IDIC) y la Carrera de Ingeniería de Sistemas (CIS) de la Universidad de Lima en Perú, respectivamente.

2.2 Ph.D. Pedro Marco Achancaray Diaz

Doctor y Máster en Ingeniería Eléctrica de la Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio - Brasil). Egresado de la Universidad Nacional de Ingeniería (UNI - Perú). Cuenta con experiencia en proyectos de Investigación, Desarrollo e Innovación (I+D+i) usando teledetección, inteligencia artificial y aprendizaje profundo en las áreas de agricultura, petróleo y gas, y preservación de patrimonio cultural. Actualmente se desempeña como investigador posdoctoral en el Institute of Geodesy and Photogrammetry (IGP) de la Technical University of Braunschweig (TU-BS) en Alemania.

2.3 Ph.D. Smith Washington Arauco Canchumuni

Máster en Ingeniería Mecánica y Doctor en Ingeniería Eléctrica de la Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio - Brasil). Egresado de la Universidad Nacional de Ingeniería (UNI - Perú) en Ingeniería Mecatrónica. Cuenta con años de experiencia en proyectos de Investigación, Desarrollo e Innovación (I+D+i) usando inteligencia artificial, aprendizaje automáticos/aprendizaje profundo, y modelos generativos en las áreas de petróleo y gas. Actualmente se desempeña como investigador y profesor en el laboratorio de inteligencia computacional de la PUC-Rio.

Referencias

- Akcil, U., Uzunboylu, H., y Kinik, E. (2021). Integration of Technology to Learning-Teaching Processes and Google Workspace Tools: A Literature Review. *Sustainability*, 13(9). <https://doi.org/10.3390/su13095018>
- Al-Omar, K. (2018). Evaluating the Usability and Learnability of the “Blackboard” LMS Using SUS and Data Mining. 386-390. <https://doi.org/10.1109/iccmc.2018.8488038>
- Easttom, C. (2021). Aplicación de técnicas matemáticas e ingeniería a la ciberseguridad. *Actas Del Congreso Internacional De Ingeniería De Sistemas*, 33-48. <https://doi.org/10.26439/ciis2021.5575>
- Google for Education. (28 de agosto del 2020). *Capacitación sobre aspectos básicos*. <https://skillshop.exceedlms.com/student/path/111629-capacitacion-sobre-aspectos-basicos>
- Hernández Sampieri, R., Fernández Collado, C., y Baptista Lucio, M. del P. (2010). *Metodología de la investigación* (Quinta). McGraw Hill.
- Mendoza, S. H. V. (2015). *Improving Communication and Building Communities with Google* [Presentación de paper]. Proceedings of the 2015 ACM SIGUCCS Annual Conference, 85-90. <https://doi.org/10.1145/2815546.28155>