

Alberto Andrés Valdés González.

Degree: Mathematical Engineer.

Work position: Data Scientist.

Mail: anvaldes@uc.cl/alberto.valdes.gonzalez.96@gmail.com

Location: Santiago, Chile.

State of Art - Deep Learning

Computer Vision:

Alexnet: AlexNet is the name of a convolutional neural network (CNN) architecture, designed by Alex Krizhevsky in collaboration with Ilya Sutskever and Geoffrey Hinton, who was Krizhevsky's Ph.D. advisor at the University of Toronto.

GoogLeNet: GoogLeNet is a type of convolutional neural network based on the Inception architecture. It utilises Inception modules, which allow the network to choose between multiple convolutional filter sizes in each block. An Inception network stacks these modules on top of each other, with occasional max-pooling layers with stride 2 to halve the resolution of the grid.

VGG19: VGG-19 is a convolutional neural network that is 19 layers deep. You can load a pretrained version of the network trained on more than a million images from the ImageNet database. The pretrained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224-by-224.

Resnet50: ResNet stands for Residual Network and is a specific type of convolutional neural network (CNN) introduced in the 2015 paper "Deep Residual Learning for Image Recognition" by He Kaiming, Zhang Xiangyu, Ren Shaoqing, and Sun Jian. CNNs are commonly used to power computer vision applications.

ResNet-50 is a 50-layer convolutional neural network (48 convolutional layers, one MaxPool layer, and one average pool layer). Residual neural networks are a type of artificial neural network (ANN) that forms networks by stacking residual blocks.

Transformers/Generative AI:

ChatGPT: ChatGPT (generalized brand from the English acronym Chat Generative Pre-Trained Transformer) is an artificial intelligence chatbot application developed in 2022 by OpenAI that specializes in dialogue. The chatbot is a language model adjusted with both supervised and reinforcement learning techniques. It is composed of the GPT-4 and GPT-3.5 models from OpenAI.

Dall-e: DALL-E is an artificial intelligence program that creates images from textual descriptions or stimuli (prompts in English), revealed by OpenAI on January 5, 2021. It uses a version of 12 billion parameters² of the GPT-3 Transformer model to interpret natural language inputs (such as a green leather bag shaped like a pentagon or an isometric view of a sad capybara) and generate corresponding images.

Gemini: Gemini is built from the ground up for multimodality — reasoning seamlessly across text, images, video, audio, and code.

BERT: Bidirectional Encoder Representations from Transformers (BERT) is a language model based on the transformer architecture, notable for its dramatic improvement over previous state of the art models. It was introduced in October 2018 by researchers at Google. A 2020 literature survey concluded that in a little over a year, BERT has become a ubiquitous baseline in Natural Language Processing (NLP) experiments counting over 150 research publications analyzing and improving the model.”
