

Deep Point of View **Generative AI**

03

Common Techniques of Generative AI

Artificial intelligence techniques have traditionally been used to clean data, enhance predictive analysis, compress data, and decrease the dimensionality of datasets for other algorithms. Novel generative AI techniques like Variational autoencoders (VAEs), for example, push this a step further by reducing errors between the raw signal and the reconstruction.

Generative models are exceptionally good at producing near-original material with a little vector. It also enables us to create previously non existing material that may be used without licensing. Some Generative AI techniques are used when working with pictures or visual data. There are certain Generative AI models that perform better in signal processing applications such as anomaly detection for predictive maintenance or security analytics. Let's discuss some of these Generative AI techniques in this section.

• Generative Adversarial Networks

Ian Goodfellow and colleagues at the University of Montreal pioneered the use of Generative Adversarial Networks (GANs) in 2014. They have showed enormous potential in producing many forms of realistic data. Yann LeCun, Meta's chief AI scientist, called GANs and their variants "the most exciting topic in machine learning in the last ten years."

For starters, they have been utilized to make realistic speech by mimicking humans and matching voices and lip movements for better translations. They have also interpreted visuals, distinguished between night and day, and defined dancing motions between bodies. They are also used in conjunction with other AI approaches to increase security and create stronger AI classifiers.

GANs use two competing neural networks, a generator and a discriminator. The generator, also known as the generative network, is a neural network responsible for producing new data or content comparable to the original data. A discriminator, also known as a discriminative

network, is a neural network that differentiates between source and produced data. The competition between these two networks is to develop their algorithms until they can create data indistinguishable from the original material.

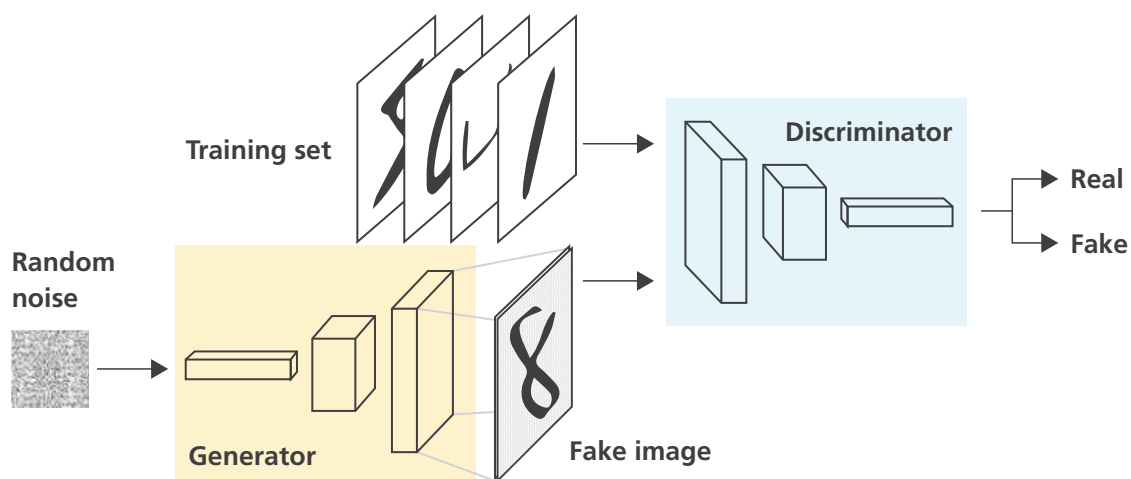


Fig 3: Generative Adversarial Networks (Thalles Silva)

• Models

DCGAN

(Deep
Convolutional
GAN)

ProGAN

(Progressively
Growing
GAN)

BigGAN

• Transformer-based Models

Transformer-based models are mainly used to analyze data with a sequential structure (such as the sequence of words in a sentence). In modern times, transformer-based techniques have become a standard tool for modeling natural language.

The ability of the transformer models to attend to various positions of the input sequence to compute a representation of that sequence is core to their architecture.



Fig 5: Transformer-based Models (Source: GitHub)

• Models

BERT

(Bidirectional Encoder Representation from Transformers)

RoBERTA

(Robustly Optimized BERT)

• Autoregressive Convolutional Neural Networks

Autoregressive refers to self-regression. The word autoregression refers to forecasting future outcomes of a series based on previously observed effects of that sequence. AR- CNNs investigate systems that change over time and believe that the

likelihood of specific data is only based on what has happened before. To create reliable new data, they rely on past time-series data. RNNs and casual convolutional networks are the most common autoregressive designs.

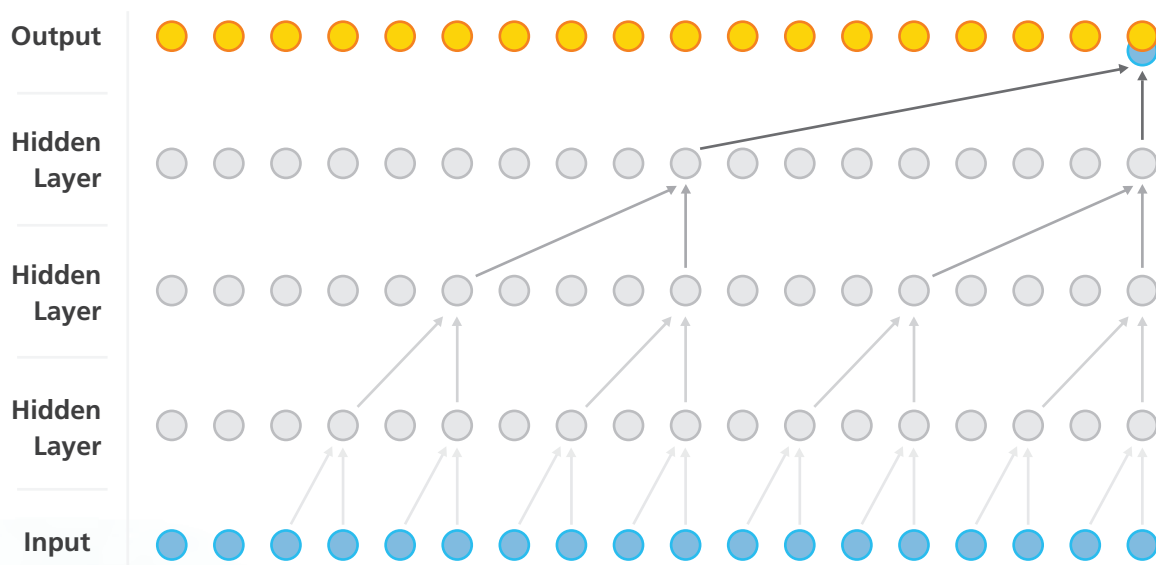


Fig 4: Autoregressive Convolutional neural networks (Source: deeppmind.com)

• Models

PixelRNN

PixelCNN

WaveNet

• Other Nascent Techniques

Bayesian Network

Bayesian Network or Bayes Network is a generative probabilistic graphical model that allows efficient and effective representation of the joint probability distribution over a set of random variables. Bayes Network consists of two main parts, which are structure and parameters. The structure is a directed

acyclic graph (DAG), and the parameters consist of conditional probability distributions associated with each node. This network can be used for various applications, such as time series prediction, anomaly detection, reasoning, etc.

Gaussian Mixture Model

Gaussian Mixture Model is a generative probabilistic model which assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. GMMs are commonly used as a parametric model of the probability distribution of features in a

biometric system, which includes vocal tract-related spectral components in a speaker recognition system. Thus, a well-trained prior model estimates GMM parameters from training data using the iterative Expectation-Maximization (EM) algorithm or Maximum a Posteriori (MAP) estimation.

Hidden Markov Model

A Hidden Markov Model (HMM) is a statistical model that can describe the evolution of observable events that depend on internal factors, which are not directly observable. The model is popularly known for its effectiveness in modeling the correlations between adjacent symbols, domains,

or events. They have been extensively used in various fields, especially in speech recognition and digital communication. A Hidden Markov Model consists of two stochastic processes: an invisible circle of hidden states and a visible process of observable symbols.

Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation (LDA) is a generative probabilistic model with collections of discrete data such as text corpora. LDA is a three-level hierarchical Bayesian model in which each collection

item is modeled as a finite mixture over an underlying set of topics. The model has applications for various problems, including collaborative filtering and content-based image retrieval.

Variational Autoencoders (VAEs)

Variational Autoencoders (VAEs) have been one of the most popular approaches to unsupervised learning of complicated distributions. They are built on top of standard function approximators, which are neural networks and can be trained with

stochastic gradient descent. The application of VAEs includes generating various kinds of complex data, including handwritten digits, faces, CIFAR images, predicting the future from static images, and more.