2022: THE YEAR OF TEXT-TO-ANYTHING

ALBERTO ARKADER KOPILER

IMAGE PROCESSING COURSE

IMPA 2022

He used AI art from Midjourney to win a fine-arts prize. Did he cheat? - The Washington Post

2022: THE YEAR OF TEXT-TO-ANYTHING

Artificial Intelligence is gaining momentum...

- AI winning rule based games (Chess, Go)
- Creation of very expressive art
- Engaging very sophisticated conversations
- Writing essays

Elon Musk in the Tesla AI Day *'now is the best time to be alive.'*

2022: the year of the text-to-anything, text-to-everything or the year of '**Generative Artificial Intelligence**.'

TEXT-TO-VIDEO

- Google: 'Imagen Video' a text-to-video generative AI model capable of producing high-definition videos from a text prompt
- Meta: 'Make-A-Video' an AI system that allows users to turn text prompts into short video clips

TEXT-TO-AUDIO

- Google: 'AudioLM' an audio generation model that learns to generate realistic speech and piano music by listening to audio-only.
- Meta: 'AudioGen,' an auto-regressive generative model that generates audio samples based on text inputs.

TEXT-TO-3D

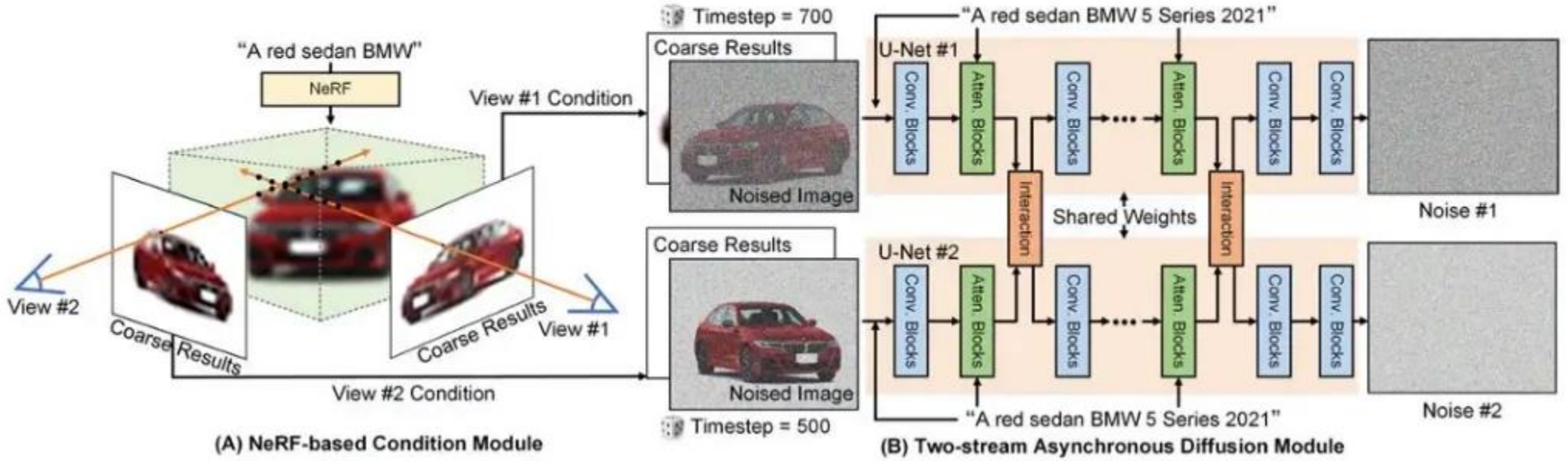
- Google:'DreamFusion', discovered a method to produce 3D models based on a user's word input The new technology, dubbed 'DreamFusion', employs 2D Diffusion and is expected to make significant advances in text-to-image generation.

The researchers optimised a randomly-initialised 3D model called 'NeRF' (a Neural Radiance Field) via gradient descent so that the renderings in 2D from random angles achieve lesser loss.

TEXT-TO-3D

3DDesigner: Towards Photorealistic 3D Object Generation and Editing with Text-guided Diffusion Models

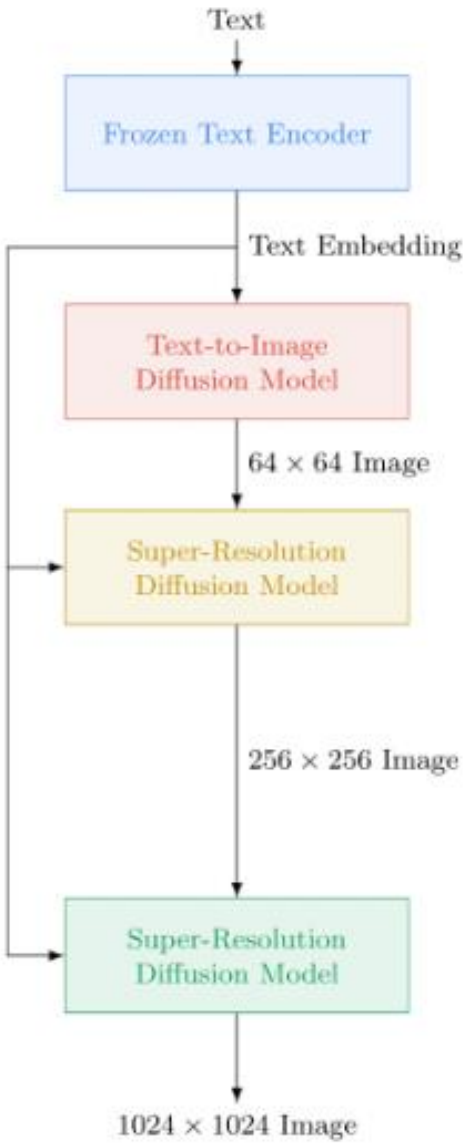
Text-guided 3D-consistent generation framework (training phase).

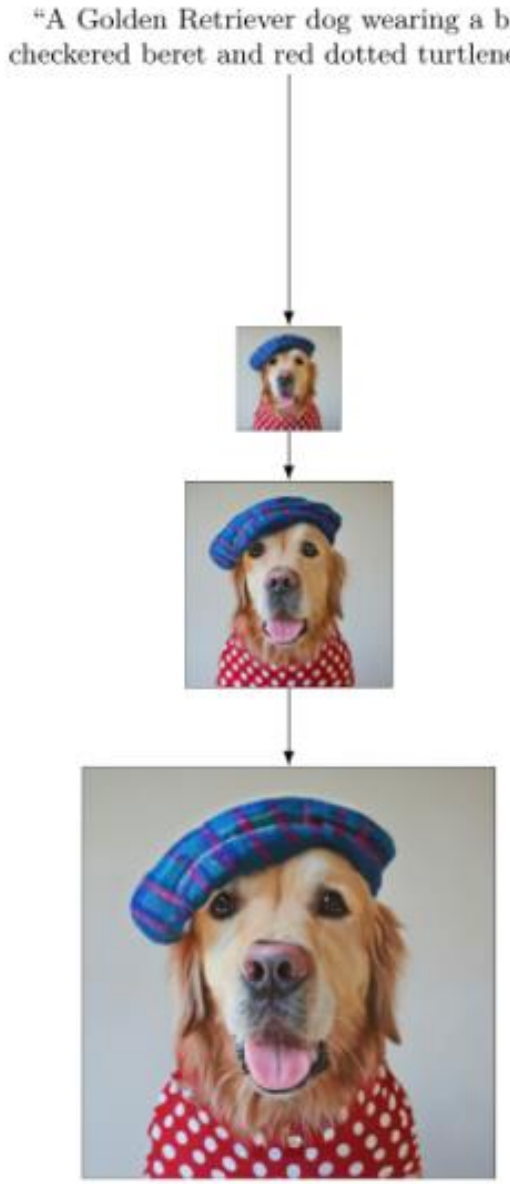


TEXT-TO-3D

A novel model that learns 3D from a single photo (Text-to-3D)

TEXT-TO-IMAGE





Credit: Google

TEXT-TO-IMAGE

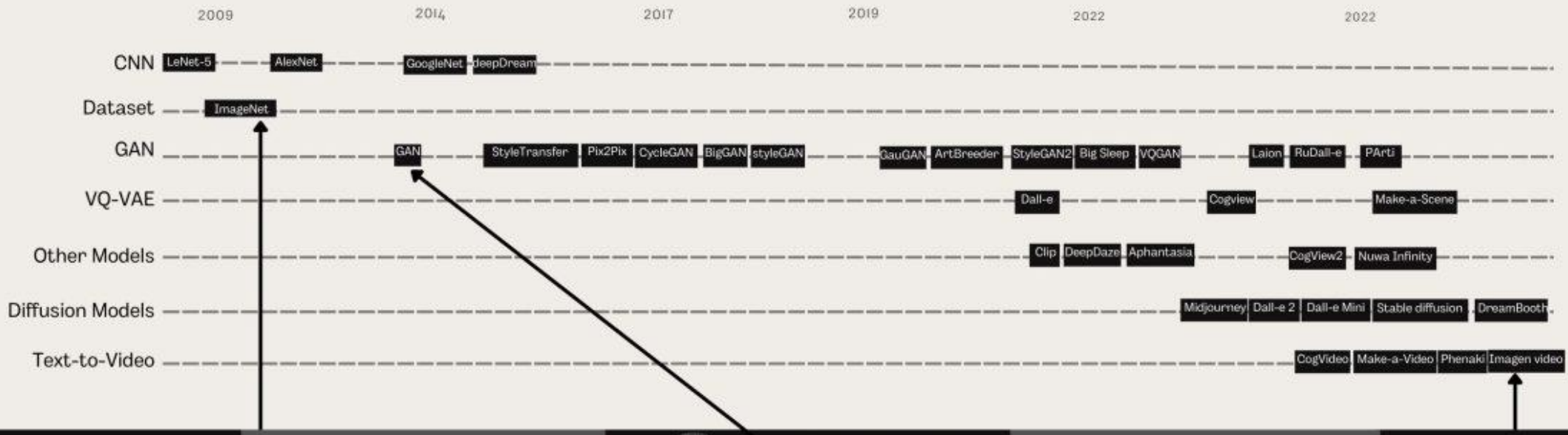
- OpenAI: DALLE-2
- Google: Imagen
- Meta: Make a Scene
- Midjourney: Midjourney Showcase
- stability.ai: Stable Diffusion 2.0 Release — Stability.Ai

TEXT-TO-HISTORY

- OpenAI: GPT-3 https://openai.com/api/ **Generative Pre-trained Transformer 3** (**GPT-3**; stylized **GPT·3**) is an autoregressive language model that uses deep learning to produce human-like text. Given an initial text as prompt, it will produce text that continues the prompt. It can auto-generate code as well! **175 billion parameters**
- Turing Test: The quality of the text generated by GPT-3 is so high that it can be difficult to determine whether or not it was written by a human (https://en.wikipedia.org/wiki/GPT-3)
- Imagenet and Wordnet: **https://www.web3.lu/wordnet-imagenet/**
- Google Search: By text, by image, by style (google scholar)
- Translation is the beginning of everything
- Is this the new Esperanto?

2022 IS THE YEAR OF TEXT-TO-ANYTHING

2022 saw many developments around art generators, starting with text-to-image generator Open AI's DALL E-2. Not just text-to-image generators, text-to-audio, text-to-video and even text-to-shop have become the talk of the town. Let's see some of the most popular systems over years.



One of the first convolutional neural networks, Lenet-5, was created in 1998 for character identification in both handwriting and machine printing.

The ImageNet project is a large visual database designed for use in visual object recognition software research. More than 14 million images have been hand-annotated by the project.

Generative adversarial networks (GAN) were designed by Ian Goodfellow and his colleagues in June 2014. Two neural networks contest with each other in the form of a zero-sum game, where one agent's gain is another agent's loss.

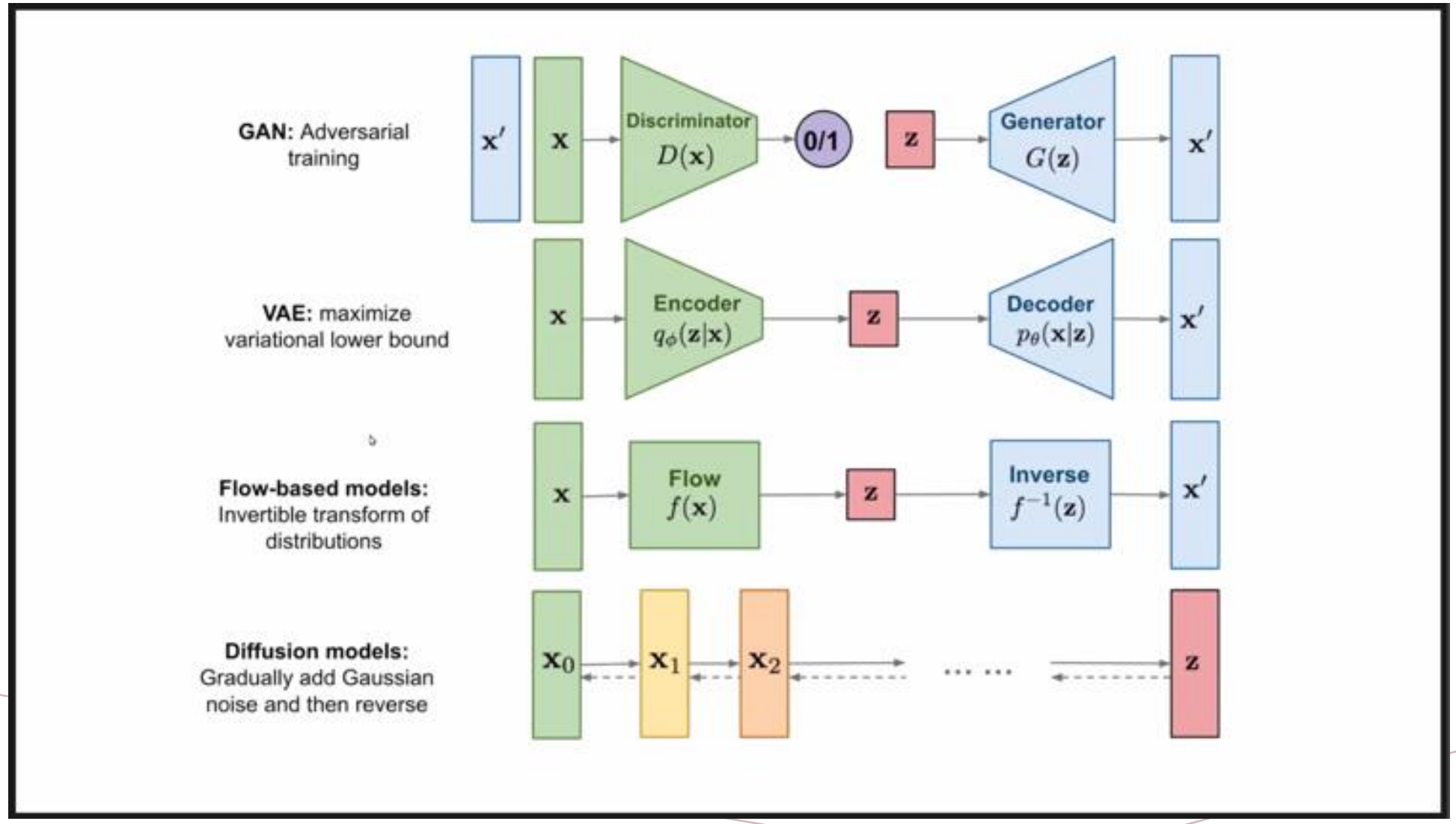
DALL·E is a 12-billion parameter version of GPT-3 trained to generate images from text descriptions, using a dataset of text–image pairs.

Imagen is the text-to-video model developed by Google In 2022 as a replacement for their T2I paradigm.

Copyright: Analytics India Magazine

With audio, image and video being created just by giving a prompt, there is no doubt that 2022 has been the year of text-to-anything.

TEXT-TO-HISTORY



2022: THE YEAR OF TEXT-TO-ANYTHING OR EVERYTHING

THANK YOU!

ALBERTO ARKADER KOPILER
akopiler@gmail.com

Credits: AIM – Analytics India Magazine