DALL·E

By: Isaiah Singer
CS 131
FALL 2022

What is DALL·E?

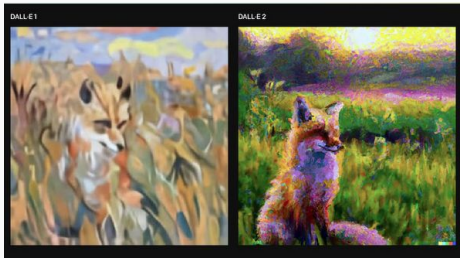
- DALL-E is an artificial intelligence tool that is able to render text prompts into realistic images.
- DALLE-2 is the new and improved program they created which generates more realistic and accurate images with 4x greater resolution.
- DALL-E and DALL-E 2 were created by OpenAI; an artificial intelligence research and deployment company.
- It is able to create simple and/or complex images such as anthropomorphized versions of animals and objects.

How Does It Work?

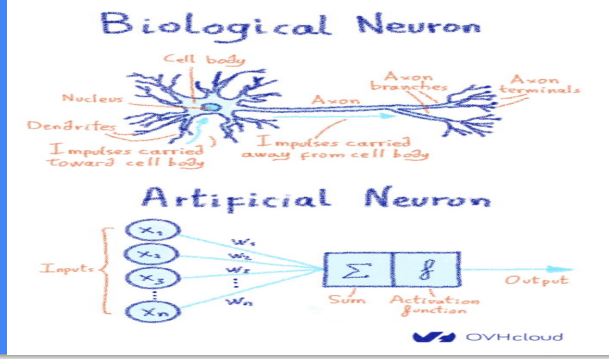
- A text prompt is input into a text encoder that is trained to map the prompt to a representation space.
- A model known as a prior maps the text encoding to a corresponding image encoding that captures the semantic information of the prompt contained in the text encoding.
- An image decoder generates an image which is a visual manifestation of the semantic information.

Differences Between DALL-E And DALL-E 2

- DALLE was able to render AI-created images in a cartoonish fashion while DALLE-2 can produce realistic images.
- DALLE-2 employs a new strategy known as diffusion which begins with a pattern of random dots and gradually changes that pattern to resemble a picture.
- DALLE-2 has a new feature where it is able to edit and retouch photographs accurately based on a simple description.
- It is able to produce multiple iterations of an image.



Neural Networks



- In order to build a good ANN(Artificial Neural Network), you would need artificial neurons(processing nodes), multiple neuron input connections(dendrites), a computation unit(nucleus) which is composed of a linear function(ax+b), an activation function, and an output(axon).
- In order to pass the image to the neural network, the image would be represented by an array of numbers, where each pixel is represented by a different number depending on the brightness which is then passed as a function.

Why Is It Important/Special?

- It is an enhancement of the way we understand machine learning.
- In the future, real estate sites can generate models of unbuilt homes and/or redesigned homes tailored to the buyer's specifications.
- DALL-E takes us one step closer to machine learning "bridging the gap between language expression and pictures."
- The most profound implications will only become visible over time, so we will have to wait to see what is created as a result of our current AI developments.

Downsides

- DALL-E can potentially create harmful images due to the stereotypes they have learned from their training. DALL-E has a filter to blacklist keywords such as blood or other racial slurs but DALL-E Mini does not.
- The artwork quality may be low for some prompts landscapes, human faces, or replicating popular art styles but DALL-E 2 has made significant improvements as far as that issue is concerned.
- It may take a while to get DALL-E since there is a waiting time in queue for the software and it costs money as well.

Conclusion

- While DALL-E 2 has improved significantly from it’s former version DALL-E, there is still a lot of room for improvement that is needed.
- It can be used as an excellent tool for professionals, educational institutions, and general use as well for companies.
- The future of deep learning is bright and in the future, could be used to create realistic images for movies and video games, among other things mentioned.