УДК 811.111:004.8
CREATING IMAGES WITH NEURAL NETWORKS

Shkantov I. N., student
Scientific supervisor – Vanik I. Y., senior lecturer
English language department №1
Belarusian National University of Technology
Minsk, Republic of Belarus

Nowadays, neural networks have very wide applications, including creating images. This process, often called neural style transfer, has revolutionized the way we think about art and technology.

A neural network is a machine learning type where a computer program mimics the human brain. Each network consists of artificial neurons that mimic human ones. These are software modules or nodes that interact and exchange information to solve a task. A basic neural network contains three layers of artificial neurons. The input layer processes external information, analyzes or classifies it, and passes it to the next hidden layer. This layer analyzes the output data of the previous layer, processes it, and passes it further to the output layer which gives the final result after processing all the data.

Neural networks are large systems of polynomial formulas with many variables. They require training with input data (like pixel color or letter codes) to produce an output (like object type from an image or a line of letters in another language).

It should be noted that training involves adjusting the coefficients within the network's formulas using a special algorithm until the network can accurately identify the input data. This process is repeated millions of times with different training images. The trained network, often referred to as a "model", doesn't store all the examples shown during training. Instead, it generates a value from the input data based on patterns identified during training.

By the mid-2010s, scientists and programmers had developed fast algorithms and selected good sets of training images to streamline the creation of image-recognizing neural networks.

By early 2021, the integration of BTM and the use of intermediate values of neural network formulas were combined by scientists and programmers. The challenge was to figure out how to use the codes of text

characters and picture pixels simultaneously in the neural network formulas so they wouldn't interfere with each other, but would be encoded in similar number formats and could participate together in the same formulas.

The intermediate results of applying pre-trained networks for both text and pictures will be similar-looking numbers, encoding some details of the content of the original input data.

This led to the creation of DALL-E, Midjourney, and several other large "multimodal" transformer models. The term "multimodal" denotes that they work with input data of several different formats or "modes". The models themselves are massive multi-terabyte files of formula coefficients. Powerful expensive video cards are used for fast simultaneous calculation of millions of formulas.

Midjourney is able to create something completely new and unique, because random number generators are used in the algorithms. The model has derived and fixed in itself three types of patterns: human language; the appearance of the world as a person sees it; connections between language and images. There are millions of these patterns, and their combinations - the factorial of these millions. The model derived them from data available on the Internet, since the Internet is a repository for an enormous amount of data. Midjourney composes patterns about our world, which it drew from our texts and pictures, adds a little randomness from the random number sensor, and generates a picture that best "according to humanity" corresponds to the entered phrase taking into account all additional settings.

In conclusion, the use of neural networks in creating images has opened up a new frontier in the intersection of art and technology. Through the process of training and manipulating these networks, we have been able to generate images that are not only visually appealing but also carry the distinct styles of different artworks.

References
1. Silva, T.W. An image generator based on neural networks in GPU / Silva, T.W., Reis, H., Melcher, E.U.K., Lima, A.M.N., Brito, A.V. // Multimedia Tools and Applications, 2022. V. 81. – P. 36353–36374.

2. Generative models [Electronic resource] – Mode of access: https://openai.com/research/generative-models. – Date of access: 27.03.2024.