**UDC 004.8**

# APPLICATION OF ARTIFICIAL INTELLIGENCE FOR IMPROVING THE EFFICIENCY OF IMAGE GENERATION

D.E. Mulyavin, R.V. Malcheva, E.N. Kushnirenko
vertik5555@mail.ru

***Abstract.*** *The aim of this study is to validate the potential utilization of neural networks and machine learning in graphic systems. It explores the components of artificial intelligence and provides an overview of their functioning. By conducting a concise analysis, a mathematical model is elucidated. The characteristics, types, and principles of model construction are explicated.*

***Key words:*** *artificial intelligence, machine learning, neural networks.*

**Introduction.** Image generation with the increasing humanity digitalization can serve as a good supporting tool in all fields of work. Let us consider some of the opportunities this technology offers:

1) Creation of unlimited number free images. Generation allows not to use the services of an expensive artist.
2) Design visualization. Can be useful when you need a quick design assessment, such as the interior of a room.
3) Correction of existing images. If an existing sample is damaged or requires modifications, image generation will help to correct inaccuracies.

Following the above arguments, we can conclude that the relevance of this technology is constantly growing.

The study analyzes the most common method of image generation using artificial intelligence (AI) with text queries. In the course of the analysis give a detailed description of AI as the main generation tool. The description considers the basic concepts of this technology, such as neural networks and machine learning

**Pattern recognition.** Pattern recognition technology solves the problem of identifying objects depicted in an illustration. The simplest implementation of such a task is text recognition.

The letter recognition algorithm is a long sequence of conditions like *"If the pixels are stacked in two circles on top of each other, it is either the number eight, the Cyrillic letter 'B' or the Latin letter 'B'"*. A number of similar conditions are applied sequentially a large number of times. The result of these transformations is electronic text.

The same task, but with an increased number of conditions, can also recognize objects that are more complex. It can be handwritten letters or images in real photos and drawings. For example, a modern phone is able to find faces to focus on. To do this, it applies pattern recognition using a set of over one million conditions [1].

There are many ways of quick applying such condition lists. One of the leading technologies is neural networks. Many algorithms for working with neural networks have been invented, including so-called *large text models*. A bright representative of

this family is the *GPT model*, which is capable of generating text. The concept behind using models is straightforward: rather than inputting image pixels, the input text's character codes are fed into the system, and character codes of the resulting text, which continues the input text, are obtained from the outputs. The network generates a continuation, and the algorithm selects formulas in order to ensure that this continuation aligns with the input text. [1].

**Image generation.** The process of generating images follows a similar concept. The difference is that a neural network generates an image instead of text. To begin generating an image, the neural network relies on textual descriptions. To comprehend these descriptions, a specific process is employed, which converts the text into a numerical format that the machine can understand. A *natural language processing (NLP)* model activates this process.

For example, a user enters a text query into an image generator: "white clouds in a blue sky". The NLP model converts this text into a numeric format that captures different elements: "white", "clouds", "blue" and "sky". It then determines the relationships between them. This numerical representation acts as a navigational map for an AI-based image generator [2].

**Artificial intelligence and its types.** Artificial intelligence (AI) is a field of information technology in which special computer programs are developed. The programs solve problems that require imitation of human thinking activity. Such programs can also summarize information and draw conclusions based on it. There are several types of artificial intelligence:

- *artificial narrow intelligence, ANI*. It is a hardware-software complex specializing in one specific area;
- *artificial general intelligence, AGI*. It is a hardware-software complex whose intelligence resembles human intelligence;
- *artificial superintelligence, ASI*. Intelligence that is superior to human intelligence in most areas [3].

The main concepts in AI systems are neural networks and machine learning. A neural network is a mathematical model and its implementation in the form of a program that recreates the work of a human neural network in a simplified form. Machine learning is a set of special algorithms that embody the key property of neural networks - the ability to self-learn based on the received data. [3].

**A neuron structure.** A neural network is similar to the brain from two points of view:

- knowledge enters the neural network from the environment and using in the learning process;

- connections between neurons, called synaptic weights, are used to accumulate knowledge.

Artificial neural network is a simplified model of the biological neural network of the human brain. Similarly, to the biological one, it accepts information from the outside and processes it, and then, based on the results obtained, performs the actions embedded in it. The basis of a neural network is a neuron. The simplified structure of a biological neuron is shown in Figure 1.

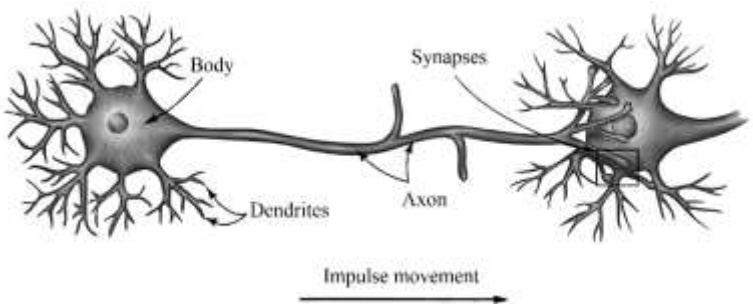


Figure 1 – The simplified structure of a biological neuron

A biological neuron is a nerve cell consisting of a *body* (where input signals are processed), *dendrites* (receive impulses from neurons in the network earlier than the neuron in question) and an *axon* (along which the neuron transmits impulses to other neurons in the network).

The artificial neuron is designed in the same way as a biological neuron. Its simplified model is shown in Figure 2.

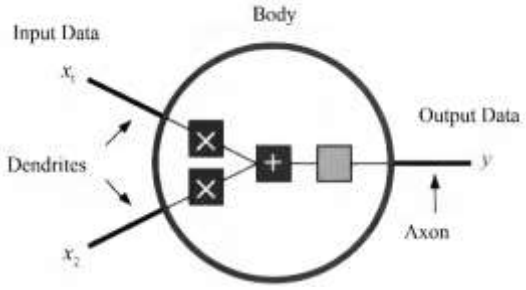


Figure 2 – The simplified structure of a artificial neuron

The neuron body also receives input data through *dendrites* $(x_1, x_2)$, processes them and outputs the result through the *axon* $(y)$. Neurons in the network interact through *synapses* - the places where the output axon of one neuron connects with the dendrites of another. An impulse transmitting through it is forming in the body of the neuron, which acts as an adder that receives all input impulses. After receiving the signals, it weighs them and passes them to some activating function, and then decides whether to continue running the impulse along the axon. Decision making consists of the following: if the total input pulses exceed some predetermined threshold, the output pulse is triggered. [3].

To implement an artificial neuron, a mathematical model realizing its basic functions is required. An example of such a model is shown in Figure 3.

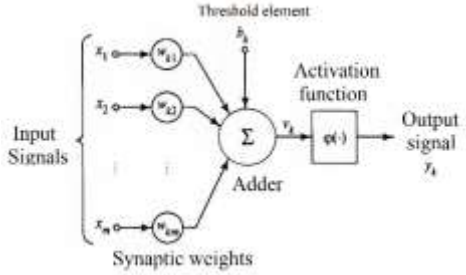


Figure 3 – Example of a mathematical model of an artificial neuron

The input of a mathematical neuron receives some number of input parameters: $x_1, x_2, \ldots x_m$, each of which has a weight: $w_{k_1}, w_{k_2} \ldots w_{k_m}$. The neuron body contains an adder which multiplying each element by some valid weight coefficient, after which the final sum is formed. As a result, we obtain the sum of products of input signal values by their weights:

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \quad (1)$$

$$y_k = \varphi(u_k + b_k) \quad (2)$$

The result is passed to the activation function $\varphi(\cdot)$. Before leaving the body of the neuron, the value of the function is checked: if it is above the threshold, a one is formed at the exit from the body. In this case, the neuron is activated and the signal is transmitted through the axon. If the value of the function is below the threshold, the output value is zero, and the neuron is not activated. The *threshold element* $(b_k)$ reflects the increase or decrease of the input signal applied to the activation function [3, 4].

**Neural network architecture.** Three main classes of neural network architectures are distinguished.

*Single-layer feed-forward networks.* In single-layer networks, signals from the input layer are fed to the output layer. It performs the necessary calculations, the results of which are immediately fed to the outputs. In counting the number of layers, source nodes are not taken into account as they do not perform the computations. Figure 4 shows the structure of such a network [3].

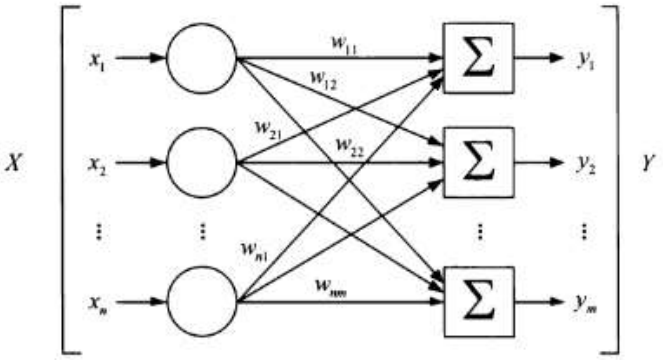


Figure 4 – Single-layer network structure

*Multilayer feed-forward networks.* This class is characterized by the presence of one or more hidden layers, the nodes of which are called hidden neurons. The function of hidden neurons is to mediate between an external input signal and the output of the neural network.

The source nodes of the input layer of the network generate an input vector, which is the input signal to the neurons of the second layer. The output signals of the second layer are the input signals to the third layer, and so on. An example of such a network is shown in Figure 5.

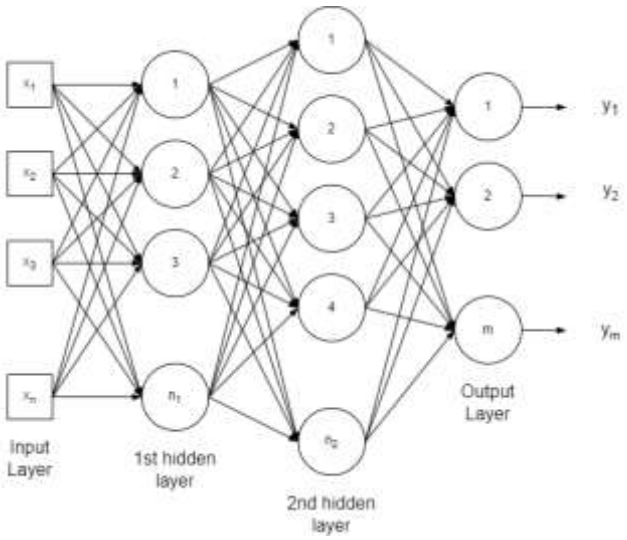

Figure 5 – Multilayer network structure

The above example network is generally called an $m - h_1 - h_2 - q$ network, where $m$ is the input layer, $h_1, h_2$ are the hidden layers, $q$ is the output layer. It is also *fully connected* since since all nodes of each layer are connected to all nodes of other layers. If some links are missing, then such a network is called *non-fully connected* [3, 4].

*Recurrent networks.* A recurrent neural network is distinguished from a feedforward network by the presence of at least one feedback. The presence of feedbacks has a direct impact on the learning ability and performance of such networks. Also, feedback implies the use of unit delay elements, which leads to nonlinear dynamic behavior if the network contains nonlinear neurons [4]. The structure of such a network shows in Figure 6.

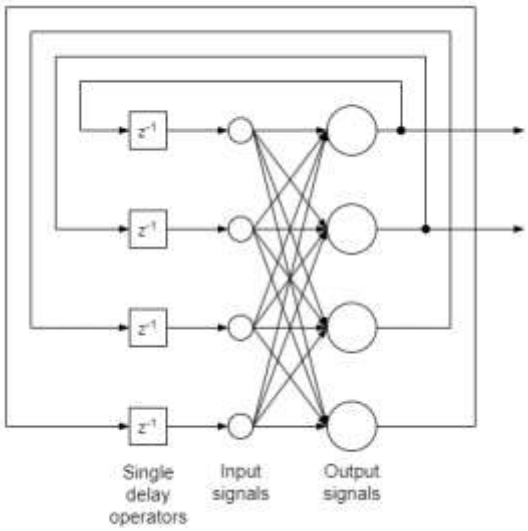

Figure 6 – Recurrent network structure

Machine learning. One of the most important properties of neural networks is their ability to learn from incoming data and experience.

The process definition above assumes the following sequence of events the NN receives stimuli from the external environment; as a result, the free parameters of the

neural network change; after changing the internal structure, the neural network responds to the stimuli in a different way.

Such a list of learning rules is called a learning algorithm. Learning algorithms differ in the way of tuning the synaptic weights of neurons [3, 4, 5].

There are several types of learning: supervised learning, unsupervised learning and reinforcement learning.

*Supervised machine learning.* This technique is the most common. A type of neural network training in which the weights of neural connections are selected so that the answers at the output of the network are minimally different from the already prepared correct answers. The essence of this approach is that test signals from the training sample feeding to the input of the network, and then the obtained answer is compared with the known correct one. With the help of special algorithms, the weights of links of the neural network are changed, one more time it gives a test input signal, and then again its answer is compared with the standard. This process repeating until the network begins to respond correctly with acceptable accuracy.

Supervised machine learning dividing into two subcategories: regression and classification. Regression involves drawing a line through a set of data points in such a way that it matches the overall shape of the data as closely as possible. Classification uses predicting categories of samples based on their attributes [3, 5].

*Unsupervised machine learning.* Type of neural network training in which the network independently classifies input signals. At the same time, the correct output signals not showing to it. Unsupervised machine learning is used when there is no possibility to provide correct answers to the input signals. Then the entire training sample consists of a set of input signals, and the network independently parses the input material. [3, 4, 5].

*Reinforcement learning.* Uses trial and error to train algorithms and create models. During the training process, algorithms operate in specific environments and then provide with feedback following each outcome. Much like how a child learns, the algorithm slowly begins to acquire an understanding of its environment and begins to optimize actions to achieve particular outcomes. For instance, an algorithm may be optimized by playing successive games of chess, which allow it to learn from its past success and failures playing each game.

Reinforcement learning usually uses to create algorithms that must effectively make sequences of decisions or actions to achieve their aims, such as playing a game or summarizing an entire text [4, 5].

**Machine learning rules.** There are many learning algorithms, let us consider one of them as an example. Learning based on error correction. To explain this rule, consider the case of the simplest neuron $k$, which is the only computational node of the output layer of a feed-forward network (Fig. 4, Fig. 5). Neuron $k$ drives by a signal vector $x(n)$ produced by one or more hidden layers of neurons, which in turn receive information from an input vector that passes to the initial nodes of the neural network. By $n$ means the step number of the process of iteratively adjusting the weights of neuron k. The output of neuron k is denoted as $y_k(n)$. It is compared with the desired output, denoted $d_k(n)$. As a result, we obtain the error signal$e_k(n)$:

$$e_k(n) = d_k(n) - y_k(n). \quad (3)$$

The error signal initializes a control mechanism whose goal is to apply a sequence of adjustments to the weights of neuron k. The changes aim at step-by-step approximation of the output signal $y_k(n)$ to the desired one $d_k(n)$, achieved by minimizing the cost function or performance index $E(n)$, defined as follows:

$$E(n) = \frac{1}{2} e_k^2(n), \quad (4)$$

where $E(n)$ – current error value.

The stepwise adjustment of the weights of neuron $k$ continues until the system reaches a steady state [4].

**Conclusion.** As a result of the work, the subject area is analyzed, in which the most popular method of image generation - with the help of artificial intelligence based on text queries - was considered.

A detailed analysis of artificial intelligence, as well as neural networks and machine learning as the main tools for generation is performed. A mathematical model of the structure of an artificial neuron is considered. It defines that it consists of dendrites, body and axon. The functioning of the neuron analyzing: the input of the neuron receives a certain number of input parameters, each of which has a weight. Types of neural networks are also considered: single-layer, multilayer and recurrent. To understand machine learning, one of the learning algorithms is selected and described.

## References


1. Капранов А. Как работают нейронные генераторы картинок / А. Капранов. – Текст: электронный // Хабр. Сообщество IT-специалистов. : [сайт]. – 2022. – 8 авг. – URL: https://habr.com/ru/articles/681310/

2. Лупиков О. Как нейросети создают изображения / О. Лупиков. – Текст: электронный // Интернет-издание о бизнесе, стартапах, инновациях, маркетинге и технологиях.: [сайт]. – 2023. – 22 сент. – URL: https://vc.ru/design/846807-kak-neyroseti-sozdayut-izobrazheniya?ysclid=lr1y7va66r768197647

3. Постолит А.В. Основы искусственного интеллекта в примерах на Python. Самоучитель / А.В. Постолит. – 2-е изд., перераб. и доп.– СПб.: БХВ-Петербург, 2023. – 448 с.: ил.

4. Хайкин С. Нейронные сети: полный курс / С. Хайкин – 2-е издание. Пер. с англ. – М.: Издательский дом «Вильямс», 2006. – 1104 с.: ил.

5. Харбанс Р. Грокаем алгоритмы искусственного интеллекта / Р. Харбанс. – СПб.: Питер, 2023 – 268 с.: ил.

***Аннотация.*** *Целью работы является обоснование возможности применения нейронных сетей и машинного обучения в графических системах. Рассмотрены элементы искусственного интеллекта и основы его функционирования. Приведен их краткий анализ и объяснена математическая модель. Даны характеристика, виды и правила построения моделей. Намечены результаты дальнейших исследований.*

***Ключевые слова:*** *искусственный интеллект, машинное обучение, нейронные сети.*

**Сведения об авторах:**

Мулявин Дмитрий Евгеньевич – ст. гр. КСм-23, Донецкий национальный технический университет

Мальчева Раиса Викторовна – доцент, к. тех. н., доцент кафедры «Компьютерная инженерия», Донецкий национальный технический университет

Кушниренко Елена Николаевна – ст. преподаватель кафедры английского языка, Донецкий национальный технический университет