Application of neural network platforms for text-based image generation

Oleh Yasniy[1,†], Abdellah Menou[2,†], Andriy Mykytyshyn[1,†], Vitalii Kubashok[1,*,†] and Iryna Didych[1,†]

[1] *Ternopil Ivan Puluj National Technical University, 56, Ruska Street, Ternopil, 46001, Ukraine*
[2] *ASTI LAB, SAP+D School, Mohammed VI Polytechnic University, Benguérir, Morocco*

Abstract
This article takes a closer look at various aspects of neural networks and their use in the modern world. Neural networks are a powerful artificial intelligence tool with many advantages in numerous areas, from image processing to content generation. Using examples of specific neural networks, such as DALL-E 3, Midjourney, ImageFX, Adobe Firefly, and Leonardo, a small experiment was performed to compare different neural networks based on the results of their work in response to a specific query. The Copilot neural network on the DALL-E 3 platform showed excellent results in generating images and text, making it one of the most successful neural networks among the tested ones.

Keywords
artificial intelligence, neural networks, platforms

1. Introduction

Artificial neural networks, or neural networks, are one of the most exciting and dynamic areas in information technology. Over the past few decades, they have evolved from concept to reality, creating incredible opportunities for development and innovation in various industries [1-3].

Neural networks are based on imitating the human brain, where many artificial neurons interact, processing input data and generating answers based on it. This approach has proved extremely powerful, allowing neural networks to solve tasks previously considered impossible for computers [4].

This article analyses in detail the application of neural networks in various spheres of life, from education to medicine and art [5-6]. It also analyses the principles of neural networks and their advantages and disadvantages. Using examples of specific neural

BAIT'2024: The 1st International Workshop on "Bioinformatics and applied information technologies", October 02-04, 2024, Zboriv, Ukraine
* Corresponding author.
† These authors contributed equally.
oleh.yasniy@gmail.com (O. Yasniy); abdellahmenou1@gmail.com (A. Menou); mikitishin@gmail.com (A. Mykytyshyn); Vitali.kubash@gmail.com (V. Kubashok); iryna.didych1101@gmail.com (I. Didych)
0000-0002-9820-9093 (O. Yasniy); 0000-0002-0606-6020 (A. Menou); 0000-0002-2999-3232 (A. Mykytyshyn); 0009-0007-9449-883X (V. Kubashok); 0000-0003-2846-6040 (I. Didych)

networks, such as DALL-E 3, Midjourney, and others, it was explored how these technologies affect modern society and their potential prospects in the future.

Next, the article discusses practical cases of neural networks in creating text-based content, such as images. Finally, the main directions of development and challenges faced by researchers in this area are analysed.

The aim of this paper is to compare different neural networks such as DALL-E 3, Midjourney, ImageFX, Adobe Firefly, and Leonardo based on their performance in response to a specific query

2. Application of neural networks

Neural networks are among the most innovative technologies affecting various aspects of our lives. They are widely used in various industries and solve various tasks, from pattern recognition to market trends forecasting. Let's take a closer look at some of the most important areas of neural networks and provide specific examples [7-8]:

Medicine:

- *Disease diagnosis:* Neural networks are employed to analyse medical images, such as X-rays or magnetic resonance imaging (MRI) images, to detect signs of diseases such as cancer or heart disease. For example, neural networks can help to detect cancer in X-ray images of the gastrointestinal tract.
- *Treatment prediction:* Neural networks can analyse patients' medical data and medical history to predict the effectiveness of different treatments. For example, they can help doctors choose the most appropriate cancer treatment for a particular patient.

Finance:

- *Financial market prediction:* Neural networks are utilized to analyse financial data and predict trends in stock and currency markets. They can analyse large amounts of data, considering various factors such as economic indicators, political events, and social trends.
- *Fraud detection:* Using neural networks, financial institutions detect fraud and anomalous transactions. They can analyse large amounts of transactional data and detect unusual patterns that may indicate fraud.

Technology:

- *Pattern recognition:* Neural networks can recognise faces, driver's licences, car licence plates, etc. For example, neural networks can automatically recognise and identify faces in images or videos.
- *Autonomous systems:* In autonomous cars, neural networks analyse data from sensors, such as radars, cameras, and lidars, and make real-time decisions. They help the car react to its environment and avoid accidents.

Education:

- *Personalised learning:* Neural networks help to create individualised learning programmes and materials, taking into account the characteristics of each student. They analyse student performance, personality, and learning style data to provide an optimal learning experience. For example, AI platforms can recommend individualised tasks and materials for each learner according to their needs and abilities [9].

- *Adaptive testing:* Neural networks can create tests that adapt to the learner's level of knowledge. They analyse the learner's answers to previous questions and, based on this, determine subsequent questions difficulty level. This helps assess students' knowledge effectively and offer them the appropriate tasks to increase motivation and learning outcomes.

2.1. The principle of neural networks

Neural networks are complex mathematical models designed based on the properties of biological neural networks in the human brain. Their operating principles include the following aspects [10-13]:

1. Architecture of the neural network:

- *Layers of neurons:* Neural networks typically have several layers, including an input layer, internal (hidden) layers, and an output layer. The input layer receives data, the hidden layers process this data, and the output layer generates the final result.
- *Connections between neurons:* Every neuron in one layer is connected to every neuron in the next. These connections contain weights that determine how strongly the input signals influence the neuron activation.

2. Activation function:

- *Linear or nonlinear:* The activation function determines how the signal is transmitted through the neuron. Neurons can have a linear or nonlinear activation function. Non-linear activation functions, such as a relu or sigmoid, allow neural networks to model more complex non-linear relationships in the input data.

3. Training of neural networks:

- *Backpropagation:* This is a basic neural network training method in which the neurons weights are adjusted according to the error between the predicted and expected outcomes. The weights are updated in a way that reduces this error.
- *Method of changing weights:* Various methods for adjusting weights, such as the momentum method, Adam's optimisation, and others help neural networks learn faster and more efficiently.

4. Regularisation:

- *Reducing overfitting:* When the neural network becomes overly adapted to the training data, regularisation techniques such as dropout, L1 or L2 regularisation are applied to prevent overfitting.

5. Output layer and loss function:

- *Loss function:* This function determines how closely the neural network prediction matches the actual data. It can include a squared error for regression or cross-entropy for classification.
- *Activate the original layer:* The activation function of the output layer may differ depending on the task. For example, a sigmoid function is used for binary classification, while for multi-class classification, a softmax function is utilized.

Understanding these principles is the key to designing and optimising neural networks for various machine learning and artificial intelligence tasks. Neural networks can solve real-world problems and form the basis of innovative applications.

2.2. Generative artificial neural networks

Neural networks have a wide range of applications in various industries, including generating content from text to images, video, and sound. Let's take a closer look at each type of application [14]:

1. Generation from text to image (Midjourney, Leonardo.ai, DALL-E 3, ImageFX):

- *Principle of operation:* Such neural networks are trained to associate textual descriptions of images and generate corresponding images. The neural network first receives an input text description, which it then converts into the components of a vector representation. This vector is then passed through a generative model that creates the corresponding image.
- *Application:* This method is widely used in art, design, and creative projects. For example, it can automatically generate illustrations for articles or books based on description text.

2. Generation from text to video (Synthesia, Pika):

- *Principle of operation:* In this case, the neural network learns correlations between text descriptions and video scenes. It receives an input text describing a plot or sequence of events and generates a corresponding video. This requires processing a large amount of video and text data to train the model.
- *Applications:* This method can find applications in film, advertising, and social media content production. For example, it can automatically generate commercials or short films based on scripts.

3. Generation from text to audio (Sound of text):

- *Principle of operation:* In this case, the neural network converts a text description into an audio file. It can learn the relationship between the text and the acoustic properties of the sound, such as voice timbre, intonation, and speech rate. This can include speech synthesis or music generation from text.
- *Applications:* This method is useful for various tasks, including audiobooks, podcasts, audio ads, and speech synthesis for apps and services. For example, it can be used to automatically create audio ads from a textual description of a product.

These applications of neural networks have significant potential to create different types of content using textual data. Although they require significant computing resources and training data, they open up new opportunities for automating the creative process and creating innovative content.

2.3. Neural networks MIDJORNEY, LEONARDO, DALL-E, IMAGEFX

Midjourney neural network: Midjourney is an advanced neural network for image and graphics processing. It offers a wide range of capabilities, from image editing and restoration to creating graphic effects and filters. Midjourney allows for the automation of image processing processes, making the tasks of graphic designers and photographers easier.

Leonardo.ai: Leonardo.ai is a platform that uses neural networks and artificial intelligence to create various graphic solutions. It provides tools for creating artwork, logo

design, image processing, and more. Leonardo.ai allows users to quickly and efficiently create impressive graphic elements using artificial intelligence.

DALL-E 3, ImageFX: DALL-E 3 and ImageFX are innovative neural networks designed to generate images from text descriptions and apply various graphic effects to images, respectively. They can create complex, realistic graphic objects using artificial intelligence algorithms [15].

2.4. Advantages and disadvantages of neural networks

Neural networks have several advantages, making them a powerful tool in artificial intelligence. First, they are known for their high accuracy. This means that they can solve complex problems with great accuracy and efficiency. For example, neural networks are successfully used for pattern recognition, text classification, and data analysis.

Secondly, neural networks can automate many processes that would have previously required significant human involvement. They can quickly and efficiently perform routine tasks, such as data processing, image editing, or speech synthesis.

In addition, neural networks can learn from a large amount of data, allowing them to improve over time. They can adapt to new conditions and tasks, which makes them versatile tools for various industries and fields of activity. It is also important to note the flexibility of neural networks. They can be applied to different simple and complex tasks and adapt to different data types. This makes them useful for various applications, from medicine to finance, advertising, and art.

Despite their advantages, neural networks also have several disadvantages. For example, they require a large amount of data to train, and insufficient data can lead to poor model accuracy. In addition, training neural networks can be time-consuming and require significant computing resources.

Another drawback is the difficulty of interpreting the results. Neural networks are often considered "black boxes" because their decisions can be difficult to understand and explain. This can complicate decision-making processes and the implementation of models in practice. In addition, as the complexity of neural networks increases, so does the need for computing resources. Large and complex neural networks can require significant computing power to train and use, which can be difficult for many organisations and researchers.

3. Operation of neural networks in practice and comparison

To analyse the effectiveness of neural networks in improving individualised learning for students, a small experiment was used to compare different neural networks based on the results of their work in response to a specific query. This experiment became the basis for analysing the effectiveness and capabilities of neural networks in the context of the learning process.

The experiment used the following promo generated in a chat with the GPT artificial intelligence language model:

"Create a high quality, realistic image of Lesya Ukrainka wearing a traditional Ukrainian embroidered shirt, reflecting her modern spirit and cultural identity. She should be standing

in a busy place in modern Ukraine, against the backdrop of Khreshchatyk. In the background, the inscription "Ukraine" should be visible, embedded in the cityscape, emphasising the connection between the figure and the place. Every detail should be carefully reproduced."

Figure 1 shows the image built by Google's neural network, ImageFX.

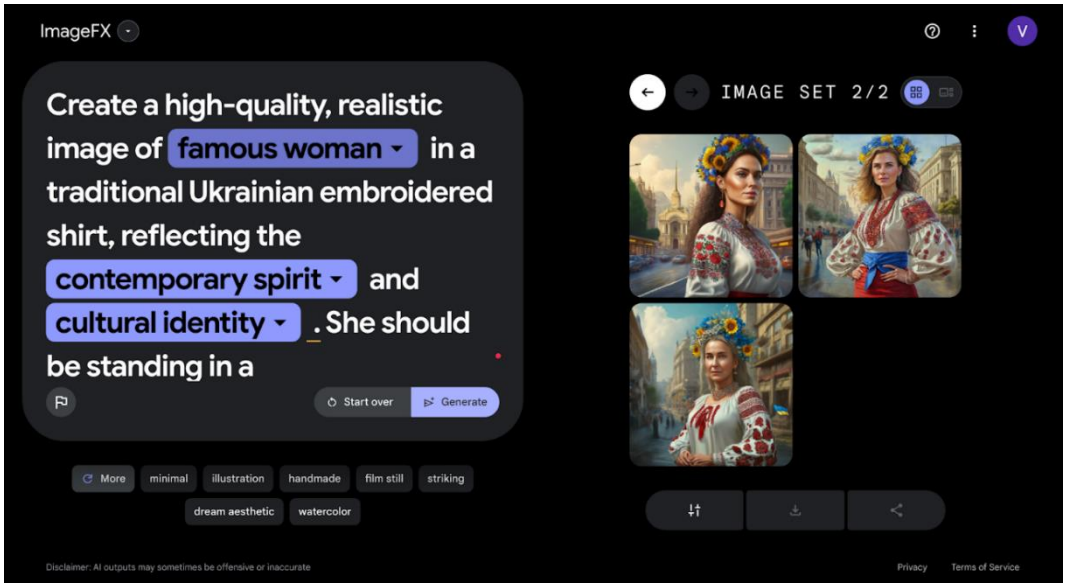

Figure 1: The image built by Google's neural network, namely ImageFX

ImageFX, a neural network from Google, provided quite beautiful images with good human detail, including Lesya Ukrainka in a Ukrainian embroidered shirt. However, there was a problem with the clarity of the background and the generation of text in the background, which made it difficult to create a realistic image.

Figure 2 shows the result of the Adobe Firefly neural network.

Figure 2: The image was built by the neural network from Adobe Firefly.

Firefly, a neural network from Adobe, also provided high-quality images. Still, it was not always possible to generate the city background and text, significantly affecting the created image realism.

Figure 3 shows the image generated by the Leonardo neural network.

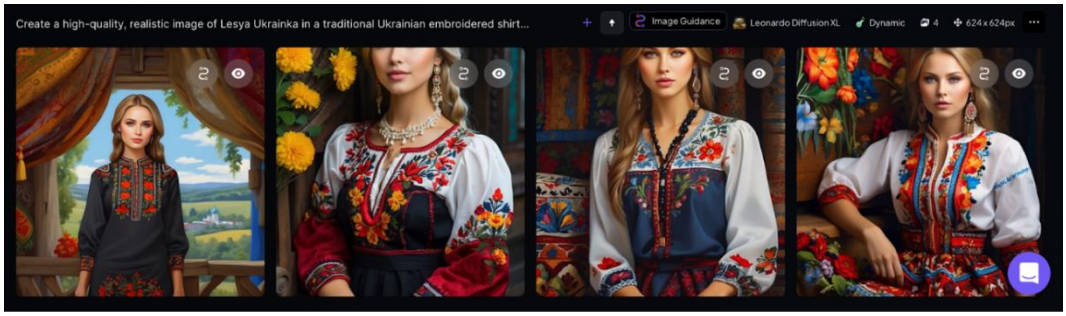

Figure 3: The image was built by the neural network from Leonardo

The Leonardo neural network did not provide the expected results, which may be due to its technical features and limitations.

Figure 4 shows the image built by Midjourney neural network.

Figure 4: The image was built by the neural network from Midjourney

Midjourney also failed to meet expectations, which indicates that this neural network likely has limitations in creating realistic images.

Figure 5 shows the image built by Copilot neural network on the DALL-E 3 platform.

Figure 5: The image was built by the Copilot neural network on the DALL-E 3 platform

The Copilot neural network on the DALL-E 3 platform demonstrated excellent results in generating images and text, making it one of the most successful neural networks among the tested ones.

The results show that neural networks can be an effective tool in improving personalised learning, but they have their limitations. Some neural networks, such as Copilot on the DALL-E 3 platform, have shown good results in generating images and text, but they are imperfect and require further improvement.

Table 1 compares different neural networks such as DALL-E 3, Midjourney, ImageFX, Adobe Firefly, and Leonardo based on their performance on a specific query.

Table 1
Parameters for comparing neural networks

	ImageFX	Adobe Firefly	Leonardo	Midjourney	DALL-E 3
Representation of Lesya Ukrainka's face	+	+	+	+	+
Embroidery details	+	+	+	+	+
Clarity of the background	-	-	-	+	+
Generation of text in the background	-	-	-	-	+

4. Conclusions

The article describes different types of neural networks, including deep neural networks, delving into their structure and principles of operation. Popular neural network

architectures and their application in solving various tasks are also considered. In addition, the advantages and disadvantages of using neural networks are discussed. In particular, the Copilot neural network on the DALL-E 3 platform demonstrated excellent results in generating images and text, making it one of the most successful neural networks among the tested ones.

In conclusion, it is important to emphasise that neural networks have great potential in many fields and continue to evolve and improve daily. They have the potential to change how we work, learn, and communicate, and it is important to continue exploring their capabilities to maximise their usage in the future.

References

[1] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

[2] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[3] Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85-117.

[4] Deng, L., & Yu, D. (2014). Deep learning: methods and applications. Foundations and Trends® in Signal Processing, 7(3–4), 197-387.

[5] Karpathy, A., & Fei-Fei, L. (2015). Deep visual-semantic alignments for generating image descriptions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3128-3137).

[6] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI blog, 1(8), 9.

[7] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 5998-6008).

[8] Brock, A., Donahue, J., & Simonyan, K. (2019). Large scale GAN training for high fidelity natural image synthesis. In International Conference on Learning Representations.

[9] Yasniy, O., Mykytyshyn, A., Didych, I., Kubashok, V., & Boiko, A. (2023). Application of artificial intelligence to improve the work of educational platforms. In ITTAP (pp. 605-609).

[10] Haykin S. Neural Networks - A Comprehensive Foundation - Simon Haykin.pdf. McMaster University, Hamilton, Ontario, Canada, 2006. P. 823.

[11] N. Richard: Applied regression analysis, third ed., John Wiley & Sons, New York, 1998

[12] Philip D. Wasserman. Neural Computing: Theory and Practice, New York: Coriolis Group (Sd), 1989

[13] K. Gurney: An introduction to neural networks, first ed., Taylor & Francis Group, London, 1997

[14] Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2019). A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 4401-4410).

[15] OpenAI. (2023). DALL-E 3: Creating Images from Text. Retrieved from https://openai.com/dall-e-3