

**PROBLEMS OF USING AI-GENERATED IMAGES
AND AN OVERVIEW OF POSSIBLE SOLUTIONS**

Perederii I.A.

e-mail: illia.perederii@nure.ua

Scientific supervisor – Ovcharenko D. R.

Kharkiv National University of Radio Electronics, Dep. CRETISS

With the development of image generation technologies based on artificial intelligence (AI), the possibilities of automated creation of visual content are significantly expanding. In addition, new ethical and legal challenges arise, such as issues of copyright, transparency of learning algorithms, and responsibility for the use of generated content. Special attention should be paid to issues of cyber security, in particular, prevention of misuse in the field of digital identification and protection of personal data. This work analyzes the key stages of image generation development using AI, current challenges, and possible ways to overcome them.

The development of image generation with the help of artificial intelligence (AI) can be divided into several key stages, characterized by changes in the architecture of neural networks, data processing algorithms and methods of training models.

Before the advent of deep neural networks, image generation was based on computer vision algorithms and statistical methods. These approaches had low availability, significant quality limitations, and difficulty in reproducing realistic visual data. The main areas of research during this period were statistical image analysis methods, neural network coding and image reconstruction algorithms.

The next stage of development was marked by the appearance of Generative Adversarial Networks (GANs), proposed by Ian Goodfellow in 2014. GANs consisted of two main components: a generator (which creates new images) and a discriminator (which evaluates their realism). This technology made it possible to significantly improve the quality of synthesized images and became the basis for creating realistic portraits of people who do not exist. Among the main achievements of this era, we can highlight: StyleGAN (2018, NVIDIA, a method for creating photorealistic faces with deep control over stylization) and BigGAN (2018, Google DeepMind, improving the detail and variety of generated images).

GANs have been replaced by diffusion models based on iterative learning through the gradual removal of noise from a random distribution. This approach made it possible to significantly improve the detail and accuracy of image generation. Significant breakthroughs: DALL·E (2021, OpenAI) is the first model capable of creating images based on textual requests from users; CLIP (2021, OpenAI) – a model that optimized the relationship between text and images, significantly improving the interpretation of requests; Stable Diffusion (2022,

Stability AI) is an open-source image generation software that has made this technology available to a wide range of users.

At the current stage, generative neural networks are widely used in combination with video, 3D graphics and integration into design tools.

The main trends are: DALL·E 3 (2023) – a model with improved understanding of text queries; MidJourney (2023-2024) – generation of artistic and photorealistic images; Generative Fill in Photoshop (2023) – integration of generative AI into graphic editors; Google Imagen, Runway ML, OpenAI Sora – extension of video generation based on text queries.

However, the rapid development of image generation technologies, as well as the gradual transition to an increasingly high level of photorealism, create numerous challenges that are already the subject of discussions in scientific circles today and may cause significant ethical, legal and social conflicts in the future

Let's take a look at the main ethical, legal and cybersecurity challenges posed by AI-generated imagery.

Copyright and legal aspects: lack of a clear legal framework regarding the authorship of AI-generated images, since the neural network is not subject to copyright; discussions about the legality of using training datasets containing copyrighted images; lawsuits over the use of artists' original works to train generative models (e.g. Getty Images v. Stability AI).

Moral and ethical challenges: fake news and deepfakes – using AI to create fake images and videos that can be misleading; identity theft – generating realistic but fictitious faces for fraudulent schemes; illegal content – the use of generative models to create prohibited or unethical materials.

Cybersecurity issues: hacking and malicious use of generative models – the possibility of creating forged documents, manipulated images and content; phishing attacks – use of realistic images for deception and cybercrimes; malicious algorithms – the use of generative AI models to create dangerous or manipulative material.

Now let's take a look at the possible ways to overcome the aforementioned problems and challenges.

Legal regulation and standardization:

- Development of clear legislative norms regarding copyright for AI-generated content, including the definition of rights holders and licensing mechanisms.

- Introduction of global standards for the ethical use of AI, which would regulate restrictions on the creation of manipulative or harmful materials.

Identification and labeling of AI-generated content:

- Implementation of digital watermarks and metadata to label generated images, helping to distinguish them from real photographs.

- Use of algorithms to detect fake images and deepfakes, along with the development of independent tools for verifying the authenticity of content.

Strengthening cybersecurity:

- Development of mechanisms to prevent the use of AI in fraudulent schemes, such as the creation of forged documents or compromising images.
- Application of AI to detect and neutralize threats associated with visual content manipulation.

Educational initiatives and public awareness:

- Increasing the level of digital literacy among the population to foster understanding of the opportunities and risks of generative AI.
- Incorporation of courses on AI ethics into educational programs for professionals working in the field of digital technologies.

References:

1. Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers Staphord Bengesi¹, Hoda El-Sayed¹, Md Kamruzzaman Sarker¹, Yao Houkpati¹, John Irungu³ and Timothy Oladunni²
2. Alberto Romele Images of Artificial Intelligence: Blind Spot in AI Ethics *Philosophy & Technology* 35(1) March 2022. Available: https://www.researchgate.net/publication/358217205_Images_of_Artificial_Intelligence_a_Blind_Spot_in_AI_Ethics
3. Taeshik Shon, Subin Oh, “Cybersecurity Issues in Generative AI”, 2023 International Conference on Platform Technology and Service (PlatCon), August 2023. Available: https://www.researchgate.net/publication/374171491_Cybersecurity_Issues_in_Generative_AI