Kostyuchenko Artem, master's student,
Loktikova Tamara, Senior Lecturer at the Department of Software Engineering,
Lysogor Iurii, Senior Lecturer at the Department of Software Engineering
Zhytomyr Polytechnic State University

DEVELOPMENT OF A NEURAL MODEL FOR IMAGE AND VIDEO RESOLUTION ENHANCEMENT

With the development of artificial intelligence and deep learning, effective methods for improving the quality of digital content have emerged. One of these areas is the creation of neural models to increase the resolution of images and videos, known as Super-Resolution.

Super-Resolution technology can restore details and improve the clarity of low-quality content and has a wide range of applications, from medicine to video games.

Modern neural networks use sophisticated algorithms to reconstruct a high-quality image from an input low-resolution image. The basic principle of these models is to train a neural network on a large set of paired images: a low-resolution image and its high-quality counterpart.

This principle allows the model to learn to restore fine details that are lost when the resolution is reduced.

In addition to traditional methods, approaches based on generative adversarial networks (GANs) are becoming increasingly popular [1]. These methods create more realistic and detailed images by using a discriminator that distinguishes between generated models and real photos, prompting the generator to improve the quality of the output.

Today, there are many networks, but let's mention the most famous ones. SRCNN, EDSR, ESRGAN are networks that have significant success in restoring details and improving image quality but also have certain limitations that complicate their practical use [2].

Models based on SRCNN [3] and EDSR [4] networks focus only on improving spatial resolution, without taking into account contextual details and textures, which can lead to smoothed or less natural results.

Generative models such as ESRGAN, although they can produce realistic high-frequency details, often lead to artifacts that can distort the original image, making it impossible to use, for example, in medicine.

The problems of these neural networks allow us to think about creating our own improved model that can reduce the existing shortcomings and provide a better resolution.

The proposed model is based on EDSR and VSRN network architectures with partial integration of ESRGAN principles.

The developed model has the ability to increase the resolution of both images and video, with support for processing on both the central processing unit (CPU) and graphics processing unit (GPU) to optimize performance.

The system supports automatic optimization for different hardware configurations: the ability to choose between fast mode for less powerful devices or high-quality mode for high-performance graphics cards.

The model supports noise and compression artifact removal, which is useful for processing old or low-quality materials.

The model is equipped with intuitive software that allows you to easily upload files, adjust processing parameters, and save improved results. It is possible to change the degree of magnification (2×, 4×, 8×) and choose between different super-resolution algorithms to achieve the optimal balance between detail and speed.

The developed model uses the TensorFlow library from Google and the Python programming language with the corresponding libraries for image and video processing. To implement the graphical interface, the PySide6 library is used in combination with the declarative language QML.

References:

1. Generative Adversarial Network.
URL: https://www.geeksforgeeks.org/generative-adversarial-network-gan/.
2. Image Super-Resolution: A Comprehensive Review.
URL: https://www.digitalocean.com/community/tutorials/image-super-resolution.
3. SRCNN (Super Resolution) Review.
URL: https://arxiv.org/abs/1904.07470
4. Enhanced Deep Residual Networks for single-image super-resolution. URL: https://keras.io/examples/vision/edsr/.