

УДК 004.932

О. Г. Байбуз<sup>1</sup>, М. Г. Сидорова<sup>1</sup>, А. Д. Сінєгіна<sup>1</sup>,  
Ю. Д. Сінєгіна<sup>1</sup>, О. В. Лапець<sup>2</sup><sup>1</sup>Дніпровський національний університет імені Олеся Гончара<sup>2</sup>Університет митної справи та фінансів

## ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ СТИЛІЗАЦІЇ ТА КОЛОРИЗАЦІЇ ЗОБРАЖЕНЬ

Розроблено інформаційну технологію стилізації та колоризації зображень з можливістю детального налаштування параметрів на основі застосування штучних згорткових нейронних мереж та підходу transfer learning. Запропоновано метод підвищення чіткості отримуваних після перетворення зображень.

*Ключові слова:* інформаційна технологія, стилізація зображень, колоризація, згорткові нейронні мережі, підвищення деталізації і чіткості зображень.

Разработана информационная технология стилизации и колоризации изображений с возможностью детальной настройки параметров на основе применения искусственных сверточных нейронных сетей и подхода transfer learning. Предложен метод повышения четкости получаемых после преобразования изображений.

*Ключевые слова:* информационная технология, стилизация изображений, колоризация, сверточные нейронные сети, повышение детализации и четкости изображений.

Nowadays due to the efficient algorithms and architectures of convolutional neural networks, as well as the possibility of applying the transfer learning approach, there are significant advances in various areas of automated image processing from accurate object detection to more creative areas such as colorization or style transfer. Given the practical significance of these tasks in various fields of human activity, it is important to create information technology and software to ensure effective and high-quality processing of visual content, as well as improving existing algorithms and approaches to obtain results that best meet expectations.

In this paper the information technology of stylization and colorization of images with a possibility of both automated, and thin adjustment of parameters according to user preferences has been created. For the stylization task, it is possible to save the original colors of the content image, the transfer of style in

this case is performed only in the brightness channel. This is motivated by the observation that visual perception is much more sensitive to changes in brightness than in color. Improving the detail and sharpness of the resulting images has been done by reducing the noise by the method of Total Variation, which allows, reducing the noise, to keep the edges and contours of the image unchanged. In addition, the proposed technology realizes the possibility of increasing the image resolution in the context of a stand-alone task, and it is shown that using it as a preliminary step of colorization can improve the clarity of images and the quality of the results. The proposed technology is implemented in the author's software, using Python programming language and the Tensorflow library.

*Keywords:* information technology, image stylization, colorization, convolutional neural networks, increasing the detail and sharpness of images.

**Вступ.** Останнім часом завдяки стрімкому розвитку ефективних алгоритмів та архітектур згорткових нейронних мереж, а також можливості застосування підходу transfer learning [1], спостерігаються значні досягнення у різних галузях автоматизованої обробки зображень – від точної детекції об'єктів до більш творчих напрямів, значне місце серед яких займають перенесення художнього стилю [2] та колоризація [3]. Зважаючи на практичну значимість цих задач у різноманітних галузях діяльності людини, актуальним є створення інформаційних технологій та програмних засобів для забезпечення ефективної та якісної зазначеної обробки візуального контенту, а також удосконалення існуючих алгоритмів та підходів для отримання результатів, що найкраще відповідають очікуванням.

**Постановка задачі.** Метою роботи є створення інформаційної технології колоризації та стилізації зображень з можливістю як автоматизованого, так і тонкого налаштування параметрів за уподобаннями користувача, а також підвищення деталізації і чіткості отримуваних зображень за допомогою попереднього збільшення роздільної здатності та зниження шуму методом повної варіації.

**Основний матеріал.** Нейронна передача стилю полягає у застосуванні стилю зображення-зразка до цільового зображення при збереженні вмісту цього цільового зображення [2]. В даному контексті під стилем, в основному, маються на увазі текстури, колірна палітра і візуальні шаблони в різних просторових масштабах, а під вмістом – високорівнева макроструктура зображення. Для виявлення стилю і вмісту у роботі застосовувалася згорткова нейронна мережа VGG [4] з усереднювальним пулінгом замість максимізаційного.

*Представлення вмісту.* Нехай  $\vec{p}$  та  $\vec{x}$  – це початкове зображення та зображення, що згенеровано,  $P^1$  та  $F^1$  – відповідне представлення ознак

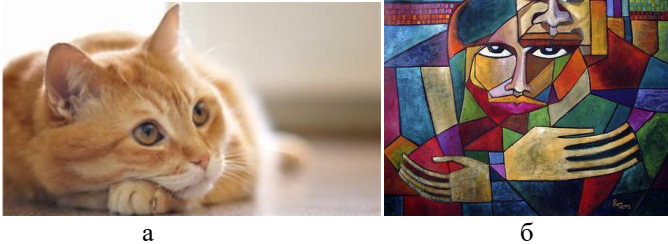
в шарі  $l$ . Тоді втрата квадратичної помилки між двома представленнями ознак обчислюється за формулою:

$$\mathcal{L}_{\text{content}}(\vec{p}, \vec{x}, l) = \frac{1}{2} \sum_{i,j} (F_{ij}^l - P_{ij}^l)^2.$$

*Представлення стилю.* Нехай  $\vec{a}$  та  $\vec{x}$  – це початкове зображення та зображення, що згенеровано,  $A^l$  та  $G^l$  – відповідні представлення стилю в шарі  $l$ . Загальна втрата стилю дорівнює:

$$\mathcal{L}_{\text{style}}(\vec{a}, \vec{x}) = \sum_{l=0}^L \omega_l E_l, \quad E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{ij}^l - A_{ij}^l)^2.$$

Розглянемо приклад. На рис. 1 наведено зображення, які бралися в якості зображення вмісту (рис. 1, а) та зображення стилю (рис. 1, б).



**Рисунок 1 – Приклади зображень, які бралися в якості: а – зображення вмісту; б – зображення стилю**

*Перенесення стилю.* Щоб перенести стиль із зображення  $\vec{a}$  на зображення  $\vec{p}$ , синтезується нове зображення, яке одночасно буде співпадати з представленням вмісту  $\vec{p}$  та представленням стилю  $\vec{a}$ . Таким чином, ми спільно мінімізуємо відстань представлення ознак зображення білого шуму з представленням вмісту фотографії в одному шарі та представленням стилю зображення, визначеного на ряді шарів згорткової нейронної мережі. Функція витрат, яку ми мінімізуємо:

$$\mathcal{L}_{\text{total}}(\vec{p}, \vec{a}, \vec{x}) = \alpha \mathcal{L}_{\text{content}}(\vec{p}, \vec{x}) + \beta \mathcal{L}_{\text{style}}(\vec{a}, \vec{x}),$$

де  $\alpha$  та  $\beta$  – це вагові коефіцієнти для реконструкції вмісту та стилю відповідно. Співвідношення  $\frac{\alpha}{\beta}$  береться або  $1 \times 10^{-3}$ , або  $1 \times 10^{-4}$ .

Результат роботи можна побачити на рис. 2, а.

*Збереження початкових кольорів при передачі стилю.* Цей метод полягає в тому, щоб виконати передачу стилю лише в каналі яскравості. Це мотивовано спостереженням, що візуальне сприйняття

набагато чутливіше до змін у яскравості, ніж у кольорі. Спочатку зображення вмісту переводиться з кольорового простору RGB до кольорового простору YIQ. В цьому кольоровому просторі компонента Y представляє яскравість, а компоненти I та Q – кольорові складові. Компонента яскравості містить зображення у відтінках сірого, а інші дві компоненти містять інформацію для відновлення потрібного кольору. Інформація про кольори зображення вмісту представлена його каналами  $I_C$  та  $Q_C$ , значення яких на початку вилучаються та зберігаються. Потім до зображень вмісту та стилю застосовується нейронний алгоритм передачі стилю для отримання вихідного зображення. Після того як отримано вихідне зображення, воно також переводиться з простору RGB до простору YIQ (матиме канали  $Y_G, I_G, Q_G$ ). Тоді канали зображення вмісту, що відповідають за кольори, поєднуються з каналом яскравості зображення, що було згенеровано, а саме кольорове вихідне зображення має значення каналів  $Y_G, I_C$  та  $Q_C$ , та вже це зображення остаточно переводиться до простору RGB [5]. Цей метод досконало зберігає кольори зображення вмісту. Результат можна побачити на рис. 2, б.



**Рисунок 2 – Результати роботи стилізації зображень: а – з повним перенесенням стилю; б – зі збереженням початкових кольорів**

*Зниження шуму методом повної варіації.* Повна варіація представляє собою суму абсолютних різниць для сусідніх значень пікселів у вхідних зображеннях. Цю функцію застосовано як функцію помилки під час оптимізації для зменшення шуму та згладжування зображення, оскільки вона забезпечує просторову гладкість. Таким чином, алгоритм TVD (Total Variation Denoising) [6] зсуває зображення на один піксель по осях  $x$  та  $y$ , обчислює різницю з початковим зображенням, приймає абсолютне значення, щоб забезпечити різницю додатним числом, та сумує по всіх пікселях на зображенні. Таким чином і створюється функція помилки, яка мінімізується, щоб уникнути різких переходів у значеннях пікселів. Ця технологія

видалення шуму має переваги над простими методами, такими як лінійне згладжування або медіанна фільтрація, які зменшують шум, але в той же час згладжують краї та контури зображення. Для зображень total variation loss обчислюється за формулою:

$$V(y) = \sum_{i,j} |y_{i+1,j} - y_{i,j}| + |y_{i,j+1} - y_{i,j}|,$$

де  $y$  – вхідне зображення.

Для колоризації зображень застосовувався класичний нейромережовий підхід з використанням архітектури U-net [7]. Проте для збільшення чіткості отримуваних зображень було застосовано попереднє збільшення роздільної здатності зображення, що також дозволило підвищити якість результатів (табл. 1, рис. 3).

Таблиця 1

### Оцінка якості навчання

	Колоризація		Збільшення роздільної якості	
	train	test	train	test
loss	0,00064	0,004057	0,000853	0,001077
accuracy	0,87670	0,685434	0,90350	0,842841

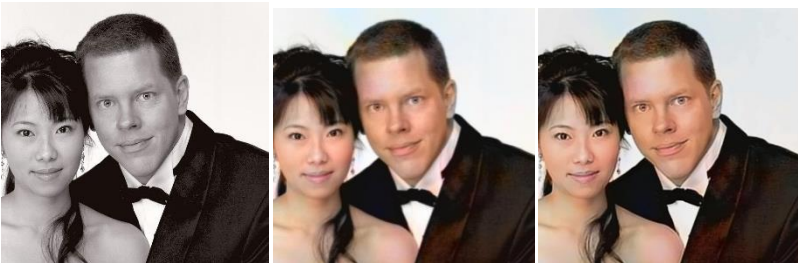


Рисунок 3 – Результати колоризації зображень (зліва направо: початкове чорно-біле зображення; звичайна колоризація; колоризація з попереднім збільшенням роздільної здатності зображення)

Запропоновану технологію реалізовано в авторському програмному забезпеченні, що написано мовою програмування Python, з використанням бібліотеки tensorflow, середовище розробки – JetBrains PyCharm Community Edition. Усі експерименти проводилися з використанням tensorflow-gpu на відеокарті NVIDIA GeForce 940MX (2GB).

**Висновки.** Створено інформаційну технологію стилізації та колоризації зображень з можливістю як автоматизованого, так і тонкого налаштування параметрів за уподобаннями користувача, в тому числі зі збереженням початкових кольорів зображення вмісту, а також підвищення деталізації і чіткості отримуваних зображень за допомогою попереднього збільшення роздільної здатності та зниження шуму методом повної варіації. Проведено ряд експериментів роботи програми на різних даних. Показано, що збільшення роздільної здатності зображення, як попередній крок колоризації, дозволяє підвищити чіткість зображень та якість отримуваних результатів.

### Бібліографічні посилання

1. Zhuang F. et al. A comprehensive survey on transfer learning. URL: <http://arxiv.org/abs/1911.02685>.
2. Gatys L. A., Ecker A. S., Bethge M. Image Style Transfer Using Convolutional Neural Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016. P. 2414–2423.
3. Bisht U., Patnaik T. Overview of Automatic Image Colorization Schemes. *International Journal of Advanced Engineering and Global Technology*. 2015. P. 1283–1287.
4. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations (ICLR)*. 2015. 14 p.
5. Gatys L. A., Bethge M., Hertzmann A., Shechtman E. Preserving Color in Neural Artistic Style Transfer. URL: <https://arxiv.org/pdf/1606.05897.pdf>
6. A. Makovetskii, Sergei Voronin An Efficient Algorithm for Total Variation Denoising. *Communications in Computer and Information Science*. 2017. P. 326–337.
7. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. *Proceedings of the Medical Image Computing and Computer-Assisted Intervention*. 2015. P. 234–241.

*Надійшла до редколегії 22.11.2020.*