

УДК 004.942:519.876

Застосування нейронної мережі для стилізованої обробки зображень

Колесник О. Ю., ORCID [0000-0002-8896-1142](https://orcid.org/0000-0002-8896-1142)e-mail: kolesnyk.oleksandra@gmail.comКафедра промислової електроніки www.kaf-pe.kpi.ua

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського» www.kpi.ua

Київ, Україна

Анотація—У статті розглянуто задачу обробки графічних зображень з використанням математичного апарату штучних нейронних мереж. Окремо розглядається структура згорткової нейронної мережі. Для вирішення задачі художньої обробки фотографії з використанням заданої стилістичної техніки використовуються два зображення – «картинка-контент» та «картинка-стиль», яка задає шаблон стилістики. При застосуванні різних стилістичних шаблонів можна сформувати фото з різною стилістикою – наприклад, у стилі імпресіонізму, абстракціонізму, тощо. Таким чином, робота представляє інтерес як з художньої, так і з дослідницької точки зору.

Бібл. 10, рис. 6.**Ключові слова**—обробка зображень; штучні нейронні мережі; згорткові мережі; згортка; передача стилю.

I. ВСТУП

Математичний апарат штучних нейронних мереж є потужним інструментом, що широко застосовується в задачах обробки та аналізу різнотипних даних. Серед цих задач важливе місце посідає задача розпізнавання об'єктів, яка у випадку необхідності обробки великих потоків даних (Big Data) та аналізу об'єктів зі складною структурою має беззаперечні переваги перед традиційними обчислювальними методами. Завдяки можливості навчання штучні нейронні мережі здатні більш ефективно вирішувати задачі розпізнавання та ідентифікації графічних образів у різних галузях – електроніці, біомедицині, географії, тощо. Додатково підвищується ефективність нейронних мереж за рахунок використання спеціалізованих пристроїв та мікросхем - нейрокомп'ютерів, нейрочипів, нейроконтролерів. На сьогоднішній день час існує багато типів архітектури штучних нейронних мереж, які в процесі обробки зображень здатні виявляти окремі деталі, налаштовувати контрастність, але архітектура згорткової нейронної мережі дозволяє вирішувати якісно нову задачу – художньої, або стилістичної обробки зображень, що дає можливість формування образів заданої стилістики. Згорткові нейронні мережі здатні знаходити певні задані об'єкти на зображеннях, а також виявляти важливі деталі цих об'єктів, тобто високорівневі ознаки. Ця здатність дозволяє розрізняти об'єкти між собою, визначати відстань між об'єктами та модернізувати зображення з подальшим збереженням високорівневих ознак. Так як штучні нейронні мережі подібні між біологічними нейронними мережами, дана робота розглядає певний шлях до алгоритмічного розуміння

того, як люди створюють та сприймають художню образність.

II. ПОСТАНОВКА ПРОБЛЕМИ

В даній роботі розглядається задача обробки фотографій та переносу певного заданого стилю на початкове зображення. Тобто на підставі наявних двох зображень – початкової фотографії та графічно-стилістичного контенту, який задає стилістику, формується нова стилізована фотографія. Ця стилізована фотографія залишається впізнаваною, але на неї накладаються певні структури, кольори та форми. При застосуванні різних графічних контентів можна сформувати фото з різною стилістикою – наприклад, у стилі імпресіонізму, абстракціонізму, тощо. Таким чином, робота представляє інтерес як з художньої, так і з дослідницької точки зору. Зокрема, важливою є розробка та реалізація алгоритму обробки зображень на базі математичного апарату штучних нейронних мереж.

III. АНАЛІЗ ЛІТЕРАТУРНИХ ДАНИХ

Нейрокомп'ютер — це пристрій переробки інформації на основі принципів роботи природних нейронних систем. Комп'ютер працює на основі навчання, а не програмування. Пристрій для вирішення задач прогнозування, діагностики, розпізнавання та інших.

Штучна нейронна мережа є системною моделлю біологічної нейронної мережі. Кожен штучний нейрон має декілька входів, на які подаються певні сигнали x_i , де $i=1,2,\dots,n$ – номер вхідного сигналу. Зв'язок між входом та нейроном називається синапсом. Синапс характеризується певною вагою – коефіцієнтом w_i . З сигналів, що надходять на входи нейрону



через синапси, формується зважена сума S , яка потім обробляється функцією активації $Y=F(S)$. Вихідний сигнал Y далі надходить через інший синапс з певною вагою на вхід наступного нейрону. Передача сигналу відбувається пошарово, причому, як правило, в структурі нейронної мережі є декілька прихованих шарів.

Згорткова нейронна мережа – структурний вид багатошарової штучної нейронної мережі, яка використовується для ефективної роботи з зображеннями. Метод базується на використанні математичної операції «згортка». До шарів в цій мережі застосовується операція згортки на вході, а результат передається до наступного шару. Така організація роботи мережі імітує реакцію біологічного нейрона на візуальні подразники.

Згортка виконується з допоміжною функцією, що задається кінцевими імпульсними характеристиками серії цифрових фільтрів, які послідовно накладаються на вхідний сигнал – зображення, що підлягає обробці. Ці фільтри є матрицями – зазвичай тривимірними наборами значень, що у просторі утворюють квадратну призму. При обробці зображень часто використовуються матриці з висотою та шириною, що дорівнюють 3 або 5. Глибина матриці відповідає кількості каналів зображення. Три кольорові канали (RGB) означають глибину, що дорівнює трьом.

Нейронна передача стилю – метод формування зображення шляхом об'єднання двох інших зображень. При цьому задається два зображення: картинка-контент та картинка-стиль. Результуюче зображення містить в собі об'єкти контенту та дрібні деталі картини-стилю. Зазвичай, контент – це фотографія, а стиль – абстрактний малюнок, в результаті одержується ніби намальована фотографія. Вхідні зображення пропускаються через згорткову нейронну мережу, після чого для вихідного зображення вираховуються значення функції втрати контенту та функції втрати стилю.

Функція втрати контенту визначається як:

$$L_{content}^l(X, Y_{content}) = \frac{1}{HWC} \sum_{c=1}^C \sum_{i=1}^H \sum_{j=1}^W (X_{ijc}^l - Y_{ijc}^l)^2, \quad (1)$$

де X^l, Y^l – ознаки, отримані після обробки зображень X та Y в l -му шарі нейронної мережі; H, W, C – розмір та кількість каналів відповідно.

Функція втрати контенту демонструє, наскільки дані зображення схожі між собою, причому порівняння здійснюється не безпосередньо (попиксельно), а шляхом визначення та подальшого порівняння високорівневих ознак зображень.

Функція втрати стилю:

$$L_{style}^l(X, Y_{style}) = \sum_i \sum_j \left(G_{ij}^{X^l} - G_{ij}^{Y_{style}^l} \right)^2, \quad (2)$$

де G^A – матриця Грама для збірки ознак A :

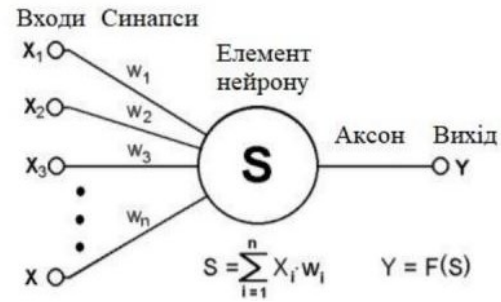


Рис.1. Структура штучного нейрону

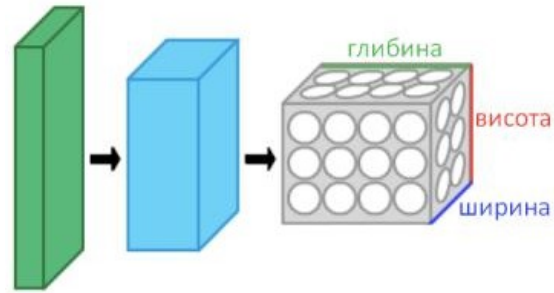


Рис.2. Модель згорткової нейронної мережі

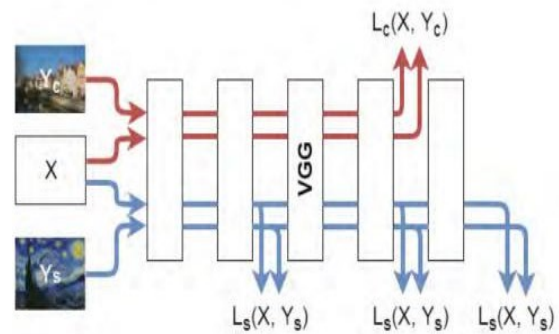


Рис.3. Принцип роботи функцій втрати контенту та стилю

$$G_{ij}^A = \frac{1}{HWC} \sum_{x=1}^W \sum_{y=1}^H a_{xy}^i a_{xy}^j, \quad (3)$$

a_{xy}^c – нейрон в координатах xu в каналі ознак A .

Матриця Грама ілюструє зв'язок між нейронами в різних каналах, а функція втрати стилю обмежує різницю в матрицях та демонструє, наскільки зображення схожі за стилем.

Далі вирішується задача оптимізації:

$$X = \arg \min_x \left[\alpha \sum_{l \in Q_{content}} L_{content}^l(X, Y_{content}) + \beta \sum_{l \in Q_{style}} L_{style}^l(X, Y_{style}) \right], \quad (4)$$

де $Q_{content}$ та Q_{style} – набори номерів тих шарів, в яких потрібно визначити функції втрат, α та β – вага контенту та стилю відповідно (чим більшим



є відношення $\frac{\alpha}{\beta}$, тим менше стилізованим буде фото, і навпаки).

IV. ЕКСПЕРИМЕНТ

Вхідні зображення масштабуємо до необхідного розміру. Завантажуємо попередньо підготовлену неймережу (VGG19). Знаючи, що ми можемо розрізняти шари, що відповідають за стиль (основні форми) і ті, що відповідають за контент (особливості зображення), ми можемо розділити шари для самостійної роботи над змістом і стилем.

Тоді ми ставимо наше завдання як завдання оптимізації, де ми будемо мінімізувати:

- втрату контенту (відстань між вхідними та вихідними зображеннями)
- втрату стилю (відстань між стилем і вихідними зображеннями)
- сумарні втрати варіації (просторове згладжування вихідного зображення)



Рис.4. Вхідне зображення контенту.



Рис.5. Вхідне зображення стилю



Рис. 6. Результат експерименту

Визначивши нашу модель, визначаємо функцію втрати контенту. Щоб зберегти оригінальний контент, зводимо до мінімуму різницю між вхідним зображенням і вихідним зображенням.

Подібно до втрати контенту, втрата стилю також визначається як різниця між двома зображеннями. Однак, для застосування нового стилю, втрату стилю визначаємо як різницю між зображенням стилю та вихідним зображенням. Визначаємо загальну втрату варіації.

Після цього, обчисливши втрату контенту, стилю та загальну втрату варіацій, визначаємо процес передачі стилю як проблему оптимізації, де ми зводимо звести до мінімуму наші глобальні втрати (це поєднання вмісту, стилю та загальних втрат варіацій). У кожній ітерації створюємо вихідне зображення так, щоб відстань (різниця) між виходом і входом / стилем на відповідних шарах об'єктів була мінімізована.

ВИСНОВКИ

В даній роботі була реалізована нейронна мережа з якісною глибиною зображення. При вивченні розпізнавання об'єкту згорткова мережа стає інваріантною до всього зображення і зберігає ідентичність об'єкта. Те, що нейронна мережа навчена передавати стиль - є великим поштовхом у можливостях виведення нашої візуальної системи.

ПЕРЕЛІК ПОСИЛАНЬ

- [1] X.-C. Liu, M.-M. Cheng, Y.-K. Lai, and P. L. Rosin, "Depth-aware neural style transfer," in *Proceedings of the Symposium on Non-Photorealistic Animation and Rendering*, 2017, pp. 1–10, DOI: [10.1145/3092919.3092924](https://doi.org/10.1145/3092919.3092924)
- [2] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image Style Transfer Using Convolutional Neural Networks," *2016 IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 2414–2423, 2016, DOI: [10.1109/CVPR.2016.265](https://doi.org/10.1109/CVPR.2016.265)



- [3] W. Chen, Z. Fu, D. Yang, and J. Deng, "Single-Image Depth Perception in the Wild," *Adv. Neural Inf. Process. Syst.*, pp. 730–738, 2016.
- [4] A. G. Howard, Some Improvements on Deep Convolutional Neural Network Based Image Classification. 2013
URL: <https://arxiv.org/abs/1312.5402>
- [5] P.-Y. Laffont, Z. Ren, X. Tao, C. Qian, and J. Hays, "Transient attributes for high-level understanding and editing of outdoor scenes," *ACM Trans. Graph.*, vol. 33, no. 4, pp. 1–11, 2014, DOI: [10.1145/2601097.2601101](https://doi.org/10.1145/2601097.2601101)
- [6] C. Li and M. Wand, "Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis," 2016, URL: <http://arxiv.org/abs/1601.04589>
- [7] V. V. Vyugin, *Matematicheskiye osnovy teorii mashinnogo obucheniya i prognozirovaniya [The mathematical foundations of machine learning theory and forecast]*. Moscow, 2013.
- [8] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," 2014, URL: <http://arxiv.org/abs/1409.1556>
- [9] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs.," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, 2018, DOI: [10.1109/TPAMI.2017.2699184](https://doi.org/10.1109/TPAMI.2017.2699184)
- [10] A. A. Lukjanitsa and A. G. Shishkin, *Obrabotka videoizobrazheniy [Video processing]*. Moscow: ISS Press, 2009.

Надійшла до редакції 10 квітня 2019 р.

УДК 004.942:519.876

Применение нейронной сети для стилизованной обработки изображений

Колесник А. Ю., ORCID [0000-0002-8896-1142](https://orcid.org/0000-0002-8896-1142)

e-mail: kolesnyk.oleksandra@gmail.com

Кафедра промышленной электроники www.kaf-pe.kpi.ua

Национальный технический университет Украины

«Киевский политехнический институт имени Игоря Сикорского» www.kpi.ua

Киев, Украина

Аннотация—В статье рассмотрена задача обработки графических изображений с использованием математического аппарата искусственных нейронных сетей. Отдельно рассматривается структура сверточной нейронной сети. Для решения задачи художественной обработки фотографии с использованием заданной стилистической техники используются два изображения – «картинка-контент» и «картинка-стиль», которая задает шаблон стилистики. При использовании различных стилистических шаблонов можно сформировать фото в разной стилистике – например, в стиле импрессионизма, абстракционизма и т.п. Таким образом, работа представляет интерес как с художественной, так и с исследовательской точки зрения.

Библ. 10, рис. 6.

Ключевые слова — обработка изображений, искусственные нейронные сети, сверточные сети, свертка, передача стиля



UDC 004.942:519.876

Application of a Neural Network for Stylized Image Processing

O. Yu. Kolesnyk, ORCID [0000-0002-8896-1142](https://orcid.org/0000-0002-8896-1142)

e-mail: kolesnyk.oleksandra@gmail.com

Department of Industrial Electronics www.kaf-pe.kpi.ua

National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» www.kpi.ua
Kyiv, Ukraine

Abstract—Manually re-drawing an image in a certain artistic style takes a professional artist a long time. Here you can use image representations derived from the Neural Networks. It is considered the problem of image processing using convolutional neural network. The article deals with the typical architecture of the convolutional neural network, the advantages of their use in problems associated with pattern recognition are presented. The problem of artist style transfer where the painting style of a given artist is applied on a real world photograph is addressed. approach uses the statistics of high-level feature representations of the images from the hidden layers of an image classification network to separate and reassemble content and style. To automatically transfer an artistic style, the first and most important issue is how to model and extract style from an image. Transferring the style from one image onto another can be considered a problem of texture transfer. In texture transfer the goal is to synthesise a texture from a source image while constraining the texture synthesis in order to preserve the semantic content of a target image. Since style is very related to texture, a straightforward way is to relate Visual Style Modelling back to previously well-studied. After obtaining the style representation, the next issue is how to reconstruct an image with desired style information while preserving its content, which is addressed by the image reconstruction technique. This is done by formulating an optimization problem that, starting with white noise, searches for a new image showing similar neural activations as the content image and similar feature correlations as the style image. A Neural Algorithm of Artistic Style is a pattern that can be separate and recombine the images is introduced. The algorithm allows us to produce new images of High quality perceptual artworks. The result provides new insights into the for high level representations image synthesis and manipulation. The result was generated on the basis of the VGG network, which was trained to perform object recognition and localisation. The usage of convolutional neural network is analyzed. Representations that factorise the variation in the content of an image and the variation in its appearance would be extremely practical for this task. In light of the striking similarities between performance optimised artificial neural networks and biological vision, the human ability to abstract content from style – and therefore our ability to create and enjoy art – might also be primarily a preeminent signature of the powerful inference capabilities of our visual system.

Ref. 10, fig. 6.

Keywords – image processing; artificial neural networks; convolutional networks; convolution; style transfer.

