

ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИРІШЕННЯ ЗАДАЧІ РОЗФАРБУВАННЯ ЧОРНО-БІЛИХ ЗОБРАЖЕНЬ

Вінницький національний технічний університет

Анотація

Розглянуто підходи до розфарбування чорно-білих зображень за допомогою нейронних мереж з виведенням сітки з пікселів, що дозволило структурувати механізм розфарбування та запропонувати мережеву архітектуру нейронної мережі для розфарбування чорно-білих зображень.

Ключові слова: зображення, нейронні мережі, колоризація, розфарбування, діапазон, піксель, кодування, контрастність.

Abstract

The principles of colorization of black-and-white images by means of neural networks with output of a grid from pixels that allowed to structure the mechanism of colorization and to offer network architecture of a neural network for colorization of black-and-white images are considered.

Keywords: images, neural networks, coloration, colorization, range, pixel, encoding, contrast.

Вступ

Автоматизоване розфарбування чорно-білих зображень на протязі багатьох років була предметом досліджень в області комп'ютерного зору і машинного навчання. Крім своєї захопливості з точки зору естетики і штучного інтелекту, така можливість має широке практичне застосування, починаючи від відновлення відео і закінчуючи поліпшенням зображення для максимальної інтерпретації.

В останні роки нейронні мережі стали де-факто стандартом для вирішення проблем класифікації зображень, досягаючи частоти помилок нижче 4% в завданні колоризації чорно-білих зображень.

Результати дослідження

Чорно-білі зображення можна представити у вигляді сітки з пікселів (рис. 1). У кожного пікселя є значення яскравості, що лежить в діапазоні від 0 до 255, від чорного до білого.

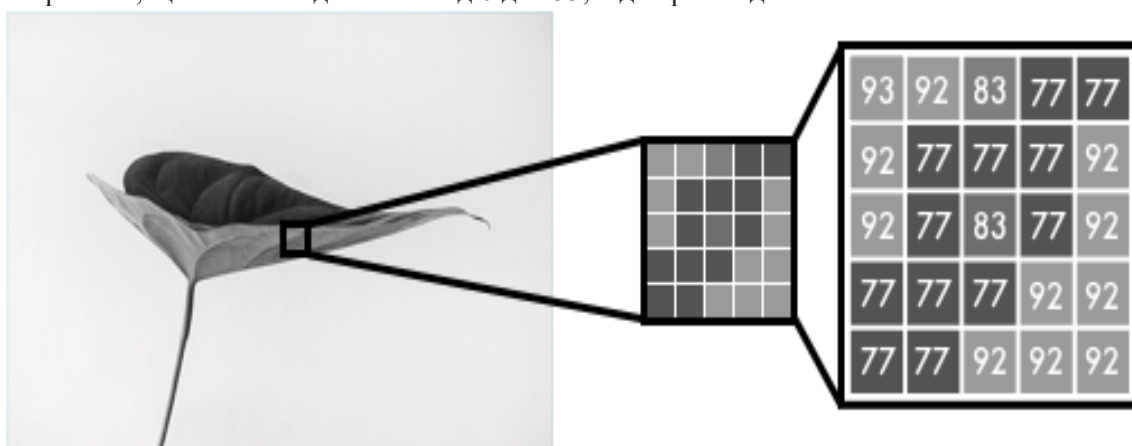


Рис. 1. Структура формування чорно-білого зображення у вигляді сітки з пікселів [1]

Кольорові зображення складаються з трьох шарів: червоного, зеленого і синього. Тобто в кольоровому зображенні за допомогою трьох шарів кодується колір і контрастність (рис. 2).

R	B	G	=	pixel
30	30	255	=	
0	0	255	=	
0	0	210	=	

Рис. 2. Структура кольорового зображення

Як і в чорно-білому зображенні, пікселі кожного шару кольорового зображення містять значення від 0 до 255. Нуль означає, що у цього пікселя в даному шарі немає кольору. Якщо у всіх трьох каналах стоять нулі, то в результаті на зображенні виходить чорний піксель.

Нейромережа встановлює взаємозв'язок між вхідним і вихідним значеннями. У нашому випадку нейромережа повинна знайти сполучні риси між чорно-білими і кольоровими зображеннями. Тобто знайти властивості, за якими можна зіставити значення з чорно-білої сітки зі значеннями з трьох кольорів.

$$f \left(\begin{bmatrix} 93 & 92 & 83 & 77 & 77 \\ 92 & 77 & 77 & 77 & 92 \\ 92 & 77 & 83 & 77 & 92 \\ 77 & 77 & 77 & 92 & 92 \\ 77 & 77 & 92 & 92 & 92 \end{bmatrix} \right) = \begin{bmatrix} 83 & 92 & 83 & 77 & 77 \\ 99 & 99 & 77 & 77 & 92 \\ 99 & 77 & 83 & 77 & 92 \\ 77 & 77 & 77 & 95 & 92 \\ 77 & 77 & 95 & 92 & 92 \end{bmatrix} \begin{bmatrix} 93 & 92 & 83 & 69 & 69 \\ 92 & 69 & 69 & 77 & 92 \\ 92 & 69 & 83 & 77 & 92 \\ 69 & 69 & 77 & 92 & 92 \\ 77 & 77 & 92 & 92 & 92 \end{bmatrix} \begin{bmatrix} 83 & 92 & 83 & 77 & 77 \\ 83 & 77 & 77 & 77 & 92 \\ 92 & 77 & 83 & 75 & 85 \\ 75 & 77 & 75 & 85 & 85 \\ 75 & 75 & 85 & 85 & 85 \end{bmatrix}$$

$f()$ – нейромережа, $[B \& W]$ – вхідні дані, $[R], [G], [B]$ – вихідні дані

Рис. 3. Формування сполучних рис між зображеннями

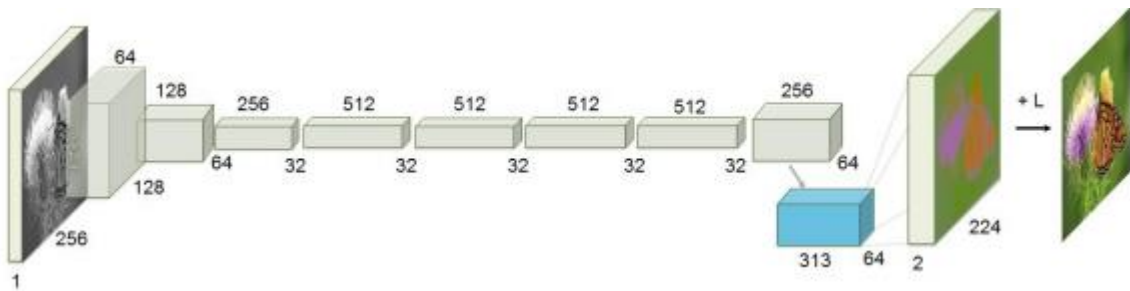


Рис. 4. Мережева архітектура

Кожен згортковий шар відноситься до блоку з 2 або 3 повторюваних шарів, за якими слідує шар нормалізації. Всі зміни в дозволі досягаються за рахунок просторової понижувальної або підвищуючої дискретизації між блоками.

Для вхідного каналу яскравості $X \in \mathbb{R}^{H \times W \times 1}$, необхідно вивчити відображення $\hat{Y} = F(X)$ в два асоційованих колірних канали $X \in \mathbb{R}^{H \times W \times 2}$, де H, W - розміри зображення.

Відстані сприйняття в цій моделі і природна цільова функція, являють собою евклідову втрату $L_2(\bullet, \bullet)$ між прогнозованим і основним істинним кольором [2]:

$$L_2(\hat{Y}, Y) = \frac{1}{2} \sum_{h,w} \|Y_{h,w} - \hat{Y}_{h,w}\|_2^2 \quad (1)$$

Однак ця втрата не стійка до властивості її неоднозначності і мультимодального забарвлення. Якщо об'єкт може приймати набір різних значень ab , оптимальним рішенням евклідової втрати буде середнє значення набору. При прогнозуванні кольору цей ефект усереднення сприяє отриманню сіруватих, ненасичених результатів. Крім того, якщо набір забарвлень не опуклий, рішення буде насправді поза набору, даючи неправдоподібні результати.

Замість цього розглядаємо проблему як поліноміальну класифікацію. Квантуємо ab вихідний простір в осередку з розміром сітки 10 і зберігаємо $Q = 313$ значень, які знаходяться в межах гамми, як показано на рисунку 5 (а). Для цього входу X вивчаємо відображення $\hat{Z} = G(X)$ до розподілу ймовірностей щодо можливих кольорів $\hat{Z} \in [0,1]^{H \times W \times Q}$, де Q – кількість квантових значень ab .

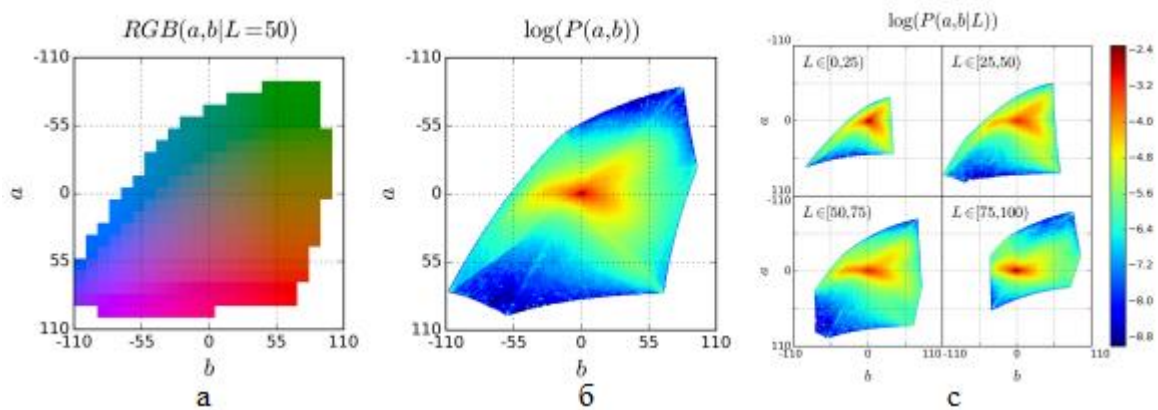


Рис. 5. а) Квантований колірний простір ab розміром сітки 10. Загалом є 313 ab пар в гамі. б) Емпіричний розподіл ймовірності значень ab , показаний у шкалі журналу. с) Емпіричний розподіл ймовірностей значень ab , обумовлений L , показаний у шкалі журналу

Щоб порівняти $\hat{Z}$ з вихідною істинністю, визначаємо функцію $Z = H_{gt}^{-1}(Y)$, яка перетворює основний колір Y у вектор Z , використовуючи схему м'якого кодування.

Потім використовуємо поліноміальну перехресну втрату ентропії, яка визначається як:

$$L_{cl}(\hat{Z}, Z) = -\sum_{h,w} v(Z_{h,w}) \sum_q Z_{h,w,q} \log(\hat{Z}_{h,w,q}) \quad (2)$$

де $v(\bullet)$ – це ваговий член, який можна використовувати для перебалансування втрат на основі кольорового класу. Нарешті, відображаємо ймовірність розподілу $\hat{Z}$ до значень кольору $\hat{Y}$ з функцією $\hat{Y} = H(\hat{Z})$.

Висновки

Попри те, що кольоровість зображень є комп'ютерною графічною задачею, вона також є прикладом складної проблеми передбачення пікселів у комп'ютерному зорі. В ході дослідження виявлено, що розфарбування, яке базується на нейронних мережах з глибоким навчанням і добре підбраною цільовою функцією може наблизитися до отримання результатів, які неможливо відрізнити від справжніх кольорових фотографій.

Запропонований метод не тільки забезпечує корисний графічний результат, але й може бути представлений як провідне завдання для репрезентативного навчання.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. R. Dahl. Automatic colorization. <http://tinyclouds.org/colorize/>, 2016.
2. Бас С. М. Бінаризація зображень на основі нейронних мереж кохонена / Станіслав Миколайович Бас, Олександр Іванович Сороковий // Новітні інформаційні системи та технології. – Полтава: ПНТУ, 2014. – Т. (1). – Режим доступу: <http://journals.nupr.edu.ua/mist/article/view/459> (дата звернення: 02.02.2021).

Павлович Роман Ігорович — аспірант кафедри комп'ютерних наук, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Україна, м. Вінниця, e-mail: pavlovich.roma97@gmail.com.

Перевозніков Сергій Іванович — д.т.н., професор кафедри комп'ютерних наук, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Україна, м. Вінниця, e-mail: perevoznikov@ukr.net

Roman I. Pavlovych — Department of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: pavlovich.roma97@gmail.com.

Serhiy I. Perevoznikov — Doct. Sc. (Eng.), Professor of the Computer Science Chair, Department of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: perevoznikov@ukr.net