

ПРОГРАМНИЙ МОДУЛЬ РОЗФАРБОВУВАННЯ ЗОБРАЖЕНЬ НА ОСНОВІ ЗГОРТКОВОЇ НЕЙРОМЕРЕЖІ

Вінницький національний технічний університет

Анотація

Дана робота присвячена розробці програмного модуля розфарбовування зображень на основі згорткової нейромережі. Було розроблено постановку задачі розфарбовування чорно-білих зображень. Розроблено архітектуру згорткової нейронної мережі, яка лежить в основі розробленого методу. Архітектура нейромережі побудована по схемі енкодер-декодер та включає гілку вилучення ознак на основі попередньо навченої нейронної мережі Inception-ResNet-v2. Описано основні етапи програмної реалізації та функціонування модуля розфарбовування зображень на основі згорткової нейронної мережі, що містять завантаження набору даних, реалізації архітектури згорткової нейронної мережі, навчання згорткової нейронної мережі, оцінювання процесу навчання. Запропонована архітектура є простішою за аналоги за рахунок використання у гілці вилучення ознак попередньо навченої нейронної мережі Inception-ResNet-v2.

Ключові слова: розфарбовування, чорно-біле зображення, згорткова нейронна мережа

Abstract

This work is devoted to the development of a software module for coloring images based on a convolutional neural network. The problem statement for coloring black-and-white images was developed. The architecture of a convolutional neural network, which underlies the developed method, was developed. The architecture of the neural network is built according to the encoder-decoder scheme and includes a feature extraction branch based on a pre-trained neural network Inception-ResNet-v2. The main stages of the software implementation and functioning of the image coloring module based on a convolutional neural network are described, which include loading the data set, implementing the architecture of the convolutional neural network, training the convolutional neural network, and evaluating the learning process. The proposed architecture is simpler than analogues due to the use of a pre-trained neural network Inception-ResNet-v2 in the feature extraction branch.

Keywords: coloring, black and white image, convolutional neural network

Вступ

Розфарбовування зображень у градаціях сірого може мати великий вплив у широкому спектрі галузей, наприклад, на ремастеринг історичних зображень та покращення потоків спостереження. Інформаційний вміст зображення у градаціях сірого досить обмежений, тому додавання кольорових компонентів може дати більше інформації про його семантику. У контексті глибокого навчання такі моделі, як Inception [1], ResNet [2] або VGG [3], зазвичай навчаються з використанням кольорових наборів даних зображень. Під час застосування цих мереж до зображень у градаціях сірого попередній крок розфарбовування може допомогти покращити результати. Однак, розробка та впровадження ефективної та надійної системи, яка автоматизує цей процес, досі залишається складним завданням.

Метою роботи є спрощення процесу розфарбовування зображень за рахунок використання згорткової нейронної мережі.

Постановка задачі

Визначимо проблему розфарбовування з точки зору колірного простору CIE Lab. Як і колірний простір RGB, це 3-канальний колірний простір, але на відміну від RGB, інформація про колір кодується лише в каналах a (зелено-червоний компонент) та b (синьо-жовтий компонент). Канал L (яскравість) кодує лише інформацію про інтенсивність.

Зображення у градаціях сірого, яке ми хочемо розфарбувати, можна розглядати як L-канал зображення в колірному просторі Lab, а наша мета — знайти компоненти а та b. Отримане таким чином зображення Lab можна перетворити в колірний простір RGB за допомогою стандартних перетворень колірного простору. Наприклад, в OpenCV цього можна досягти за допомогою `cvtColor` з опцією `COLOR_BGR2Lab`.

Для спрощення обчислень, простір ab колірного простору Lab квантується на 313 інтервалів, як показано на рисунку 1.

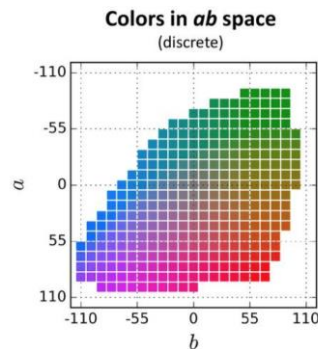


Рисунок 1 - Вид 313 квантованих кольорів в ab-просторі

Замість того, щоб знаходити значення а та b для кожного пікселя, через це квантування нам просто потрібно знайти номер інтервалу від 0 до 312. Ще один спосіб мислення про проблему полягає в тому, що у нас вже є канал L, який приймає значення від 0 до 255, і нам потрібно знайти канал ab, який приймає значення від 0 до 312. Таким чином, задача прогнозування кольору тепер перетворюється на задачу поліноміальної класифікації, де для кожного сірого пікселя є 313 класів на вибір.

Архітектура нейронної мережі

Архітектура використовуваної ЗНМ базується на роботі [4]: за заданою компонентою яскравості зображення модель оцінює його компоненти $a*b*$ і поєднує їх із вхідними даними для отримання остаточної оцінки кольорового зображення. Замість того, щоб навчати гілку вилучення ознак з нуля, було використано нейромережу Inception-ResNet-v2 (надалі називану Inception) і отримуємо вбудоване представлення сірого зображення з її останнього шару. Використовувана архітектура нейромережі показана на рис. 2.

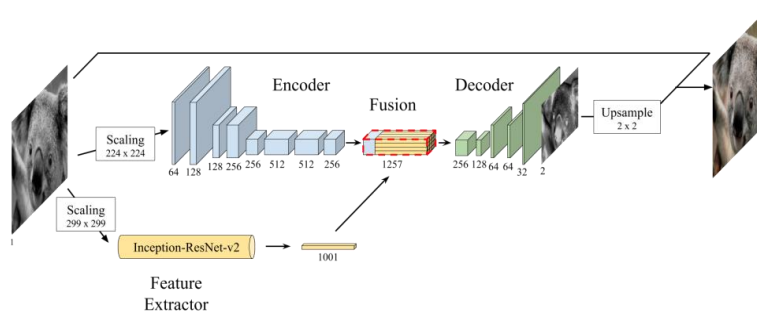


Рисунок 2.9 – Архітектура використовуваної згорткової нейромережі

Мережа логічно поділена на чотири основні компоненти. Компоненти кодування та вилучення ознак отримують ознаки середнього та високого рівня відповідно, які потім об'єднуються у шарі злиття. Нарешті, декодер використовує ці ознаки для оцінки вихідних даних. Таблиця 2.1 детальніше описує шари мережі, в якій ліворуч показана мережа кодера, посередині - мережа злиття, праворуч - мережа декодера. Кожен згортковий шар використовує функцію активації ReLU, за винятком останнього, який застосовує функцію гіперболічного тангенсу. Гілка вилучення ознак має таку ж архітектуру, як Inception-ResNet-v2, за винятком останнього шару softmax.

Вимоги до програмного модуля розфарбовування зображень на основі згорткової нейронної мережі

Розроблювана програма повинна перетворювати чорно-біле зображення у кольорове (розфарбовувати чорно-біле зображення). Зокрема, вона має робити наступне:

1. Завантажує попередньо навчану нейронну мережу Inception-ResNet-v2, яка буде використовуватись для вилучення ознак вхідного зображення.
2. Завантажує датасет зображень, сформований на основі ImageNet.
3. Змінює розміри зображень до 224x224 для подачі в енкодер і до 299x299 для подачі в нейронну мережу Inception-ResNet-v2.
4. Виконує попередню обробку зображень завантаженого датасету, яка полягає у їх нормалізації до діапазону $[-1,1]$ та отриманні ембеддінгів.
5. Виконує навчання згорткової нейромережі задачі розфарбовування зображень на основі навчальної вибірки протягом заданої кількості епох.
6. Оцінює продуктивність навченої нейромережі шляхом підрахунку середньоквадратичної помилки розфарбовування на зображеннях тестової вибірки.
7. При потребі розфарбовує невідоме мережі зображення користувача, яке записано у відповідну папку

Висновки

Таким чином, було розроблено постановку задачі розфарбовування чорно-білих зображень. Розроблюваний метод відноситься до автоматичних методів мультимодального розфарбовування зображень у градаціях сірого. Було розроблено архітектуру згорткової нейронної мережі, яка лежить в основі розробленого методу. Архітектура нейромережі побудована по схемі енкодер-декодер та включає гілку вилучення ознак на основі попередньо навченої нейронної мережі Inception-ResNet-v2. Описано основні етапи програмної реалізації та функціонування модуля розфарбовування зображень на основі згорткової нейронної мережі, що складається з завантаження набору даних, реалізації архітектури згорткової нейронної мережі, навчання згорткової нейронної мережі, оцінювання процесу навчання. Запропонована архітектура є простішою за аналоги за рахунок використання у гілці вилучення ознак попередньо навченої нейронної мережі Inception-ResNet-v2.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. CoRR abs/1512.00567 (2015)
2. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. CoRR abs/1512.03385 (2015)
3. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. CoRR abs/1409.1556 (2014)
4. Iizuka, S., Simo-Serra, E., Ishikawa, H.: Let there be color!: joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification. ACM Transactions on Graphics (TOG) 35(4) (2016) 110

Гречкосій Марк Петрович — студент групи 1КН-216, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, email: markgrks@gmail.com

Короленко Олег Олександрович — аспірант кафедри комп'ютерних наук, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, email: olegkorolenko.legoex@gmail.com

Колесницький Олег Костянтинович — канд. техн. наук, професор кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця. email: kolesnytskiy@vntu.edu.ua

Kovalyk Andriy V.— student of Intelligent Information Technologies and Automation Department, Vinnytsia National Technical University, Vinnytsia, e-mail: : markgrks@gmail.com

Korolenko Oleh O. - Postgraduate student of the Department of Computer Science, Vinnytsia National Technical University, Vinnytsia. email: olegkorolenko.legoex@gmail.com

Kolesnytskyj Oleh K. – Cand. Sc. (Eng.), Professor of the Computer Science Chair, Vinnytsia National Technical University, Vinnytsia. email: kolesnytskiy@vntu.edu.ua