

Міністерство освіти і науки України
НТУУ «Київський політехнічний інститут»
Фізико-технічний інститут

Машинне навчання
Лабораторна робота №6
Отримання навичок використання існуючого програмного коду

Виконав:
Студент 4-го курсу
групи ФІ-21
Климент'єв Максим
Перевірив:

Зміст

1	Мета роботи	3
2	Огляд обраної статті	4
2.1	Опис проблеми	4
2.2	Запропонований алгоритм та новизна	4
3	Демонстрація роботи	5
4	Шляхи покращення	6
5	Висновок	7
	Перелік посилань	8

1 Мета роботи

Отримати навички використання існуючого програмного коду.

2 Огляд обраної статті

Була обрана стаття, яку я аналізував в своєму аналізі літератури для дипломної роботи *Colorful Image Colorization* [4].

Назва: Colorful Image Colorization 2016

Автори: Richard Zhang, Phillip Isola, Alexei A. Efros.

Посилання на статтю: <https://arxiv.org/abs/1603.08511>

Посилання на код: <https://github.com/richzhang/colorization>

2.1 Опис проблеми

Традиційні методи колоризації часто використовували функцію втрат L2. Оскільки для одного сірого об'єкта може існувати багато варіантів кольору (наприклад, яблуко може бути червоним, зеленим або жовтим), модель, що навчається на L2, намагається усереднити ці варіанти. Це призводить до ненасичених, сіруватих результатів.

2.2 Запропонований алгоритм та новизна

Автори пропонують підхід, що розглядає колоризацію не як задачу регресії, а як задачу **класифікації**.

- **Квантування кольору:** Простір кольорів Lab квантується на 313 дискретних класів.
- **Class Rebalancing:** Оскільки кольори в природі розподілені нерівномірно (багато сірого/неба, мало яскравого), автори вводять вагові коефіцієнти для рідкісних класів. Це змушує мережу генерувати більш яскраві та різноманітні кольори.
- **Архітектура:** Використовується Convolutional Neural Network типу VGG, яка передбачає розподіл ймовірностей кольорів для кожного пікселя.

3 Демонстрація роботи

Код було запущено у хмарному середовищі Google Colab з використанням офіційного репозиторію авторів. Модель навчили вже в самому колабі, оскільки завантажити з посилання вже не можна було.

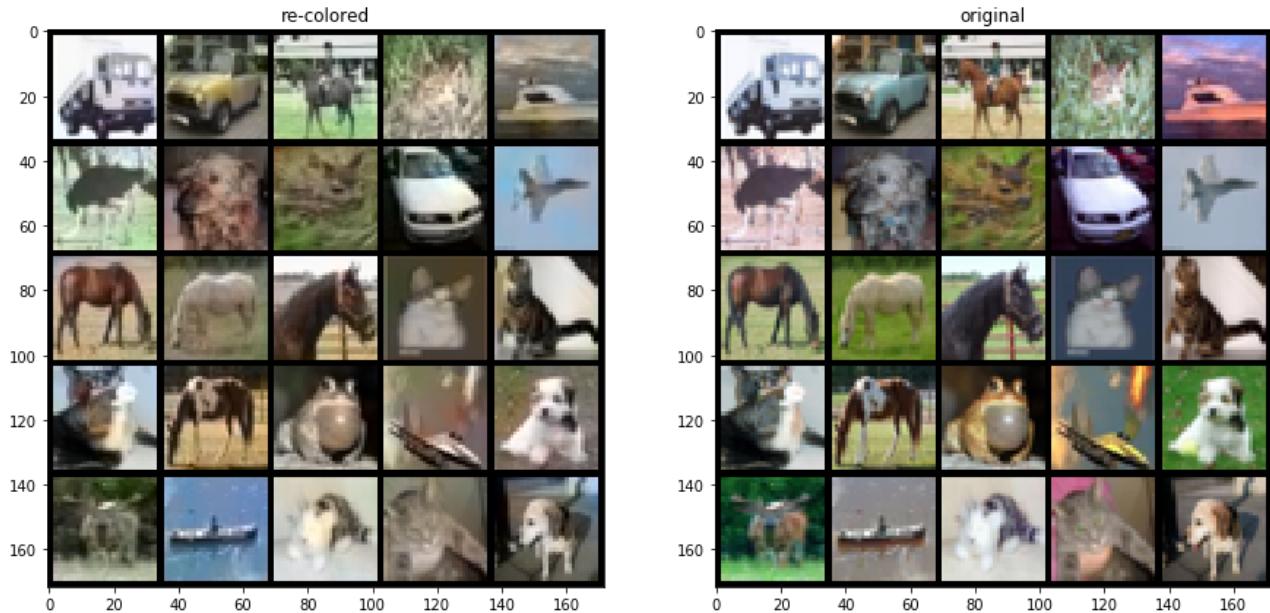


Рис. 1: Результат роботи алгоритму: зліва - результат автоматичної колоризації, справа - оригінал

На рис. 1 видно, що модель успішно відновила кольори, правильно ідентифікувавши небо, траву та об'єкт на передньому плані.

4 Шляхи покращення

Незважаючи на гарні результати, метод іноді допускає помилки. Базуючись на сучасних дослідженнях, пропонуються такі шляхи покращення:

1. **Використання GAN:** Застосування підходу *Pix2Pix*[1] або *CycleGAN*[5] дозволило б покращити текстурну відповідність та різкість зображення завдяки дискримінатору.
2. **Diffusion Models:** Сучасні методи, такі як *Palette*[3] або *Control Color*[2], використовують дифузійні моделі. Вони дозволяють генерувати більш фотorealістичні зображення та краще розуміють глобальний контекст сцени.
3. **Multimodal Colorization:** Інтеграція текстових описів дозволила б користувачеві керувати процесом, наприклад, вказати системі: "зроби машину червоною".

5 Висновок

У ході лабораторної роботи було досліджено метод *Colorful Image Colorization*. Встановлено, що перехід від регресії до класифікації кольорів значно покращує візуальне сприйняття результату.

Перелік посилань

- [1] Phillip Isola та ін. «Image-To-Image Translation With Conditional Adversarial Networks». В: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.
- [2] Z. Liang та ін. *Control Color: Multimodal Diffusion-based Interactive Image Colorization*. Z. Liang et al. International Journal of Computer Vision, 1-27. 2025. URL:
<https://link.springer.com/article/10.1007/s11263-025-02549-6>.
- [3] Chitwan Saharia та ін. «Palette: Image-to-Image Diffusion Models». В: *ACM SIGGRAPH 2022 Conference Proceedings*. SIGGRAPH '22. Vancouver, BC, Canada: Association for Computing Machinery, 2022. ISBN: 9781450393379. DOI: 10.1145/3528233 / 3530757. URL:
<https://doi.org/10.1145/3528233.3530757>.
- [4] R. Zhang, P. Isola та A. A. Efros. *Colorful image colorization*. In European conference on computer vision (pp. 649-666). Cham: Springer International Publishing. <https://arxiv.org/abs/1603.08511>. 2016. URL:
https://link.springer.com/chapter/10.1007/978-3-319-46487-9_40.
- [5] Jun-Yan Zhu та ін. «Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks». В: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2017.