

Введение. Автоматическое улучшение изображений становится всё более важным. Применение подобных систем достаточно обширно: от общей цветокоррекции любительского снимка до использования таких подходов в качестве препроцессинга для дальнейшего решения задач обработки и анализа исходного изображения. Несмотря на достижения в этой области, многие методы сталкиваются с ограничениями, такими как высокая вычислительная сложность, неэффективность в реальном времени, требования к высоким вычислительным мощностям [4]. В научной среде, как правило, рассматривают тяжеловесные решения [1,2,3,6], способные породить цветовые артефакты. Одним из способов избежать появления артефактов на изображении является отказ от использования выходов нейронной сети на пиксельном уровне. Некоторые исследователи [1,2,4] для решения проблемы также предлагают использовать либо глубокое обучение с подкреплением, либо подбирать оптимальную последовательность заранее заданных цветовых преобразований. Обеспечение легковесности подхода позволяет модели оставаться портируемой на мобильные устройства, что также является несомненным преимуществом по сравнению с использованием дополнительных серверов и облачных решений в контексте приватности данных.

Основная часть.

Модель нейронной сети должна удовлетворять следующим требованиям:

- 1) высокое качество результатов в сравнении с другими решениями;
- 2) отсутствие артефактов;
- 3) легковесность;
- 4) высокая скорость работы;
- 5) портируемость для применения на мобильном устройстве.

Для обучения будем использовать общедоступный набор MIT Adobe 5K [8], состоящий из пар фото до и после обработки неким экспертом. Набор состоит из 5000 пар изображений. Мы возьмем обработку экспертом С в качестве целевой как наиболее часто используемую в других работах.

После проведения анализа литературы выбрано следующее оптимальное решение: двухэтапный подход цветовой коррекции цветовой гаммы изображений. На первом этапе идет предсказание параметров для фильтров, которые будут в дальнейшем применены к входному изображению. Уменьшенная копия исходного изображения подается в легковесный сверточный кодировщик, состоящий из трех блоков свертки с активацией LeakyReLU и нормализацией BatchNorm2d и двух полносвязных слоев. Выход кодировщика дополняется одной из функций активации: сигмоидальная, если для фильтра требуется распределение в диапазоне $[0,1]$, гиперболический тангенс, если требуется диапазон $[-1,1]$, либо же функция активации не применяется вовсе. Второй этап – применение дифференцируемых фильтров (white balance, exposure, ...) к исходному изображению с использованием полученных на предыдущем этапе параметров. Параллельно применяется несколько фильтров с различными параметрами, после чего их значения складываются с отсечением в диапазон $[0,1]$.

Кроме того, нами предложено новое преобразование (фильтр) на основе свертки с универсальными ядрами. Данное преобразование позволяет использовать гораздо меньше параметров для обучения весов в сверточном слое. Мы показали эффективность использования данного преобразования в задаче коррекции цветовой гаммы изображений.

Были выбраны следующие функции потерь: Расстояние L1 между сгенерированным и размеченным изображением, а также функция потерь, основанная на индексе структурной схожести изображений. Метрики качества: SSIM и PSNR [5,7]. Проведен сравнительный

анализ предложенного подхода с конкурирующими методами, наш подход является самым легковесным, содержит всего 47 тысяч обучаемых параметров и работает на уровне современных решений. Стоит отметить, что скорость работы нашего решения практически в 3 раза выше предыдущего state-of-the-art подхода в группе легковесных моделей цветовой коррекции.

Выводы. В проведенной исследовательской работе была разработана легковесная нейросетевая архитектура для автоматической коррекции цветовой гаммы изображений. Предложенное решение является state-of-the-art среди легковесных нейросетевых подходов для решения задачи цветовой коррекции по метрикам SSIM и PSNR [5] и содержит всего 47 тысяч обучаемых параметров. Кроме того, в работе проведено комплексное исследование применимости и производительности фильтров на основе свертки с универсальным ядром в задаче цветовой коррекции изображения.

Список использованных источников:

1. Deep photo enhancer: Unpaired learning for image enhancement from photographs with gans / Y.-S. Chen [и др.] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. — 2018. — С. 6306–6314.
2. Distort-and-recover: Color enhancement using deep reinforcement learning / J. Park [и др.] // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2018. — С. 5928–5936.
3. Exposure: A white-box photo post-processing framework / Y. Hu [и др.] // ACM Transactions on Graphics (TOG). — 2018. — Т. 37, № 2. — С. 1–17.
4. Tatanov O., Samarin A. LFIEM: Lightweight Filter-based Image Enhancement Model // 2020 25th International Conference on Pattern Recognition (ICPR). — 2021. — С. 873–878.
5. Hore A., Ziou D. Image quality metrics: PSNR vs. SSIM // 2010 20th International Conference on Pattern Recognition. — IEEE. 2010. — С. 2366–2369.
6. Deng Y., Loy C. C., Tang X. Aesthetic-driven image enhancement by adversarial learning // Proceedings of the 26th ACM international conference on Multimedia. — 2018. — С. 870–878.
7. Image quality assessment: from error visibility to structural similarity / Z. Wang [и др.] // IEEE transactions on image processing. — 2004. — Т. 13, № 4. — С. 600–612.
8. Learning Photographic Global Tonal Adjustment with a Database of Input / Output Image Pairs / V. Bychkovsky [и др.] // The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition. — 2011.