

**РАЗРАБОТКА НЕЙРОСЕТЕВОГО МЕТОДА
ПОВЫШЕНИЯ РАЗРЕШЕНИЯ, РАБОТАЮЩЕГО В
РЕАЛЬНОМ ВРЕМЕНИ**

Богатырев Евгений Николаевич

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: evgeneey.bogatyrev@graphics.cs.msu.ru

Научный руководитель — Ватолин Дмитрий Сергеевич

Задача повышения разрешения (Super-Resolution, SR) направлена на восстановление кадров высокого разрешения по их версиям низкого разрешения и играет важную роль в системах потоковой передачи видео. Видео контент, передающийся по сети, подвергается сильному сжатию, что приводит к потере деталей и появлению искажений. Несмотря на прогресс нейросетевых методов SR, большинство современных моделей либо обладают высокой вычислительной сложностью, либо демонстрируют низкое качество при обработке сжатого видео в условиях реального времени [1].

В данной работе предлагается метод повышения разрешения EfRFLN (Efficient Residual Local Feature Network), нацеленный на работу в реальном времени. Метод комбинирует в себе идеи из других эффективных архитектур и направлен на улучшение реконструкции текстур и границ без увеличения вычислительных затрат.

Общая схема предложенного подхода представлена на рис. 1. Архитектура сети включает блок извлечения признаков, последовательность модулей обработки и блок реконструкции изображения. Основу модели составляют разработанные блоки эффективных остаточных локальных признаков (ERLFB).

Ключевым изменением является использование гиперболического тангенса в качестве функции активации вместо ReLU, что позволяет сохранять знак признаков и улучшает распространение градиентов в сети. Дополнительно в архитектуру интегрирован механизм Efficient Channel Attention [2], выполняющий адаптивное взвешивание каналов признаков при минимальных вычислительных затратах.

Для обучения сети предложена составная функция потерь, объединяющая функционал качества L1, потерю на основе признаков сети VGG-19 [3] и градиентную потерю Собеля. Такая комбинация обеспечивает одновременное повышение точности реконструкции границ и улучшение визуального качества результата. Для обучения модели был собран набор данных их 5200 пар видео разных

разрешений, полученных с видеохостинга YouTube.

Экспериментальная оценка показывает, что предложенная модель превосходит существующие методы повышения разрешения реального времени в объективном сравнении с использованием функционалов качества PSNR и LPIPS [4], а также по результатам субъективного сравнения, в котором приняло участие 3822 зрителя.

Иллюстрации

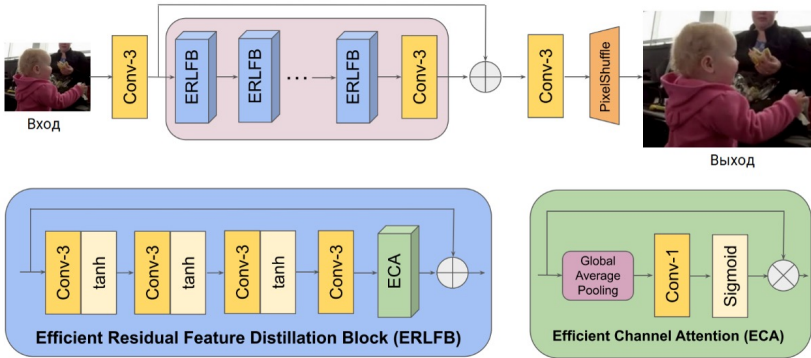


Рис 1. Общая схема предложенного метода.

Литература

1. Kong F., Li M., Liu S., Liu D., He J., Bai Y., Fu L. Residual Local Feature Network for Efficient Super-Resolution // In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, P. 766–776
2. Wang Q., Wu B., Zhu P., Li P., Zuo W., Hu Q. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks // In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, P. 11534–11542.
3. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition // In arXiv preprint arXiv:1409.1556, 2014
4. Zhang R., Isola P., Efros A., Shechtman E., Wang O. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. // In Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, P. 586–595.