

**РАЗРАБОТКА НЕЙРОСЕТЕВОГО МЕТОДА
ОБНАРУЖЕНИЯ ОТМАСШТАБИРОВАННЫХ
ИЗОБРАЖЕНИЙ**

Мещанинов Вячеслав Павлович

Студент

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

E-mail: s02180080@gse.cs.msu.ru

Научный руководитель — Ватолин Дмитрий Сергеевич

Необходимость решения задачи обнаружения отмасштабированного видео и изображений впервые возникла в сфере проверки подлинности медиафайлов. Изначально для повышения разрешения использовались методы линейной и бикубической интерполяции, которые оставляли заметные артефакты на кадрах. Но после появления нейросетевых алгоритмов повышения разрешения, качество работы которых значительно выше, чем у классических, появилась необходимость в выявлении зависимостей между пикселями, которые человеческий глаз обнаружить неспособен.

В данной работе рассматривается два подхода к решению задачи — классификация изображений на основе дискретного преобразования Фурье и нейросетевой алгоритм, работающий с градиентами входных изображений.

Первый подход мотивирован тем, что спектры частот оригинальных и отмасштабированных изображений имеют отличительные особенности рис. 1 [1]. Для классификации спектров использовались деревья решений. Но экспериментально было установлено, что на спектр частот также влияют и другие факторы, такие как камера, на которую производилась съемка, и разрешение изображения. Поэтому алгоритмы классификации дают низкую точность при данном подходе.

Второй подход заключается в использовании усовершенствованной нейросетевой архитектуры [2]. Для извлечения пространственных признаков кадр проходит через слой извлечения шума. Далее каждый поток проходит через четыре слоя извлечения нейросетевых признаков, которые включают в себя операции свертки, батч нормализации, активации и пулинга. Затем, чтобы зафиксировать корреляцию между обоими направлениями: горизонтальным и вертикальным, используется третий поток, входом которого является конкатенация выходов первых слоев горизонтального и вертикального потоков. Далее конкатенация признаков с обоих потоков пода-

ется на вход классификатору.

Для обучения методов был взят датасет изображений разрешения 512x512 пикселей Alaska и отмасштабирован с помощью нейросетевых алгоритмов повышения разрешения и классических методов интерполяции, таких как бикубическая и билинейная. В итоге размер обучающей выборки составил 420000 изображений.

Для тестирования был создан набор данных из 10-кадровых плоскопараллельных сцен высокого разрешения, используя видео платформу Vimeo, и отмасштабирован пятью методами нейросетевого повышения видео и изображений. Количество кадров выбрано в соответствии с рекомендациями алгоритмов повышения разрешения. Результаты тестирования показаны на рис. 2.

Иллюстрации

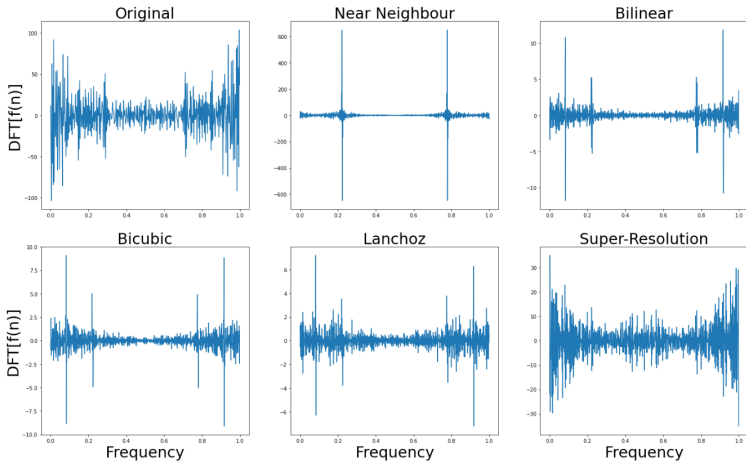


Рис. 1. Графики дискретного преобразования Фурье для оригинального изображения и различных методов повышения разрешения.

Метод нейросетевого повышения разрешения	Классический алгоритм [1]	DSCNN [2]	Предложенный алгоритм
D3D-NET	0.32	0.87	1
RBPN	0.27	0.83	0.97
Real-SR	0.23	0.049	0.65
SOF-VSR	0.24	0.81	0.99
TDAN	0.34	0.9	1

Рис. 2. Результаты тестирования алгоритмов на 5 методах нейросетевого повышения разрешения. В качестве метрики использовалась AUC ROC.

Литература

1. Gallagher, Andrew C Detection of linear and cubic interpolation in JPEG compressed images // The 2nd Canadian Conference on Computer and Robot Vision, 2005.
2. Cao G, et al Resampling detection of recompressed images via dual-stream convolutional neural network //arXiv preprint, 2019.</div>

Благодарности. Исследования выполнены при финансовой поддержке РФФИ в рамках научного проекта 19-01-00785 а.