

ВОССТАНОВЛЕНИЕ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ НЕОБУЧЕННЫХ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Хурсов П. С.

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Искра Н. А. – ст. преп.

Зачастую возникает необходимость в восстановлении искаженных изображений. Примерами искажений могут быть артефакты сжатия, наложение шума или посторонних объектов. В данной работе предлагается использование необученных сверточных нейронных сетей для решения этой задачи.

Типичным решением для задачи восстановления изображения является обучение какой-либо генеративной модели на обучающей выборке и использование её для генерации восстановленного изображения. Однако недавние исследования[1] показали, что обучение не является необходимостью, так как некоторое знание о выборке может содержаться прямо в структуре модели. Так сверточным нейронным сетям гораздо проще работать с сигналами, нежели чем с шумом. Это можно объяснить тем, что основным слоем в сверточных нейронных сетях является сверточный слой, который можно рассматривать как некоторый цифровой фильтр над выходом предыдущего слоя.

Используя это можно сформулировать задачу восстановления как генерацию изображения из некоего случайного шума:

$$\operatorname{argmin}_{\theta} \|D(x) - x_0\|, \quad x = f_{\theta}(z^*) \quad (1),$$

где $D(x)$ – некий деградационный процесс над изображением x , x_0 – исходное изображение,

$f_{\theta}(z^*)$ - процесс генерации изображения моделью с параметрами θ из шума z^* .

Таким образом производя минимизацию целевой функции, можно добиться того, чтобы модель могла генерировать целевое изображение из случайного шума.

В качестве модели используется U-Net[2](Рис. 1) с некоторыми изменениями в конфигурации. Модель состоит из 5 блоков понижения\повышения размерности. Блок понижения размерности состоит из сверточного слоя с шагом 2 и 128 фильтрами, слоем групповой нормализации, еще одного сверточного слоя из 128 фильтров, групповой нормализации и функции активации LeakyReLU. Блок повышения размерности имеет похожую конфигурацию, но вместо сверточного слоя с шагом 2 используется билинейная интерполяция для повышения размерности. Для минимизации целевой функции(1) использовался алгоритм Adam[3].

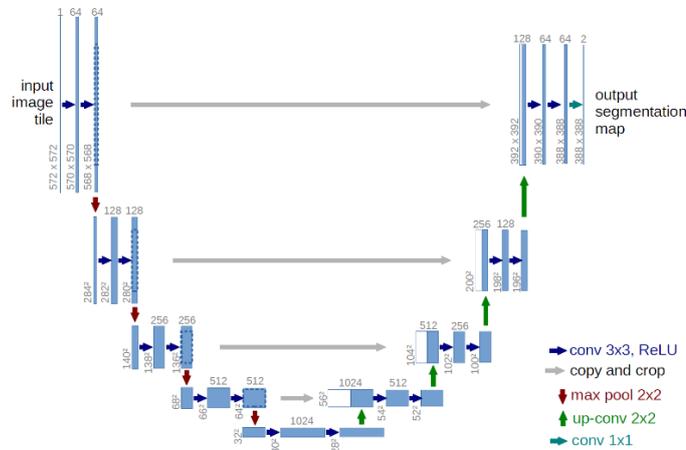


Рисунок 1 – Архитектура нейросети U-Net.

Так как алгоритму понадобится некоторое количество шагов минимизации прежде чем начать воспроизводить помимо сигнала шум, то для решения задачи удаления шума можно использовать экспоненциальное скользящее среднее над результатами каждого шага алгоритма минимизации.

$$I_{n+1} = I_n * \beta + I * (1 - \beta) \quad (2),$$

Где I_{n+1} – значение скользящего экспоненциального среднего на $n+1$ итерации, I_n – значение на n итерации, β – взвешивающий параметр.

Результат удаления шума можно увидеть на рисунке 2.



Рисунок 2 – Результат восстановления зашумленного изображения. Слева направо: оригинальное изображение, изображение с шумом, восстановленное изображение.

Уровень искажения зашумленного изображения по метрике Peak Signal to Noise Ratio(PSNR)

$$PSNR(I, K) = 10 \log_{10} \left(\frac{255}{\frac{1}{NM} \sum_{i,j} (I_{i,j} - K_{i,j})^2} \right) \quad (3),$$

Где I, K – пара изображений размера N на M пикселей.

Как видно из формулы, чем больше изображения похожи друг на друга, тем большим будет метрика PSNR. Для пары оригинального и зашумленного изображения PSNR равен 26.02, а для пары оригинального и восстановленного 30.38. Как видно необученные нейронные сети вполне могут быть использованы для удаления шума с изображений.

Другим применением необученных нейронных сетей для восстановления изображений может

быть задача супер разрешения. Задача состоит в том, чтобы по изображению низкого разрешения получить изображение большего разрешения без потери качества. Для этого нужно в выражении 1 использовать в качестве деградационного процесса функцию понижения разрешения сгенерированного изображения до оригинального.

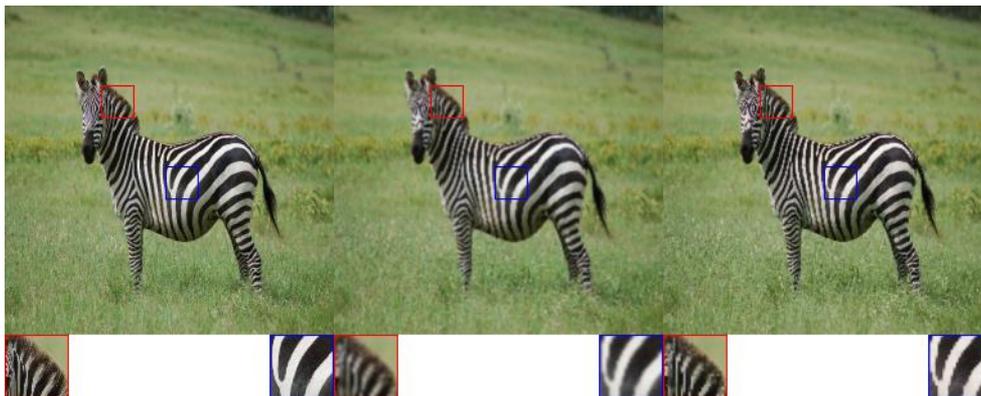


Рисунок 3 – Результат повышения разрешения изображения. Слева направо: оригинальное изображение, изображение восстановленное с помощью бикубической интерполяции, изображение восстановленное с помощью необученной нейронной сети.

Для сравнения использовалось изображение в высоком разрешении (512x512 пикселей). Далее используя метод бикубической интерполяции разрешение понижалось до 256x256 пикселей. После этого разрешение снова повышалось до 512x512 пикселей с помощью бикубической интерполяции и с помощью необученной сверточной нейронной сети (Рис. 3).

Если сравнить изображения по метрике PSNR, то пара оригинального изображения и восстановленного с помощью бикубической интерполяции имеет значение PSNR 26.16, а пара оригинального и восстановленного с помощью необученной нейронной сети 28.22.

Как видно необученные нейронные сети могут быть эффективным инструментом для решения задач восстановления изображений. Однако для этого необходимо знать деградационный процесс. Так, например, можно восстанавливать размытые изображения, если известно ядро размытия.

Список использованных источников:

1. Deep Image Prior / D. Ulyanov, A. Vedaldi, and V. Lempitsky // CVPR, 2018.
2. U-net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, and T. Brox . – MICCAI, 234–241с, 2015.
3. Adam: A method for stochastic optimization / Diederik P. Kingma and Jimmy Lei Ba // ICLR, 2015.