

Р.А. ШАКУН

ШУМОПОДАВЛЕНИЕ НА ИЗОБРАЖЕНИИ С ПОМОЩЬЮ FFDNet

Белорусский государственный университет информатики и радиоэлектроники П. Бровки, 6, Минск, 220013, Беларусь

Аннотация. FFDNet - это новейший метод шумоподавления изображений, основанный на архитектуре сверточной нейронной сети. В отличие от других существующих нейросетей, FFDNet имеет несколько преимуществ, таких как быстрое время выполнения и использование небольшого объема памяти, а также способность обрабатывать широкий диапазон уровней шума с единой сетевой моделью. Сочетание качества шумоподавления и небольшой вычислительной нагрузки делает этот алгоритм привлекательным для практического использования в шумоподавляющих приложениях. В этой работе будут обсуждаться характеристики архитектуры этого алгоритма.

Ключевые слова: Шумоподавление, CNN, FFDNet, нейронные сети.

Abstract. FFDNet is a recent image denoising method based on a convolutional neural network architecture. In contrast to other existing neural network denoisers, FFDNet exhibits several desirable properties such as faster execution time and smaller memory footprint, and the ability to handle a wide range of noise levels with a single network model. The combination between its denoising performance and lower computational load makes this algorithm attractive for practical denoising applications. In this work will be discussed the characteristics of the architecture of this algorithm.

Keywords: Denoising, CNN, FFDNet, neural networks.

Введение

В последнее время, благодаря выдающимся характеристикам, привлекли значительное внимание новые способы шумоподавления на изображениях, основанные на методах глубокого обучения. В частности, дискриминационные методы обучения (например CSF, TNRD) показывают относительно быстрое время вывода и очень хорошие шумоподавляющие характеристики. Тем не менее, существенным недостатком всех этих алгоритмов является то, что конкретная модель должна быть обучена для каждого уровня шума.

Другой популярный подход включает использование сверточных нейронных сетей (CNN), таких, как RBDN, DnCNN и FFDNet. Их производительность выгодно отличается от других современных алгоритмов как количественно, так и визуально. Эти методы состоят из последовательности сверточных слоев с нелинейными активационными функциями между ними.

Наиболее интересной из перечисленных сверточных сетей является FFDNet [1]. Главное отличие FFDNet по отношению к DnCNN заключается в включении предварительного и последующего слоев обработки до и после одного и того же нелинейного отображения DnCNN. Слой предварительной обработки реорганизовывает пиксели входного изображения в многоканальное изображение с четвертным разрешением. Большинство вычислений выполняются в этом меньшем масштабе, что снижает общую сложность алгоритма.

В целом, FFDNet примерно в три раза быстрее DnCNN и более дружелюбный к памяти. Также дополнительный канал содержащий карту шума соединяется со входом на уровне предварительной обработки. Роль карты шума в качестве входного параметра может рассматриваться как компромисс между уменьшением шума и сохранением деталей. Наконец, слой постобработки изменяет выходные данные нелинейного отображения обратно в исходное разрешение. Эти характеристики делают FFDNet привлекательным методом, даже для потребительских приложений, так как он обеспечивает интересный баланс между шумоподавлением, производительностью и сложностью.

Архитектура FFDNet

Для реализации FFDNet используется стандартная архитектура прямой связи, как показано на рисунке 1. Сеть состоит из D сверточных слоев, которые имеют одинаковую структуру. Каждый из них имеет W выходов, а пространственный размер их ядер $K \times K$.

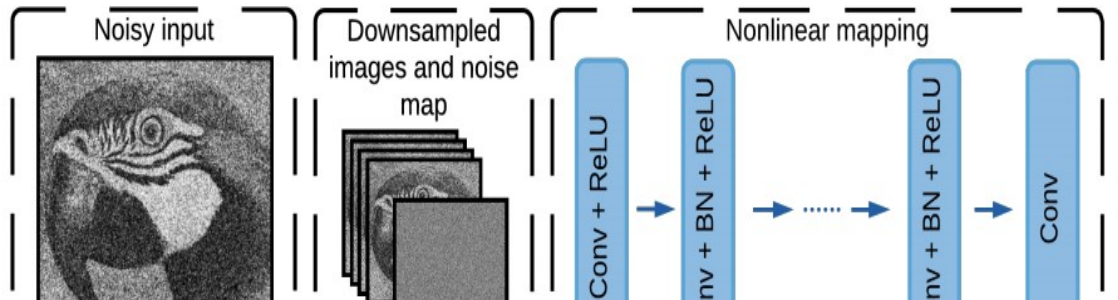


Рисунок 1: Архитектура FFDNet.

На уровне предварительной обработки сеть сначала реорганизовывает пиксели I входного изображения размером $n_{ch} \times h \times w$ в изображение с более низким разрешением размером $4n_{ch} \times h/2 \times w/2$.

Слой F^0 извлекает 2×2 патчи и реорганизовывает их пиксели в разных каналах выходного изображения в соответствии с F^0 :

$$F^0[c, x, y] = I \left[\left\lfloor \frac{c}{4} \right\rfloor, 2x + (c \bmod 2) \right]$$

где $0 \leq c < 4n_{ch}$, $0 \leq x < h$, $0 \leq y < w$. Большая часть обработки будет выполнена в этом уменьшенном масштабе.

Также ко входу добавляется карта оценки шума M такого же разрешения. Эта шумовая карта контролирует компромисс между шумоподавлением и сохранением детализации. Для пространственно

инвариантного гауссовского шума со стандартным отклонением σ , карта шума однородна и все ее элементы равны σ . На рисунке 2 показана схема этого слоя.

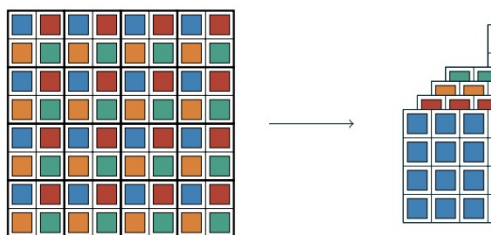


Рисунок 2: Схема слоя масштабирования.

Нелинейное отображение, состоящее из D сверточных слоев ($F^1 \dots F^D$) следует после предварительной обработки слоя F^0 (см. рисунок 1). Каждый из этих слоев состоит из W сверточных фильтров пространственного размера $K \times K$. Выходы слоев $F^1 \dots F^{D-1}$ следуют за точечной функцией активации $\text{ReLU}(\cdot) = \max(\cdot, 0)$. Во время обучения слои нормализации помещаются между сверточным и ReLU слоями от F^2 до F^{D-1} . Во время оценки слои нормализации удаляются и заменяются аффинным слоем. c -й канал d -го слоя, F_c^d можно записать как:

$$F_c^d = \text{ReLU} \left(\sum_{c'}^{D-1} w_{cc'}^d * F_{c'}^{d-1} \right) \text{ for } c$$

где $w_{cc'}^d$ - двумерное ядро свертки размера $K \times K$. Обобщая характеристики из D слоев нелинейного отображения, мы имеем

- слой F^1 : Conv + ReLU. Используется W фильтров размером $(4n_{ch} + 1) \times K \times K$, которые генерируют W карт объектов. В качестве нелинейности используется точечная функция активации ReLU.
- слои $F^2 \dots F^{D-1}$: Conv + BN + ReLU. Используются фильтры W размера $W \times K \times K$. Во время тренировки уровни нормализации размещаются между сверточным и ReLU-слоями.
- слой F^D : Conv. В этом слое используются 4-канальные фильтры размером $W \times K \times K$.

Слой постобработки F^{D+1} увеличивает разрешение выходных данных из слоя F^D до оригинального. Таким образом, он переупаковывает свой ввод размера $4n_{ch} \times h / 2 \times w / 2$ в изображение размер $n_{ch} \times h \times w$, как показано на рисунке 3. Общее количество слоев равно $D + 2$, где D - число сверточных слоев. Пространственный размер сверточных ядер K равен 3. Глубина D установлена на 15 для шумоподавления на изображениях в оттенках серого и 12 для шумоподавления на цветных изображениях. Что касается количество карт характеристик W , оно равно 64 для серого и 96 для цветного изображений. Эти настройки представляют собой хороший компромисс между сложностью и производительности.

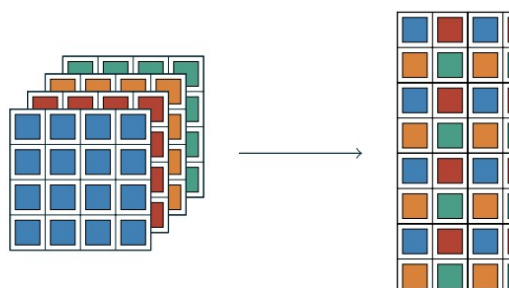


Рисунок 3: Схема слоя постобработки.

Заключение

В данной работе был проанализирован алгоритм шумоподавления на основе сверточной нейронной сети FFDNet. По сравнению с другими алгоритмами FFDNet является более быстрым, эффективным и универсальным. Эти улучшения достигаются благодаря использованию разнообразных

методов, которые описаны выше. Наиболее заметным из таких методов является то, что шумоподавление выполняется на уменьшенных изображениях, что значительно сокращает время работы и объем памяти без ущерба для производительности.

Список литературы

1. K. Zhang, W. Zuo, and L. Zhang, FFDNet: Toward a Fast and Flexible Solution for CNN based Image Denoising, IEEE Transactions on Image Processing, 27 (2018), pp. 4608–4622. <http://doi.org/10.1109/TIP.2018.2839891>.

References

1. K. Zhang, W. Zuo, and L. Zhang, FFDNet: Toward a Fast and Flexible Solution for CNN based Image Denoising, IEEE Transactions on Image Processing, 27 (2018), pp. 4608–4622. <http://doi.org/10.1109/TIP.2018.2839891>.

Сведения об авторах

Шакун Р.А., магистрант Белорусского государственного университета информатики и радиоэлектроники.

Information about the authors

Shakun R.A., master student of Belarusian State University of Informatics and Radioelectronics.

Адрес для корреспонденции

220013, Республика Беларусь,
г. Минск, ул. П. Бровки, д. 6,
Белорусский государственный университет
информатики и радиоэлектроники
тел. +375-29-851-22-25;
e-mail: shakunroman@gmail.com
Шаукн Роман Александрович

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovka st., 6,
Belarusian State University of Informatics and
Radioelectronics
tel. +375-29-851-22-25;
e-mail: shakunroman@gmail.com
Shakun Raman Alexandrovich