

ШУМОПОДАВЛЕНИЕ НА ИЗОБРАЖЕНИИ С ПОМОЩЬЮ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Шакун Р.А., Фоменок В.В.

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

Лукашевич М.М. – к.т.н., доцент

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.

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].

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 $K \times K$.

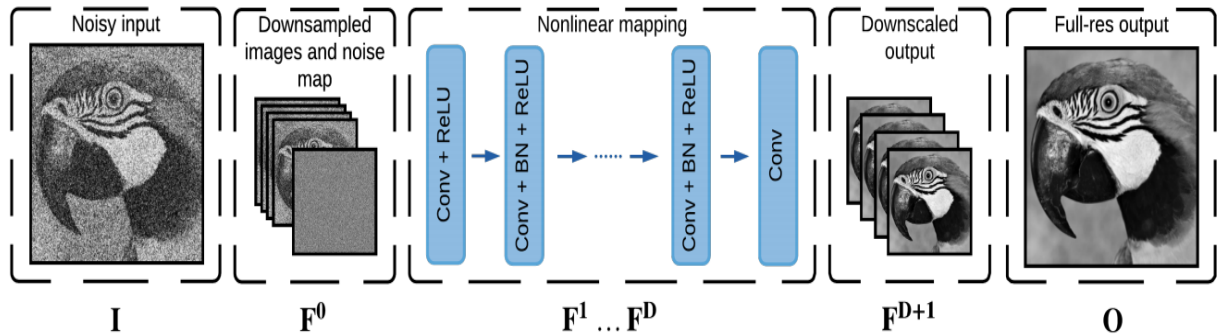


Рисунок 1 – Архитектура FFDNet.

На уровне предварительной обработки сеть сначала реорганизует 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 реорганизует их пиксели в разных каналах выходного изображения в соответствии с F^0 :

$$F^0[c, x, y] = I \left[\left\lfloor \frac{c}{4} \right\rfloor, 2x + (c \bmod 2), 2y + \left\lfloor \frac{c}{2} \right\rfloor \right] \quad (1),$$

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.

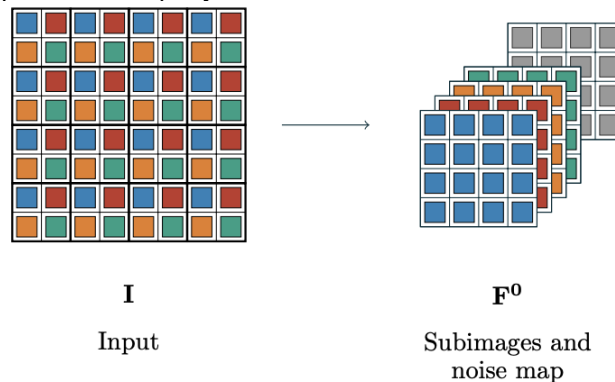


Рисунок 2 – Схема слоя масштабирования.

Нелинейное отображение, состоящее из D сверточных слоев ($F^1 \dots F^D$) следует после предварительной обработки слоя F^0 (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 $ReLU(\cdot) = \max(\cdot, 0)$. 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^{dc} можно записать как:

$$F_c^d = ReLU \left(\sum_{c'=0}^{D-1} w_{cc'}^d * F_{c'}^{d-1} \right) \text{ for } c \in \{0 \dots W - 1\} \quad (2),$$

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 перепакетывает 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 количества карт характеристик W , оно равно 64 для серого и 96 для цветного изображений. Эти настройки представляют собой хороший компромисс между сложностью и производительности.

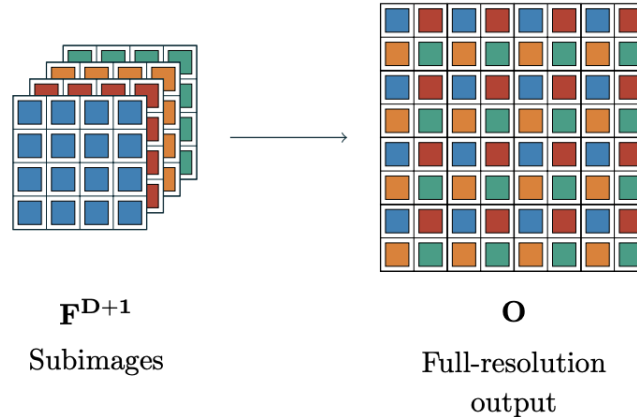


Рисунок 3 – Схема слоя постобработки.

Список использованных источников:

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>.