

Шумоподавление изображений с использованием дискретно-косинусного преобразования

В.В. Алексеев, А.Г. Шоберг

Тихоокеанский государственный университет, Хабаровск

Аннотация: В данной работе рассматривается проблема удаления шумов из изображения на основе алгоритма дискретно-косинусного преобразования (ДКП). Несмотря на свою простоту, алгоритм до сих пор пользуется популярностью в преобразовании изображений. Тем не менее с недавних пор идет сильное развитие сверточных нейронных сетей, оставляя позади «традиционные» методы обработки сигналов. В данной работе исследуется шумоподавление изображений при помощи ДКП и сверточных нейронных сетей и создание интерпретируемой сверточной нейронной сети для получения точных данных. За основу был взят язык программирования Python и библиотека для работы с нейронными сетями – PyTorch. Исходя из этого, модель нейронной сети была обучена на наборе данных «The Berkeley Segmentation Dataset». Эксперименты показали, что обученная нейронная сеть показывает результаты, сопоставимые с традиционными алгоритмами шумоподавления изображений.

Ключевые слова: шумоподавление, сверточная нейронная сеть, дискретно-косинусное преобразование, машинное обучение, обработка сигналов, оператор Кэнни.

Введение

Шумоподавление изображений – одна из самых изученных областей в обработке изображений [1, 2]. Предполагается, что обработка происходит с изображением, подверженным Гауссовскому шуму. Затем наблюдаемое изображение необходимо обработать и получить оригинальное изображение, убрав из него некоторое значение шума. Для решения этой задачи было придумано различное количество методов, в частности – дискретно-косинусное преобразование (ДКП) [1].

За последние годы развитие глубокого обучения произвело революцию в сфере компьютерного зрения. Было предложено множество сверточных нейронных сетей, превосходящих традиционные алгоритмы и позволяющих на основе обучающего набора изображений добиться высочайшей точности работы. Одновременно с этим развивалась сфера шумоподавления, где каждый год появлялись новые методы улучшения данной области. Однако, из-за относительной новизны, представлено мало разработок.

Данная работа основывается на использовании сверточных нейронных сетей в алгоритме ДКП шумоподавления. Поможет в этом язык программирования Python и специальная библиотека для работы с нейронными сетями – Pytorch. Для избежания артефактов на границах объектов, будет применена процедура классификации на основе оператора Джона Кэнни (анг John F. Canny), разработанного в 1986 году. После разделения на области, будет применяться наша нейросеть с заложенным в ней алгоритмом ДКП, и в конечном итоге сравним изображения при помощи соотношения сигнал-шум.

Теория

Локальное дискретное косинусное преобразование (DCT), применяемое в самом популярном стандарте сжатия изображений JPEG, дает разреженное представление локальных фрагментов изображения.

Данный алгоритм шумоподавления разделяет изображение на локальные участки размером 16×16 и удаляет шум с участков с оценкой пороговой обработки в области ДКП. Преобразование 16×16 в среднем приводит к лучшим результатам шумоподавления. Хотя оно дает такую же производительность, как и меньший размер, при низком уровне шума ($\sigma < 30$), при высоком уровне шума оно превосходит меньшую размерность, улучшение в сравнение с размером 8×8 идет в пределах от 0,5 до 2 дБ, при изменении σ от 50 до 100. Размер окна больше 16×16 не приводит к дальнейшему значительному улучшению. Пример базиса ДКП показан на рис.1 [2, 3].

Введение трансляционной инвариантности значительно улучшает оценку порога в ортонормированном базисе. Инвариантное к трансляции шумоподавления ДКП реализуется путем разложения изображения на скользящие перекрывающиеся фрагменты, вычисления шумоподавления ДКП в каждом фрагменте, а затем объединяем шумоподавленные участки с

изображением, усредняя перекрывающиеся пиксели. Инвариантное к трансляции шумоподавление ДКП значительно улучшает производительность шумоподавления, обычно примерно от 2 до 5 дБ, и удаляет блочные артефакты за счет затрат $N \times N$ раз вычислений по сравнению с оценкой с неперекрывающимися участками [4, 5].

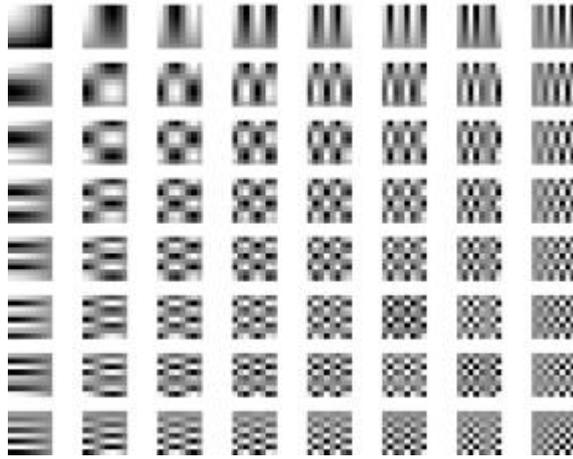


Рис. 1. – ДКП базис

Три цветовых канала цветного изображения обычно сильно коррелируют. Декоррелирование цветовых каналов значения значительно улучшает оценку шумоподавления. Декоррелирование трех каналов в каждом пикселе может привести к представлению разреженного сигнала с одним ненулевым коэффициентом из трех. Таким образом, оценка порогового значения уменьшает дисперсию шума до одной трети, что эквивалентно улучшению оценки соотношения сигнал/шум [6, 7].

В классическом варианте Дискретно-косинусное преобразование переводит определенный массив пикселей изображения в массив значений частотных составляющих Фурье. При шумоподавлении каждая область размером $n \times n$ будет изменяться независимо, т.е. один пиксель будет проходить шумоподавление n^2 раз. Для каждого пикселя, значение шумоподавления (n^2) будет получено путем его усреднения.

Таким образом, можно составить ДКП матрицу:

$$P_{i,j} = \frac{2}{n} C(n) C(m) \cos\left[\frac{(2x+1)\mu\pi}{2p}\right] \cos\left[\frac{(2y+1)\nu\pi}{2p}\right],$$

где: $C(n) = 1$ при $n \neq 1$; $C(m) = 1$ при $m \neq 1$; $C(m) = \frac{1}{\sqrt{2}}$, при $m = 0$; x, y – номер пикселя в блоке по горизонтали и вертикали; i, j – коэффициенты, характеризующие изображение в спектральной плоскости;

Из данной матрицы можно составить наше ДКП:

$$F_{i,j} = \frac{2}{n} C(n) C(m) \sum_{x=0}^{N-1} \sum_{x=0}^{N-1} \cos\left[\frac{(2x+1)\mu\pi}{2p}\right] \cos\left[\frac{(2y+1)\nu\pi}{2p}\right]$$

Данную формулу можно уже использовать для реализации шумоподавления на основе свёрточной нейронной сети.

Свёрточная нейронная сеть

В последнее время, благодаря легкому доступу к крупномасштабным наборам данных и достижениям в методах глубокого обучения, свёрточные нейронные сети продемонстрировали большой успех в решении различных зрительных задач. Для гауссовского шумоподавления легко сгенерировать достаточное количество обучающих данных из набора изображений высокого качества.

Рассмотрим два метода работы со свёрточной нейронной сетью:

1) Остаточное обучение. Изначально оно было придумано для решения проблем снижения производительности, точность обучения снижалась вместе с увеличением входных данных. Это сеть с пропущенными соединениями, которые сопоставляют идентификаторы, объединённые с выходными данными слоя, путем сложения. Данный метод легок в изучении но обладает достаточно низкой скоростью.

2) Пакетная нормализация. Широко используется при обучении моделей свёрточной сети. Метод подает на вход предварительно обработанные данные и имеющие нулевое математическое ожидание и единичную дисперсию. Пакетная нормализация обладает рядом

преимуществ, таких, как быстрое обучение, более высокая производительность и низкая чувствительность к инициализации.

Наша сеть состоит из сверточного слоя размерностью $p \times p$ и ядра свертки с выходом на p^2 каналов. Задача ядра - поэлементно выполнять умножение входных данных, суммировать и подавать на выход одно значение пикселя. В результате свертки, получим массив значений весов или матрицу признаков. Пример операции свертки изображена на рисунке 2. Затем данные с этого массива попадают на функцию активации, которая определяет выходной сигнал. Важная составляющая сверточной сети – нахождение определенных признаков и объектов, позволяющих добиться наивысшей точности [8, 9].

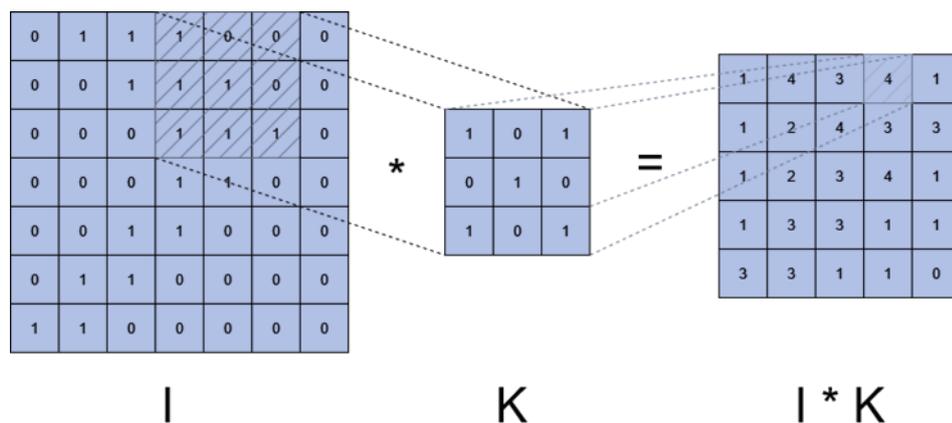


Рис. 2. – Операция свертки

Затем данные с этого массива попадают на функцию активации, которая определяет выходной сигнал. Важная составляющая сверточной сети – нахождение определенных признаков и объектов, позволяющих добиться наивысшей точности. Все эти операции могут происходить по несколько раз, т.е в несколько слоев. Полная схема работы сверточной сети изображена на рис.3.

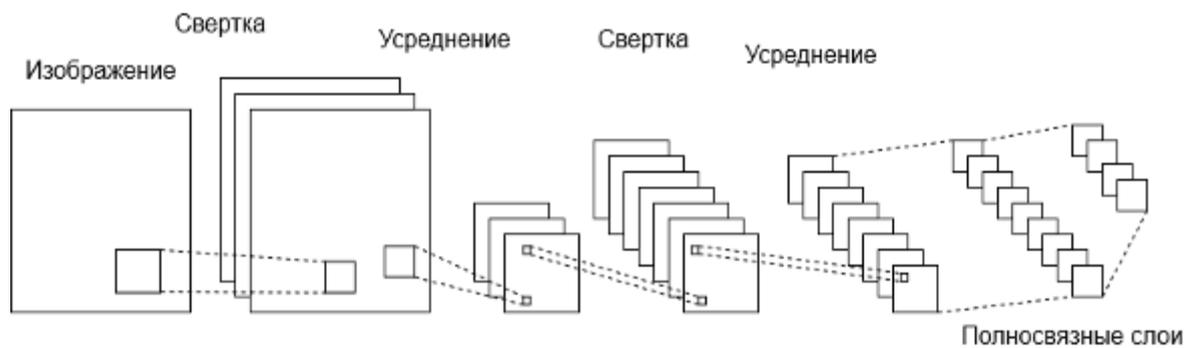


Рис. 3. – Схема работы сверточной сети

Первый слой представляет собой обратимый оператор понижающей дискретизации, который преобразует зашумленное изображение у размером $m \times n$ в четыре вложенных изображения с пониженной дискретизацией. Далее идет объединение с настраиваемыми картами уровня шума с пониженной дискретизацией $m = \{m^1, \dots, m^c\}$ с вложенными изображениями с пониженной дискретизацией, чтобы сформировать тензор \hat{y} размером $\frac{m}{2} \times \frac{n}{2} \times 5c$, в качестве входных данных для сверточной сети, где m – это однородная карта, все элементы которой равны шуму, при этом при понижающей дискретизации не будет потеряна никакая информация [10, 11].

С данным тензором \hat{y} в качестве входных данных, сверточная сеть будет состоять из серий сверточных слоев размером 3×3 . Каждый слой состоит из трех типов операций: Свертки, функции активации и пакетной нормализации. Свертка и функции активации используются для первого сверточного слоя. Для второго комбинация всех перечисленных операций, для финального слоя обычная операция свертки. Для сохранения размера карт объектов неизменных используется операция добавления нулей после каждой свертки.

После финального сверточного слоя, применяется операции увеличения масштаба в качестве оператора, обратного оператору понижающей дискретизации, применяемому на этапе ввода для получения

остаточного зашумленного изображения \hat{y} размером $m \times n \times c$. Затем изображение с убраным шумом получается с помощью $\hat{x} = y - \hat{y}$ [12].

Проверка результатов

В данной работе использовался язык программирования Python и библиотека для работы с нейросетями Pytorch. Первым делом нужно было на основе данных, полученных выше обучить нашу нейросеть и получить готовую модель для работы с шумоподавлением. Модель обучалась на наборе данных «The Berkeley Segmentation Dataset». Обучение происходило на 500 изображениях, в два сверточных слоя.

В первую очередь на изображение добавлялся гауссовский шум с величиной 15, 25 и 50. Модель обучена работать только с данными значениями шумов. На рис.4 показан последовательный пример работы программы. На вход подается зашумленное изображение, затем изображение переводится в частотную область, определяются объекты и их границы и затем происходит шумоподавление изображения [13-15].

Для того, чтобы понять эффективность работы шумоподавления, была произведена оценка по трем различным характеристикам: пикового соотношения сигнал/шум (PSNR), Среднеквадратичное отклонение, и Индекс структурного сходства (SSIM); занесенные в таблицу 1,2,3 соответственно, а также сравнение с работой стандартного алгоритма ДКП.

Для оценки соотношения сигнал/шум сначала измеряется среднеквадратичная ошибка (СКО или MSE), вычисляемая по формуле:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |I(i, j) - K(i, j)|^2,$$

где I и K – входное и выходное изображение; m,n – размеры изображения.

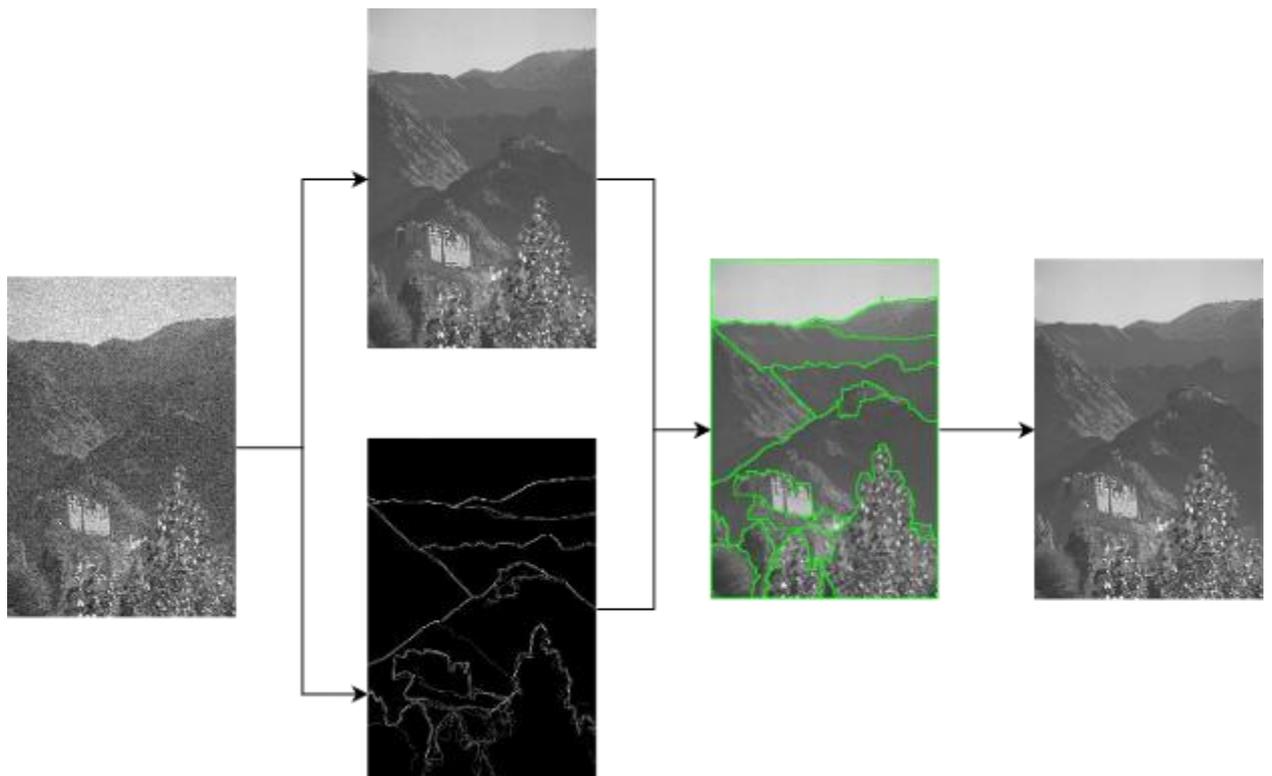


Рис. 4. – Пример работы программы

Затем на основе СКО вычисляем соотношение сигнал/шум

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right),$$

где MAX – это максимальное значение, принимаемое пикселем изображения;

В случае индекса структурного сходства, он высчитывается по следующей формуле:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)},$$

где: μ_x - среднее x и y; σ_x^2, σ_y^2 - дисперсия x и y; C – переменные;

На рис.5, продемонстрирована работа фотографий, принимавших участие в сборе характеристик и тестировании программы.

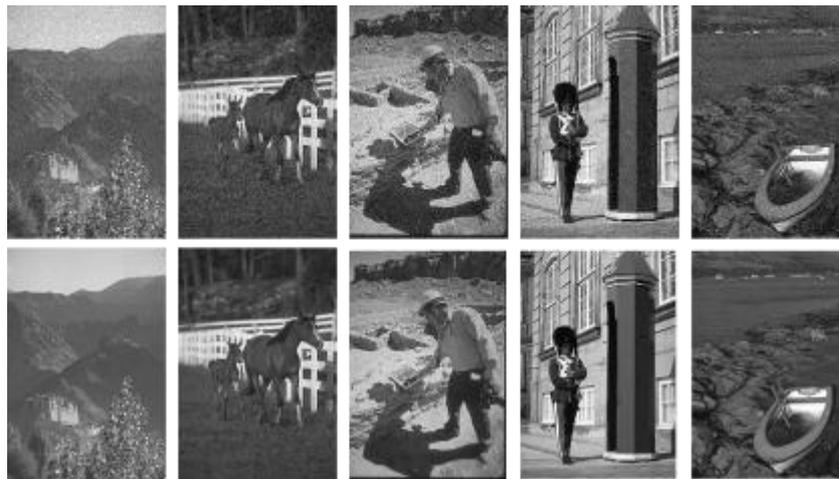


Рис. 5. – Работа программы шумоподавления при значении шума 25.

Таблица №1

Соотношение сигнал/шум при различных вхождениях шума

Изображения	1	2	3	4	5
Уровень шума	15				
ДКП + Сверточная сеть	32,43	31,24	30,31	30,64	32,58
ДКП	28,91	27,32	29,19	29,56	31,22
Уровень шума	25				
ДКП + Сверточная сеть	29,85	27,81	28,52	28,48	30,56
ДКП	28,91	27,32	29,19	29,56	31,22
Уровень шума	50				
ДКП + Сверточная сеть	28,55	26,2	26,59	26,32	28,78
ДКП	25,67	27,73	25,48	25,67	27,73

Таблица №2

Сравнение характеристик изначального и конечного изображения на основе среднеквадратичного отклонения.

Изображения	1	2	3	4	5
Уровень шума	15				
ДКП + Сверточная сеть	37,3	60,89	49,52	55,62	35,78
ДКП	40,23	63,71	52,11	57,03	37,97
Уровень шума	25				
ДКП + Сверточная сеть	67,02	90,04	91,66	93,07	57,92
ДКП	70,23	94,12	93,25	95,94	60,04
Уровень шума	50				
ДКП + Сверточная сеть	139,7	210,7	211,87	167,31	107,1
ДКП	145,32	215,43	216,74	171,5	112,5

Таблица №3

Сравнение характеристик на основе SSIM.

Изображения	1	2	3	4	5
Уровень шума	15				
ДКП + Сверточная сеть	0,898	0,863	0,924	0,827	0,846
ДКП	0,857	0,827	0,887	0,784	0,808
Уровень шума	25				
ДКП + Сверточная сеть	0,841	0,789	0,877	0,738	0,776
ДКП	0,803	0,745	0,841	0,694	0,731
Уровень шума	50				
ДКП + Сверточная сеть	0,725	0,659	0,776	0,606	0,69
ДКП	0,681	0,618	0,744	0,575	0,69

На основе данных, по уровню шума, равному 15, построим диаграммы, на которых будут наглядно видны улучшения в работе со сверточной нейронной сетью (рис.6,7,8).

На графиках видно, что нейронная сеть показывает лучшие соотношения в сравнении с обычным алгоритмом ДКП. Лучше соотношение сигнал/шум, ниже среднеквадратичное отклонение, и выше процент индекса соответствия. Несмотря на данные показатели, это все же не самые лучшие результаты, для процедуры шумоподавления.

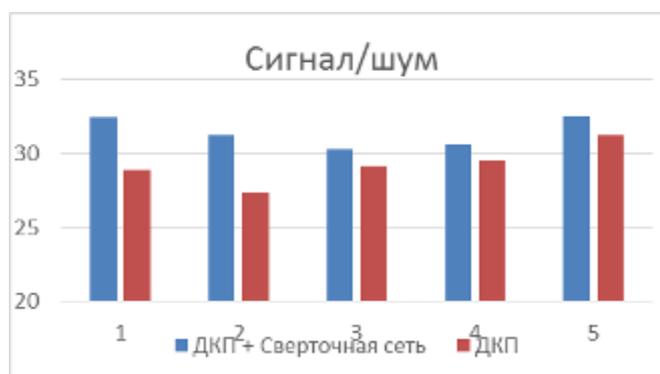


Рис. 6. – Диаграмма на основе характеристик Сигнал/шум

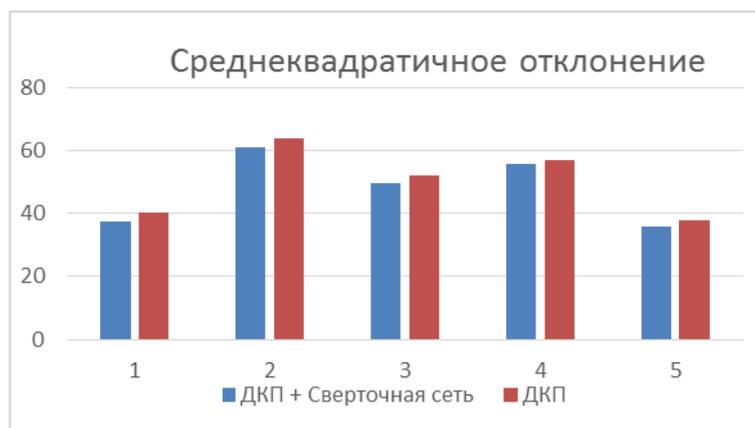


Рис 7. – Диаграмма на основе характеристик среднеквадратичного отклонения.

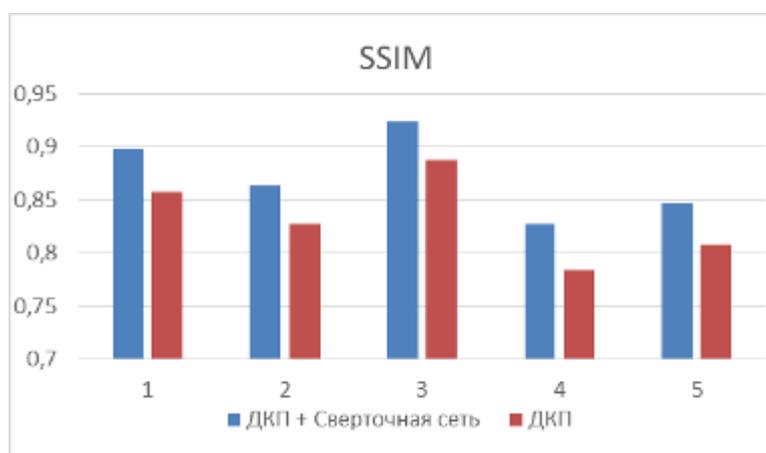


Рис 8. – Диаграмма на основе характеристик SSIM

В результатах с более высоким гауссовским шумом, работа обоих алгоритмов начала давать результаты хуже. Несмотря на это, нейросеть продемонстрировала лучшие результаты. Однако это не предел, и в дальнейшем есть возможность по улучшению шумоподавления.

Заключение

ДКП – один из самых популярных методов обработки изображений. Совместно со сверточной нейронной сетью, можно добиться превосходных результатов шумоподавления. Использование данной комбинации показало себя по значению отношения сигнал/шум выше классических алгоритмов.

Это определенно не самый лучший результат, которого можно добиться, но даже так работа нейросети показала себя эффективнее.

Литература

1. Хрящев Д.А. Повышение качества изображений, полученных в условиях недостаточной освещенности // Инженерный вестник Дона. 2013. №3. URL: ivdon.ru/ru/magazine/archive/n3y2013/1796.
2. Толкачев Д.С. Повышение точности калибровки внешних параметров видеокамеры // Инженерный вестник Дона. 2013. №3. URL: ivdon.ru/ru/magazine/archive/n3y2013/1840.
3. Галушкин А. И. Нейронные сети: основы теории. М.: РиС, 2014. 496 с.
4. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005. 1066 с.
5. Ростовцев В.С. Искусственные нейронные сети. Киров: ВятГУ, 2014. 208 с.
6. Yu G., Sapiro G. DCT image denoising: a simple and effective image denoising algorithm // Image Processing On Line. 2011. V. 1. pp. 292–296.
7. Donoho D.L., Johnstone I.M. Ideal spatial adaptation by wavelet shrinkage // Biometrika. 1994. V. 83. pp. 425-455.
8. Coll B.B., Morel J.-M. A non-local algorithm for image denoising // IEEE Conference on Computer Vision and Pattern Recognition. 2005. V. 2. pp. 60–65.
9. Elad M., Aharon M. Image denoising via sparse and redundant representations over learned dictionaries // IEEE Transactions on Image Processing. 2006. V. 15. №12. pp. 3736-3745.
10. Coll B.B., Morel J.-M. A review of image denoising algorithms, with a new one // SIAM Journal on Multiscale Modeling and Simulation. 2005. V. 4. №2. pp. 490-530.

11. Xie J., Xu L., Chen E. Image denoising and inpainting with deep neural networks // Advances in Neural Information Processing Systems. 2012. V. 25. pp. 341–349.
12. Zhang K., Zuo W., Gu S., Zhang L. Learning deep CNN denoiser prior for image restoration // Proceedings IEEE Conference on Computer Vision and Pattern Recognition. 2017. pp. 3929–3938.
13. Zhang K., Zuo W., Zhang L. FFDNet: Toward a Fast and Flexible Solution for CNN based Image Denoising // IEEE Transactions on Image Processing. 2018. V. 27. №9. pp. 4608–4622.
14. Zhang K., Zuo W., Chen Y., Meng D., Zhang L. Beyond a gaussian denoiser: residual learning of deep CNN for image denoising // IEEE Transactions on Image Processing. 2017. V. 26. №7. pp. 3142–3155.
15. Lefkimmiatis S. Non-local color image denoising with convolutional neural networks // Proceedings IEEE Conference on Computer Vision and Pattern Recognition. 2017. pp. 3587–3596.
16. Lebrun M. An Analysis and Implementation of the BM3D Image Denoising Method // Image Processing On Line. 2012. V. 2. pp. 175–213.

Bibliography

1. Khryashchev D.A. Inzhenernyj vestnik Dona. 2013. №3. URL: ivdon.ru/ru/magazine/archive/n3y2013/1796.
 2. Tolkachev D.S. Inzhenernyy vestnik Dona. 2013. №3. URL: ivdon.ru/ru/magazine/archive/n3y2013/1840.
 3. Galushkin A. I. Neyronnye seti: osnovy teorii [Neural networks: fundamentals of theory]. M.: RiS, 2014. 496 p.
 4. Gonsales R., Vuds R. Tsifrovaya obrabotka izobrazheniy [Digital image processing]. M.: Tekhnosfera, 2005. 1066 p.
 5. Rostovtsev V.S. Iskusstvennye neyronnye seti [Artificial neural networks]. Kirov: VyatGU, 2014. 208 p.
-



6. Yu G., Sapiro G. Image Processing On Line. 2011. V. 1. pp. 292–296.
7. Donoho D.L., Johnstone I.M. Biometrika. 1994. V. 83. pp. 425-455.
8. Coll B.B., Morel J.-M. IEEE Conference on Computer Vision and Pattern Recognition. 2005. V. 2. pp. 60–65.
9. Elad M., Aharon M. IEEE Transactions on Image Processing. 2006. V. 15. №12. pp. 3736-3745.
10. Coll B.B., Morel J.-M. SIAM Journal on Multiscale Modeling and Simulation. 2005. V. 4. №2. pp. 490-530.
11. Xie J., Xu L., Chen E. Advances in Neural Information Processing Systems. 2012. V. 25. pp. 341–349.
12. Zhang K., Zuo W., Gu S., Zhang L. Proceedings IEEE Conference on Computer Vision and Pattern Recognition. 2017. pp. 3929–3938.
13. Zhang K., Zuo W., Zhang L. FFDNet: IEEE Transactions on Image Processing. 2018. V. 27. №9. pp. 4608–4622.
14. Zhang K., Zuo W., Chen Y., Meng D., Zhang L. IEEE Transactions on Image Processing. 2017. V. 26. №7. pp. 3142–3155.
15. Lefkimmiatis S. Proceedings IEEE Conference on Computer Vision and Pattern Recognition. 2017. pp. 3587–3596.
16. Lebrun M. Image Processing On Line. 2012. V. 2. pp. 175–213.

Дата поступления: 22.12.2023

Дата публикации: 1.02.2024