

Частотное разделение характеристик изображения в задачах сверхразрешения

В.В. Алексеев, А.Г. Шоберг

Тихоокеанский государственный университет, Хабаровск

Аннотация: Сверхразрешение изображений – это популярная задача, цель которой состоит в переводе изображений из низкого разрешения в высокое. Для этой задачи часто используются свёрточные сети. Свёрточные нейронные сети имеют большое преимущество в обработке изображений. Но, несмотря на это, часто информация может теряться при обработке, а увеличение глубины и ширины сети может усложнить дальнейшую работу. Для решения этой проблемы используется преобразование данных в частотную область. В данной работе изображение делится на высокочастотные и низкочастотные участки, где больший приоритет дается первым. Затем, при помощи проверки качества и визуальной оценки, проведен анализ метода и вывод относительно работы алгоритма.

Ключевые слова: сверхразрешение, низкое разрешение, высокое разрешение, дискретно-косинусное преобразование, сверточные-нейронные сети.

Введение

Задача сверхразрешения (СР) – это задача восстановления изображения высокого разрешения (ВР) из низкого разрешения (НР). На практике это позволяет преодолеть ограничения по выполнению задач, связанных с компьютерным зрением, улучшить обработку в медицинской сфере, а также использовать в сфере безопасности и наблюдения. [1, 2]

При этом большинство методов решения данной проблемы имеют значительные требования в объемах памяти и вычислительных ресурсов. [3-5]. В связи с этим задача СР является до сих пор актуальной. Вместо разработки глубоких сетей со сложной архитектурой в данной работе, представлена легковесная версия сети, которая разделяет изображение на низкие и высокие частоты, при этом обеспечивая связь между ними. Так как низкие частоты сохраняются при операциях понижающей дискретизации,

следовательно, они могут быть восстановлены из входного сигнала. Большая часть мощности сети же будет направлена на высокие частоты, при этом сохраняя низкочастотную информацию, что позволит улучшить качество реконструкции. Плюсом данного метода является возможность применять его в различных моделях СР без необходимости изменения архитектуры сети или настройки параметров. Еще одной задачей будет улучшение производительности сети без потери качества.

Сверточная нейронная сеть

Основой данной сети, служит сверточная нейронная сеть. Данная сеть позволила достичь, множественных успехов в сфере обработки изображений. Сеть состоит из множественных слоев называющихся сверточными. Каждый слой начинает обрабатывать изображение полностью, затем рассматривая с каждым слоем меньшие по объему участка [6]. Пример такой сети изображен на рисунке 1.

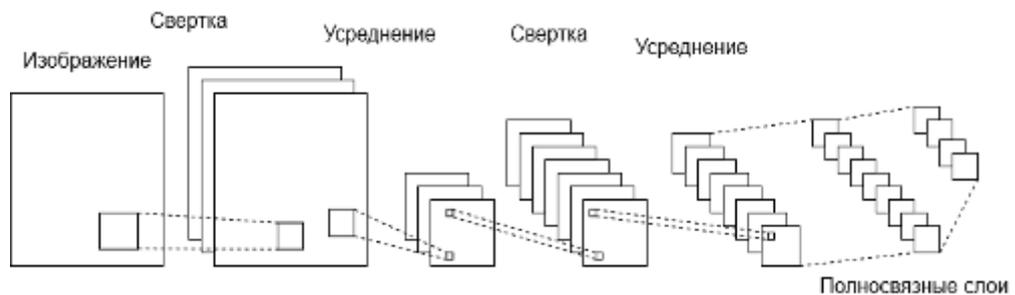


Рис. 1. – Пример архитектуры сверточной нейронной сети

Работа сверточного слоя основана на операции свертки. Пример операции изображен на рисунке 2. На каждом слое за сверточным следует функция активации, которая вызывает нелинейность в поведении сети. Таких функций существует множество видов, например: сигмоида, бинарная, тождества и т.д. В данной работе используется сигмоидная функция активации. Это нелинейная функция преобразующая входное значение в

диапазон от минус до плюс бесконечности. График сигмоиды изображен на рисунке 3. [7]

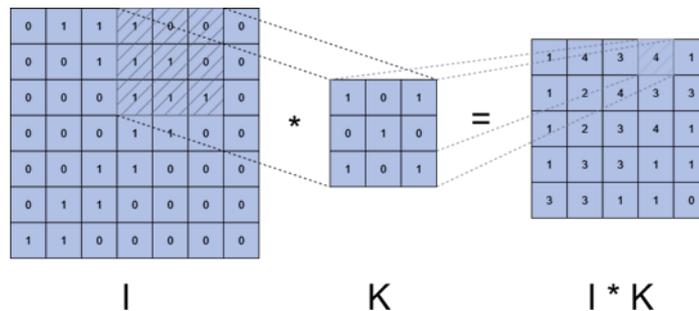


Рис. 2. – Операция свертки

Важное условие при работе с сетью, сделать ее легко применимой к различным моделям СР без необходимости глобально изменять параметры сети и ее архитектуру. Это особенно важно при работе со сверточной сетью, так как расширение или увеличение ширины и глубины сети, может повысить потребление памяти и вычислительной мощности.

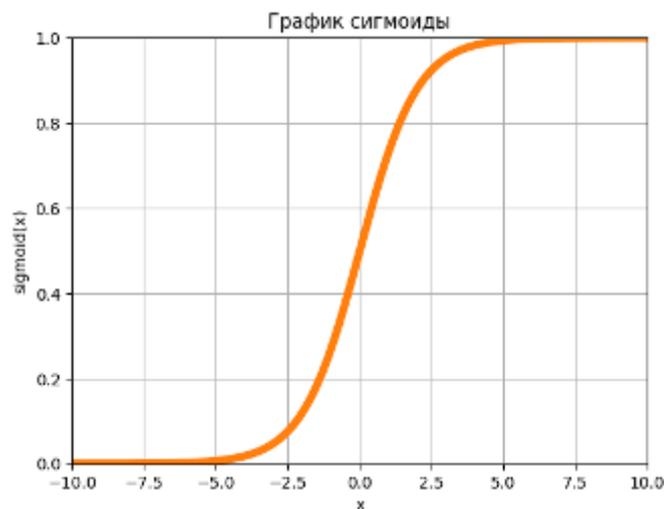


Рис. 3. – График сигмоиды

Методы СР на основе сверточных сетей часто ограничивают высокочастотную информацию в изображениях НР [8, 9]. Выходные карты

признаков сверточного слоя также рассматриваются как сочетание информации на низких и высоких частотах. Низкочастотная информация может быть передана в конечный выход без существенных вычислений. Высокочастотная информация состоит из более мелких деталей, соответственно, на нее лучше делать упор при обработке, т.к это наиболее важная часть в задачах СР.

Частотное преобразование

Дискретно косинусное преобразование (ДКП) – это вид преобразования, в котором пространственная информация представляется в виде частот. Для изображений ДКП переводит определенный массив пикселей изображения в массив значений составляющих Фурье [10-11].

Для двумерного дискретного сигнала f и его представления в частотной области F , ДКП описывается следующим образом:

$$F[u, v] = a(u)a(v) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f[x, y] \gamma(x, y, u, v)$$
$$\gamma(x, y, u, v) = \cos \frac{\pi(2x+1)u}{2N} \cos \frac{\pi(2y+1)v}{2M},$$
$$a(u) = \begin{cases} \sqrt{\frac{1}{N}}, u = 0 \\ \sqrt{\frac{2}{N}}, u \neq 0 \end{cases}$$

Архитектура сети

Задача ядра - постепенно умножать входные данные. Пример операции свертки изображен на рисунке 4. Слой объединения снижает размерность этих матриц, деля их на блоки определенного размера, где для каждого блока вычисляется функция.

Сеть состоит из модуля нелинейного отображения и модуля реконструкции. Для извлечения исходных значений из входного изображения НР применяется только один слой 3×3 .

$$H_0 = H(I_{LR})$$

где I_{LR} - вход; H_0 - исходные значения; H - сверточный слой;

Затем используем модуль нелинейного отображения, который состоит из нескольких наложенных друг на друга объектов. Для извлечения большего количества иерархической контекстной информации было использовано слияние признаков для всех объектов. Затем используем модуль реконструкции, содержащий сверточные слои и слои перетасовки пикселей, для увеличения выборки объектов до размера ВР.

Для оптимизации параметров сети, в качестве функции потери L мы используем обучающий набор из N пар изображений НР и ВР. Сеть обучена для минимизации функции потерь.

$$L = \frac{1}{N} \sum_{i=1}^N \| CP - HP \|$$

Изображение можно разделить на низкочастотную и высокочастотную составляющие. Низкочастотные составляющие представляют собой плавно меняющиеся структуры большой величины [12-14]. Высокочастотная составляющая, наоборот, мелкие, быстро меняющиеся структуры. Аналогичным образом, выходные карты характеристик сверточного слоя могут быть разложены на такие же объекты с разными частотами. Высокочастотная информация обрабатывается с помощью операций более высокой сложности (по количеству параметров и нелинейности), низкочастотная обрабатывается с помощью операций низкой сложности, чтобы компенсировать увеличение объема вычислений.

На рисунке 4 представлена схема архитектуры сети. Она состоит из двух путей, каждый из которых отвечает за различные функции. Каждый путь имеет вначале сверточный слой 1×1 .

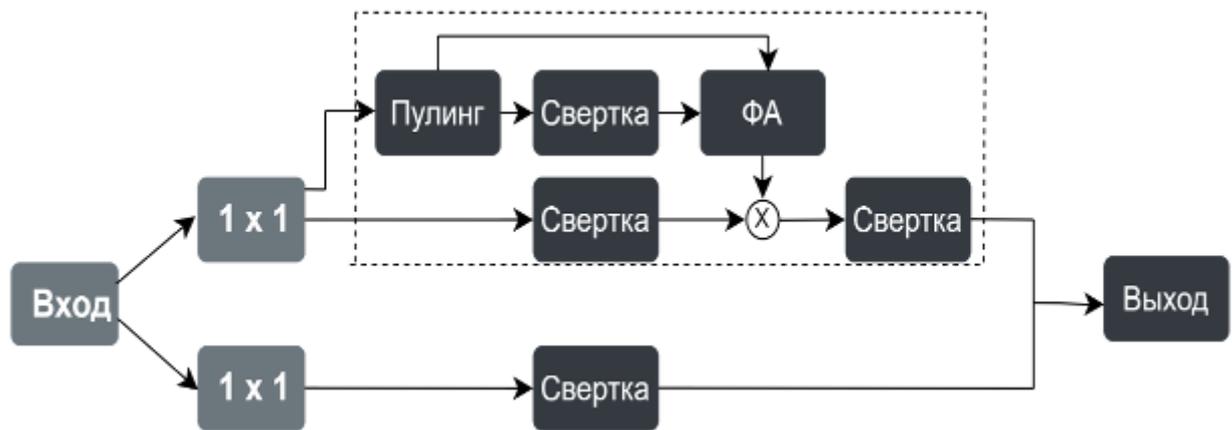


Рис. 4. – Архитектура работы сверточной сети

Над низкочастотной информации выполняется простая операция свертки 3×3 , где захватываются крупные участки изображений, для экономии вычислений.

По второму пути идет высокочастотная информация. Там она попадает на слой объединения, где используется функция среднего значения. Затем информация подвергается дополнительной дискретизации с помощью бикубической интерполяции для получения нового тензора того же размера, как у входных данных.

Сеть использует бикубическую понижающую дискретизацию в качестве стандартной модели понижения для генерации изображений НР из оригинальных, в масштабе $\times 2$, $\times 3$, $\times 4$.

Экспериментальная часть

Для данной работы использовался специальный набор изображений с целью обучения, содержащий около 1000 изображений высокого качества. Для проверки результатов были использованы различные наборы изображений, подготовленных специально для работы с нейронными сетями.

Для количественной оценки работы использовались два распространенных параметра – отношение сигнал-шум (PSNR) и индекс структурного сходства (SSIM). [14]

Прежде чем вычислять PSNR, необходимо вычислить среднеквадратичную ошибку (СКО или MSE), которая вычисляется 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 изображения; m, n – размеры изображения.

Затем на основе вычисляем соотношение сигнал/шум (PSNR):

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

где MAX – это максимальное значение, принимаемое пикселем изображения;

Индекс структурного сходства вычисляется по 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;

Все данные будут занесены в таблицу 1, в таблице 2 будут занесены характеристик из классического метода для сравнения с нашим. На рисунке 5, изображена визуальная оценка работы сети, взятых из набора изображений, под каждый вид масштабирования (x2, x3, x4) соответственно. Для лучшего понимания, из каждого ВР изображения взяты мелкие участки для лучшей проверки качества. Также, на основе данных с таблиц, сделаны графики, для лучшего понимания разницы между моделями, они изображены на рисунках 6 и 7.

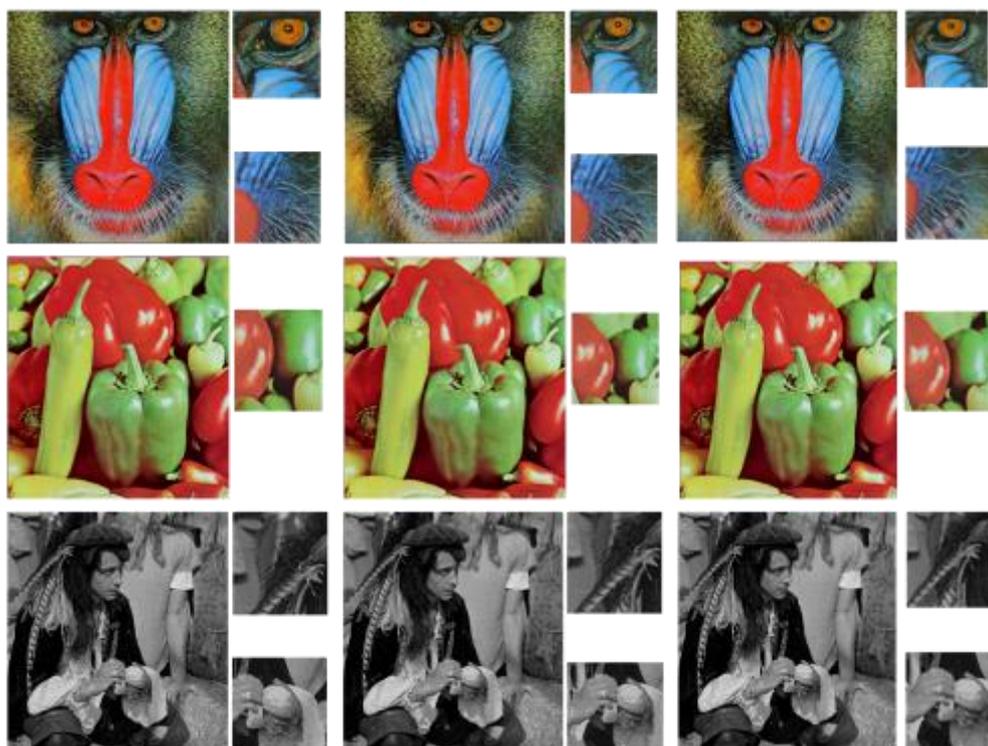


Рис.5. – Изображения СР с масштабом x2, x3 и x4.

Таблица №1

Количественные характеристики для нашей сети

Скалирование	1 набор		2 набор		3 набор		4 набор		5 Набор	
	Частотное разделение									
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
x2	37,84	0,9592	33,24	0,9120	32,09	0,8972	32,05	0,9201	37,36	0,9752
x3	33,62	0,9207	29,96	0,8384	28,94	0,8032	28,12	0,8421	32,94	0,9420
x4	31,94	0,8892	28,75	0,7754	27,42	0,7294	25,52	0,7741	28,91	0,8952

Таблица №2

Количественные характеристики через классический метод

Скалирование	1 набор		2 набор		3 набор		4 набор		5 Набор	
	Классический метод									
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
x2	37,53	0,9556	33,03	0,9025	31,82	0,8960	30,76	0,9102	37,20	0,9705
x3	34,01	0,9224	28,74	0,8217	28,06	0,7966	27,14	0,8520	37,17	0,9456
x4	31,42	0,8841	27,05	0,7730	27,19	0,7251	25,12	0,7524	28,82	0,8901

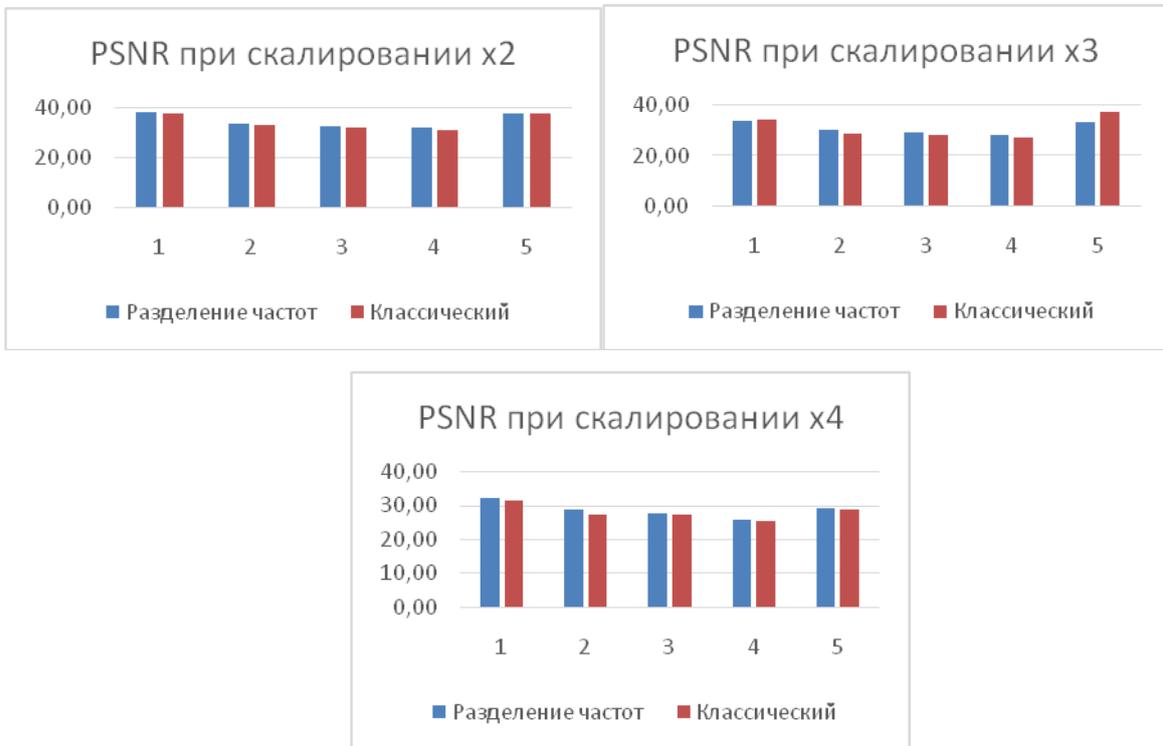


Рис. 6. – Соотношение сигнал/шум при различном масштабировании.

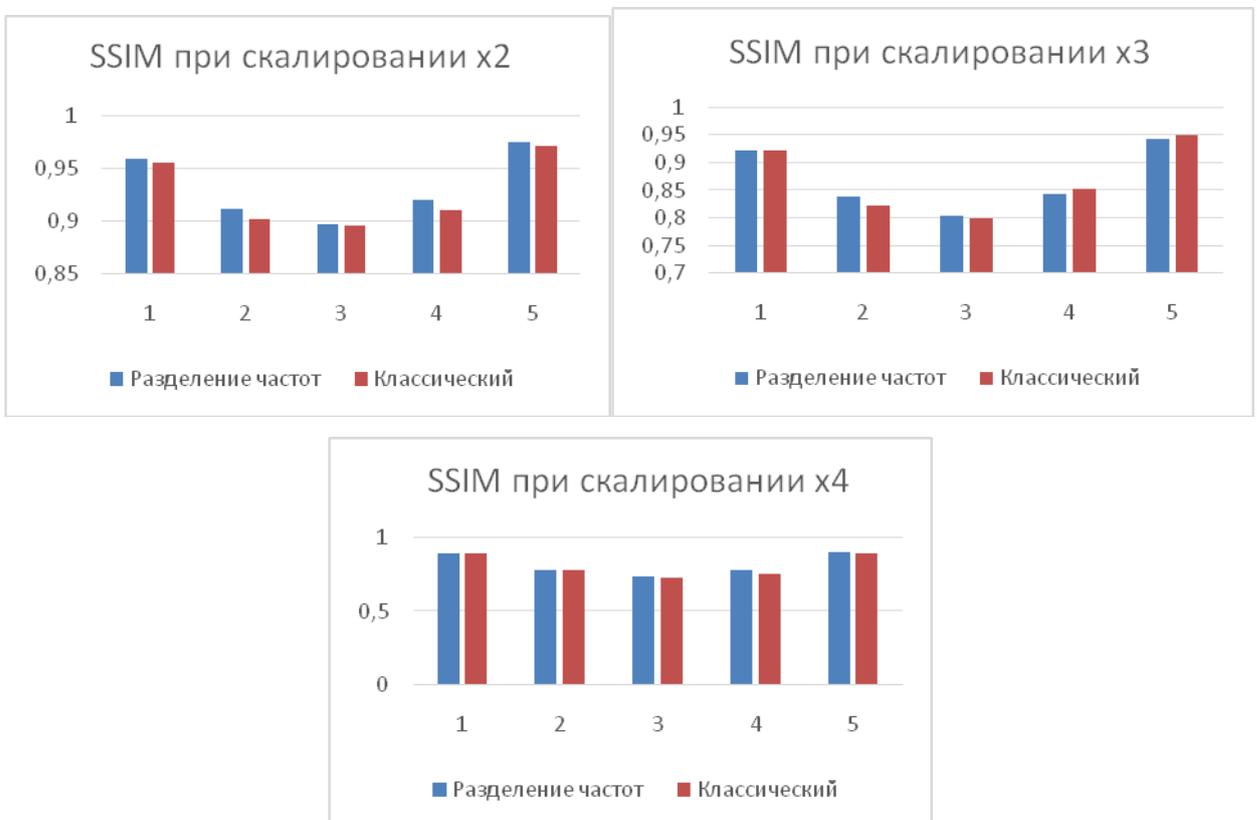


Рис. 7. –Индекс структурного сходства при различном масштабировании.

Заключение

На базе преобразования в частотную область и разделения сигналов на низкочастотные и высокочастотные была выполнена задача сверхразрешения. В основе её лежала попытка улучшить существующие методы и уменьшить объем памяти и потребление вычислительной мощности. На основе полученных данных можно сделать вывод об эффективности сети, ее результатов, схожих с привычными методами выполнения сверхразрешения, как с точки зрения визуальной оценки, так и параметрической, на основе оценок сигнал/шум и индекса структурного сходства. Хотя и не во всех случаях, но в большинстве, сеть показывала лучшие результаты, при меньшей вычислительной мощности.

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. Kim J., Lee J.K., Lee K.M. Accurate image super-resolution using very deep convolutional networks // IEEE Conf. on Computer Vision and Pattern Recognition, 2016, pp. 1646–1654.
4. Kim. J, Kwon Lee. J, Mu Lee. K. Deeply-recursive convolutional network for image super-resolution // IEEE Conf. on Computer Vision and Pattern Recognition, 2016, pp. 1637–1645.
5. Lai.W.-S, Huang .J.-B, Ahuja. N, and Yang .M.-H. Deep laplacian pyramid networks for fast and accurate super-resolution // IEEE Conf. on Computer Vision and Pattern Recognition, 2017, pp. 624–632.

6. Галушкина А. И. Нейронные сети: основы теории. – М. // РиС, 2014. – 496с
 7. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М. // Техносфера, 2005. – 1066 с.
 8. Ростовцев В.С. Искусственные нейронные сети: учебник // Киров: ВятГУ, 2014. – 208 с
 9. Donoho D.L., Johnstone I.M. Ideal spatial adaptation by wavelet shrinkage. // *Biometrika*, 1994. V.83. pp. 425-455
 10. Guo T., Mousavi H. S., Vu T.H., Monga V. Deep wavelet prediction for image super-resolution // *IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 1100–1109.
 11. Dong C., Loy C.C., Tang X. Accelerating the super-resolution convolutional neural network // *ECCV*, 2016. pp 1-17.
 12. Rao K.R., Yip P. Discrete Cosine Transform: Algorithms, Advantages, Applications // *Academic Press Professional, Inc., USA*, 1990. 490 p.
 13. Sun L., Ma C., Chen Y., Zheng Y., Shim H.J., Wu Z., Jeon B. Low rank component induced spatial-spectral kernel method for hyperspectral image classification // *IEEE Transactions on Circuits and Systems for Video Technology*, 2020. V. 30. № 10. pp. 3829–3842.
 14. Wang Y., Wang L., Wang H., Li P. Resolution-aware network for image super-resolution, // *IEEE Transactions on Circuits and Systems for Video Technology*, 2018. V. 29. №. 5. pp. 1259–1269.
 15. Ledig C., Theis L., Husz' F., Caballero J., Aitken A. P., Tejani A., Totz J., Wang Z., Shi W. Photo-realistic single image super-resolution using a generative adversarial network // *IEEE Conference on Computer Vision and Pattern Recognition*, 2018. pp. 105–114.
-

16. Sheikh, H.R., Bovik, A.C., De Veciana, G. An information fidelity criterion for image quality assessment using natural scene statistics // IEEE Trans image processing, 2005. V 14. №12. pp 2117–2128.

References

1. Khryashchev D.A. Inzhenernyj vestnik Dona. 2013. №3. URL: ivdon.ru/ru/magazine/archive/n3y2013/1796.
2. Tolkachev D.S. Inzhenernyj vestnik Dona. 2013. №3. URL: ivdon.ru/ru/magazine/archive/n3y2013/1840.
3. Kim. J, Kwon Lee. J, Mu Lee. K IEEE Conf. on Computer Vision and Pattern Recognition, 2016, pp. 1646–1654.
4. Kim. J, Kwon Lee. J, Mu Lee. K. IEEE Conf. on Computer Vision and Pattern Recognition, 2016, pp. 1637–1645.
5. Lai.W.-S, Huang .J.-B, Ahuja. N, and Yang.M.-H. IEEE Conf. on Computer Vision and Pattern Recognition, 2017, pp. 624–632.
6. Galushkin A. I. Neyronnyeseti: osnovy teorii [Neural networks: fundamentals of theory]. M.: RiS, 2014. 496 p.
7. Gonsales R., Vuds R. Tsifrovaya obrabotka izobrazheniy [Digital image processing]. M.: Tekhnosfera, 2005. 1066 p.
8. Rostovtsev V.S. Iskusstvennyye neyronnyye seti [Artificial neural networks]. Kirov: VyatGU, 2014. 208 p
9. Donoho D.L., Johnstone I.M., Ideal spatial adaptation by wavelet shrinkage. Biometrika, 1994. V.83. pp. 425-455
10. Guo T., Mousavi H. S., Vu T.H., Monga V. Conf. on Computer Vision and Pattern Recognition Workshops, 2017. pp. 1100–1109.
11. Dong C., Loy C.C., Tang X. ECCV, 2016. pp. 1-17.
12. Rao K.R., Yip P. Discrete Cosine Transform: Algorithms, Advantages, Applications in Academic Press Professional, Inc., USA, 1990. 490 p.



13. Sun L., Ma C., Chen Y., Zheng Y., Shim H.J., Wu Z., Jeon B. IEEE Transactions on Circuits and Systems for Video Technology, 2020.V. 30.№ 10. pp. 3829–3842.
14. Wang Y., Wang L., Wang H., Li P. IEEE Transactions on Circuits and Systems for Video Technology, 2018.V. 29.№. 5. pp. 1259–1269.
15. Ledig C., Theis L., Huzs' F., Caballero J., Aitken A. P., Tejani A., Totz J., Wang Z., Shi W. IEEE Conference on Computer Vision and Pattern Recognition, 2018. pp. 105–114.
16. Sheikh, H.R., Bovik, A.C., De Veciana, G.: IEEE Trans image processing, 2005. V 14. №12. pp 2117–2128.

Дата поступления: 21.05.2024

Дата публикации: 3.07.2024