

Сравнение зависимости эффективности нейронных сетей по улучшению разрешения изображений от формата и размера

А.А. Журавлев

Уральский федеральный университет

Аннотация: Дороги имеют огромное влияние на жизнь современного человека. Одной из ключевых характеристик дорожного полотна является его качество. Для оценки качества дорожного покрытия существует множество систем. Такие технологии работают лучше с изображениями высокого разрешения (ИВР), поскольку на них проще выделять какие-либо признаки. Для улучшения разрешения фотографий существует достаточное количество способов, среди которых числятся и нейронные сети. Однако каждая нейросеть обладает определенными характеристиками. Например, для некоторых нейронных сетей достаточно проблематично работать с фотографиями большого исходного размера. Чтобы понять, насколько эффективна конкретная нейронная сеть, необходим сравнительный анализ. В данном исследовании за основной показатель эффективности берется среднее время получения ИВР. В качестве нейронных сетей выбраны EDSR, ESPCN, ESRGAN, FSRCNN и LapSRN, каждая из которых увеличивает ширину и высоту изображения в 4 раза (количество пикселей возрастает в 16 раз). Исходным материалом являются 5 фотографий 5 различных размеров (141x141, 200x200, 245x245, 283x283, 316x316) форматов png, jpg и bmp. Наилучшие показатели эффективности, согласно предложенной методологии, демонстрирует ESPCN, нейросеть FSRCNN также обладает хорошими результатами. Поэтому они являются более предпочтительными для решения задачи по улучшению разрешения изображений.

Ключевые слова: сравнение, зависимость, эффективность, нейронная сеть, нейросеть, улучшение разрешения, изображение, фотография, формат, размер, дорожная поверхность.

Введение

Оценка качества дорожного покрытия является одной из самых популярных задач во всем мире. Данной области посвящено множество монографий [1, 2] и научных работ, которые можно разделить на две основные категории: с использованием машинного обучения [3, 4] и традиционные методы (без использования машинного обучения) [5]. В настоящее время наиболее актуальной является первая группа в связи с большой популярностью нейронных сетей.

Системы, оценивающие качество дорожного покрытия, как правило, эффективнее работают с изображениями высокого разрешения, поскольку на них проще выделять какие-либо признаки. Однако на практике не всегда

получается взаимодействовать с такими фотографиями. В настоящее время существует достаточное количество нейронных сетей, предназначенных для получения изображения высокого разрешения (ИВР) из исходной фотографии. К ним относятся: EDSR [6], ESPCN [7], ESRGAN [8], FSRCNN [9], LapSRN [10]. Каждая из нейросетей обладает своими характеристиками. Одним из ключевых показателей качества нейросети является время получения ИВР. Также важна скорость роста функции зависимости времени от размера фотографии. Чем меньше оба показателя (абсолютный показатель времени и скорость роста функции зависимости функции), тем эффективнее нейронная сеть. Кроме того, фотографии могут иметь разный формат (png, jpg, bmp), что, скорее всего, также оказывает влияние на эффективность нейронной сети.

Цель работы – сравнить эффективность нейронных сетей по улучшению разрешения изображений разных форматов.

Материал исследования – фотографии дорожного покрытия форматов png, jpg и bmp.

В исследовании используется эмпирический метод, поскольку основными источниками результатов являются сравнение и эксперимент.

Научная новизна заключается в способе определения эффективности нейронных сетей по улучшению разрешения изображений.

Задачи исследования:

1. Описать методологию проведения эксперимента для определения зависимости эффективности нейронных сетей по улучшению изображений от формата и размера.
 2. Дать краткую информацию о выбранных нейросетях.
 3. Провести сравнительный анализ нейронных сетей на основе предложенной методологии.
 4. Описать полученные результаты.
-

Описание анализируемых нейронных сетей

В качестве нейронных сетей по улучшению разрешения изображений выбраны EDSR, ESPCN, ESRGAN, FSRCNN и LapSRN, увеличивающие размер фотографии в 4 раза по высоте и ширине одновременно (количество пикселей увеличивается в 16 раз). Краткая информация о каждой нейросети представлена ниже.

EDSR (Enhanced Deep Residual Network for Single Image Super-Resolution) – усовершенствованная глубокая остаточная сеть для сверхразрешения одного изображения [6].

ESPCN (Efficient Sub-Pixel Convolutional Neural Network) – эффективная субпиксельная сверточная нейронная сеть [7].

ESRGAN (Enhanced Super-Resolution Generative Adversarial Network) – усовершенствованная генеративно-сопоставительная сеть сверхвысокого разрешения [8].

FSRCNN (Fast Super-Resolution Convolutional Neural Network) – быстрая сверточная нейронная сеть со сверхразрешением [9].

LapSRN (Laplacian Pyramid Super-Resolution Network) – сеть сверхразрешения пирамиды Лапласа [10].

Методология эксперимента

Для эксперимента используются 5 фотографий 5 различных размеров (141x141, 200x200, 245x245, 283x283, 316x316) форматов png, jpg и bmp. Размеры изображений подобраны таким образом, чтобы количество пикселей в фотографии изменялось равномерно. Если перевести размеры в пиксели, то получаем: 20000 ($\approx 141 \times 141$), 40000 (200×200), 60000 ($\approx 245 \times 245$), 80000 ($\approx 283 \times 283$), 100000 ($\approx 316 \times 316$). Шаг изменения равен 20000 пикселей.

Ключевым показателем эффективности нейросетей выбранного типа является время получения изображения высокого разрешения (ИВР). Также большое влияние на эффективность оказывает размер данных, с которым

работает нейросеть. В нашем случае размером данных является разрешение фотографии.

Компьютер, на котором проводится исследование, имеет следующие характеристики: операционная система – Windows 10 Pro, тип системы – x64, процессор – Intel Core i5, оперативная память – 8 гигабайт, частота процессора – 1,6 гигагерц. В качестве среды выполнения выбрана Visual Studio Code 2023 язык программирования – Python (версия 3.11.7), тип приложения – Jupyter-блокнот.

Описание эксперимента для определения зависимости среднего времени получения ИВР нейросетью от формата и размера в пикселях:

1. В качестве исходного материала используются 5 фотографий 5 различных размеров форматов png, jpg и bmp.

2. Для каждой нейронной сети и формата изображения проводится серия из 5 экспериментов, в которых последовательно увеличивается количество пикселей, содержащихся в фотографии, на 20000: 20000 ($\approx 141 \times 141$), 40000 (200×200), 60000 ($\approx 245 \times 245$), 80000 ($\approx 283 \times 283$), 100000 ($\approx 316 \times 316$). Количество пикселей у ИВР увеличивается в 4 раза по высоте и ширине одновременно (количество пикселей изменяется в 16 раз). При этом 5 фотографий для каждого эксперимента хранятся в отдельной компьютерной папке (всего 75 различных папок: 5 (нейросетей) \times 3 (формата) \times 5 (количество экспериментов)). Изображения высокого разрешения сохраняются в новую папку, расположенную в том же месте, что и исходные фотографии.

3. Для определения среднего времени получения изображения высокого разрешения суммарное время, которое затрачивается на обработку всех фотографий, делится на общее количество изображений (1).

$$t_{\text{ср}} = \frac{t_{\text{общ}}}{n} \quad (1)$$

где t_{cp} – среднее время получения ИВР, $t_{общ}$ – общее время обработки всех фотографий, n – количество изображений (в нашем случае 5).

Результаты экспериментов

Исходные изображения и соответствующие ИВР (после проведения эксперимента) на примере формата jpg представлены на рис. 1.



Рис. 1. – Исходные фотографии формата jpg (слева) и соответствующие ИВР (справа)

Результаты эксперимента для определения эффективности нейронных сетей по улучшению разрешения изображений от формата и размера в пикселях представлены в таблице 1.

Таблица № 1

Среднее время (в миллисекундах) получения изображения высокого разрешения в зависимости от формата и исходного размера фотографии

PNG					
Нейронная сеть	Размер изображения				
	20000	40000	60000	80000	100000
EDSR	9015	20535	31414	43481	56139
ESPCN	300	426	542	667	803
ESRGAN	2516	4326	6625	9015	11974
FSRCNN	218	373	598	774	1022
LapSRN	954	2009	3119	4188	5290
JPG					
Нейронная сеть	Размер изображения				
	20000	40000	60000	80000	100000
EDSR	9105	21561	34429	47134	60525
ESPCN	48	62	80	98	128
ESRGAN	1616	3900	5511	7545	9118
FSRCNN	32	74	124	176	226
LapSRN	842	1715	2733	3715	4729
BMP					
Нейронная сеть	Размер изображения				
	20000	40000	60000	80000	100000
EDSR	10019	21042	33504	46165	60283
ESPCN	28	51	76	108	155
ESRGAN	3046	5040	6634	8235	10181
FSRCNN	43	98	140	185	227
LapSRN	935	1734	2624	3597	4750

Для того, чтобы однозначно понять, какой характер зависимости имеет конкретная нейронная сеть от формата и размера, необходимо построить соответствующие графики (рис. 2).

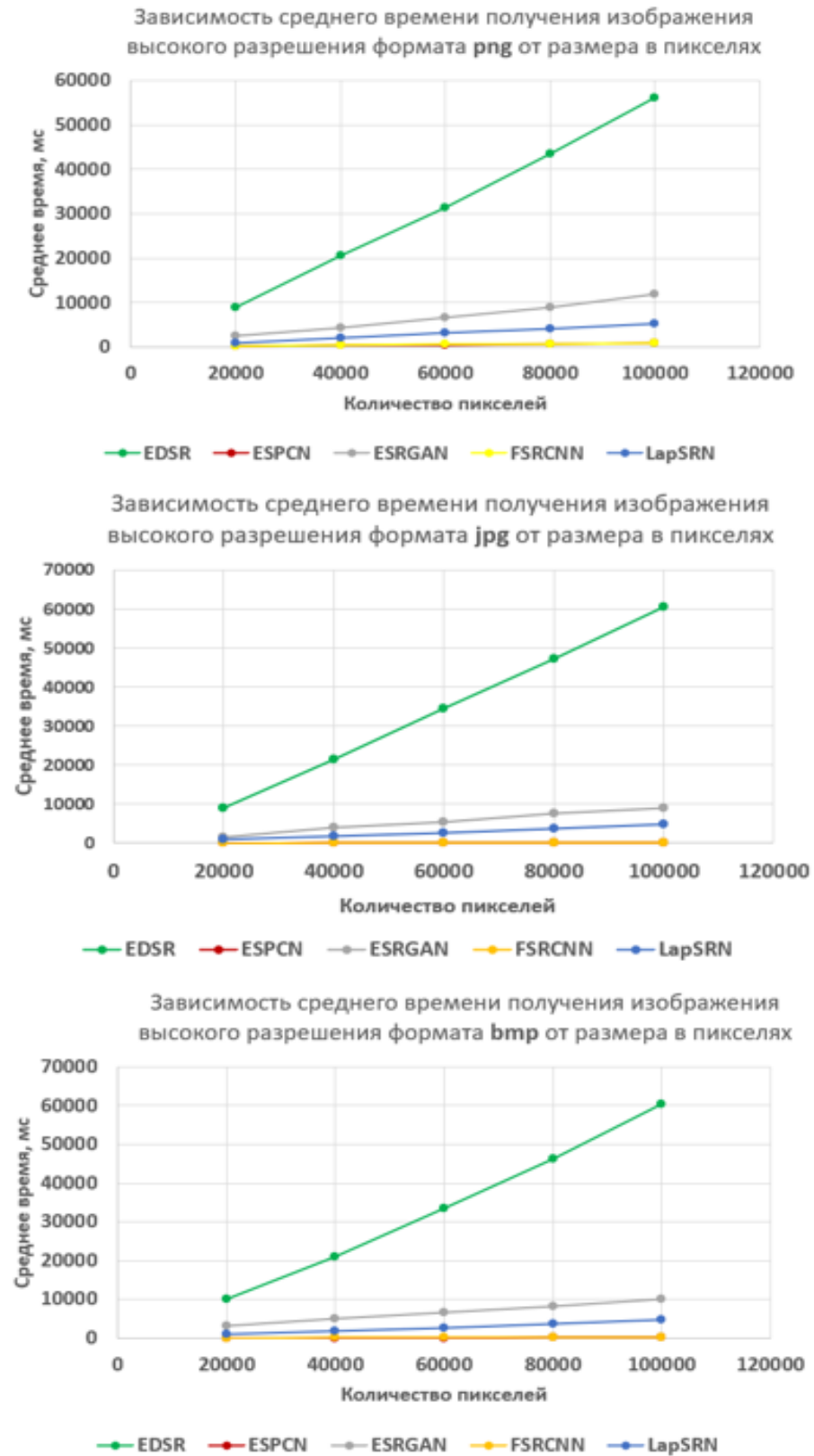


Рис. 2. – Графики зависимости среднего времени получения изображения высокого разрешения от формата и размера в пикселях

Обсуждение результатов

Как видно из графиков, каждая нейронная сеть для всех форматов имеет линейную зависимость $O(n)$ среднего времени получения ИВР от разрешения. Однако, абсолютные показатели времени и скорость роста функций отличаются. Чтобы иметь четкую картину эффективности конкретной нейросети в выбранной системе из пяти элементов (EDSR, ESPCN, ESRGAN, FSRCNN, LapSRN), рассчитаем эффективность, зависящую от скорости изменения функции таким образом, чтобы сумма всех показателей (эффективности) давала 1. Необходимы следующие формулы:

$$v_{\text{cp}} = \frac{\Delta t}{\Delta x} = \frac{t_{\text{max}} - t_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \quad (2)$$

где v_{cp} – средняя скорость изменения функции, Δt – разница максимального (t_{max}) и минимального (t_{min}) абсолютного значения времени, Δx – разница максимального (x_{max}) и минимального (x_{min}) значения размера изображения в пикселях.

$$k_{\text{норм}} * \left(\sum_{i=1}^N \frac{1}{v_{\text{cp} i}} \right) = k_{\text{норм}} * \left(\sum_{i=1}^N \frac{\Delta x_i}{\Delta t_i} \right) = 1 \quad (3)$$

где $k_{\text{норм}}$ – коэффициент нормализации, $v_{\text{cp} i}$ – скорость изменения функции конкретной нейросети, N – количество нейронных сетей (в нашем случае 5).

$$E = \frac{k_{\text{норм}}}{v_{\text{cp}}} = \frac{k_{\text{норм}} * \Delta x}{\Delta t} \quad (4)$$

где E – эффективность конкретной нейросети.

$$E_{\text{cp}} = \frac{\sum_{i=1}^N E_i}{M} \quad (5)$$

где E_{cp} – средняя эффективность нейросети, E_i – эффективность нейронной сети для определенного формата, M – количество форматов (равно 3).

Подставляя нужные значения в формулы (2), (3), (4) и (5), получаем результаты, представленные в таблице 2 (значения округлены до 3 знаков после запятой).

Таблица № 2

Показатели эффективности выбранных нейросетей по улучшению разрешения изображений для форматов png, jpg и bmp

Формат	Нейронная сеть				
	EDSR	ESPCN	ESRGAN	FSRCNN	LapSRN
PNG	0,006	0,554	0,029	0,347	0,064
JPG	0,001	0,692	0,008	0,285	0,014
BMP	0,002	0,573	0,010	0,396	0,019
Среднее значение	0,003	0,606	0,016	0,343	0,032

Наилучший средний показатель эффективности имеет ESPCN, также высоким значением обладает FSRCNN. Следовательно, ESPCN и FSRCNN являются наиболее подходящими для решения задачи по улучшению разрешения изображений.

Выводы

В работе проведено сравнение зависимости эффективности нейронных сетей по улучшению разрешения от формата и размера. Исходный материал - 5 фотографий пяти различных размеров (141x141, 200x200, 245x245, 283x283, 316x316) форматов png, jpg и bmp. В качестве нейросетей для анализа выбраны EDSR, ESPCN, ESRGAN, FSRCNN и LapSRN, увеличивающие размер фотографии в 4 раза по ширине и высоте, одновременно (количество пикселей увеличивается в 16 раз). Каждая из нейросетей обладает линейной зависимостью $O(n)$ времени получения изображения высокого разрешения (ИВР) от размера фотографии для всех форматов. Однако, наилучшими показателями среднего времени и скорости

изменения функции обладает ESPCN, хороший результат также имеет FSRCNN.

Литература

1. Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В., Ососков М.В., Моржин А.В. Обработка и анализ изображений в задачах машинного зрения. - М.: ФИЗМАТКНИГА, 2010. - 672 с.
 2. Гонсалес Р. Цифровая обработка изображений. - М.: Техносфера, 2012. - 1101 с.
 3. Arya D., Maeda H., Ghosh S. K., Toshniwal D., Mraz A., Kashiya T., Sekimoto Y. Deep learning-based road damage detection and classification for multiple countries // Automation in Construction. 2021, Vol. 132, P. 103935.
 4. Silva L.A., Leithardt V.R.Q., Batista V.F.L., González G.V., Santana J.F.D.P. Automated Road Damage Detection Using UAV Images and Deep Learning Techniques // IEEE Access. 2023, Vol. 11, pp. 62918–62931.
 5. Zhuravlev A.A., Aksyonov K.A. Comparison of Contour Detection Methods in Images on the Example of Photos with Road Surface Damage // Institute of Electrical and Electronics Engineers Inc., 2023 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT). 2023, pp. 183-186.
 6. Enhanced Deep Residual Networks for Single Image Super-Resolution (EDSR). URL: huggingface.co/eugenesiow/edsr#:~:text=EDSR%20upscaling%20x2.-,Model%20description,the%20end%20of%20the%20network.
 7. ESPCN. URL: learnopencv.com/super-resolution-in-opencv/#sec4
 8. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. URL: esrgan.readthedocs.io/en/latest/
 9. Accelerating the Super-Resolution Convolutional Neural Network. URL: mmlab.ie.cuhk.edu.hk/projects/FSRCNN.html
 10. LapSRN. URL: learnopencv.com/super-resolution-in-opencv/#sec6
-

References

1. Vizil'ter Yu.V., Zheltov S.Yu., Bondarenko A.V., Ososkov M.B., Morzhin A.V. Obrabotka i analiz izobrazhenij v zadachax mashinnogo zreniya [Image processing and analysis in machine vision problems]. M.: FIZMATKNIGA, 2010. 672 p.
2. Gonsales R. Cifrovaya obrabotka izobrazhenij [Digital image processing]. Texnosfera. 2012. 1101 p.
3. Arya D., Maeda H., Ghosh S. K., Toshniwal D., Mraz A., Kashiyama T., Sekimoto Y. Automation in Construction. 2021, Vol. 132, 103935.
4. Silva L.A., Leithardt V.R.Q., Batista V.F.L., González G.V., Santana J.F.D.P. IEEE Access. 2023, Vol. 11, pp. 62918–62931.
5. Zhuravlev A.A., Aksyonov K.A. Institute of Electrical and Electronics Engineers Inc., 2023 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT). 2023, pp. 183-186.
6. Enhanced Deep Residual Networks for Single Image Super-Resolution (EDSR). URL: huggingface.co/eugenesiow/edsr#:~:text=EDSR%20upscaling%20x2.-,Model%20description,the%20end%20of%20the%20network.
7. ESPCN. URL: learnopencv.com/super-resolution-in-opencv/#sec4.
8. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. URL: esrgan.readthedocs.io/en/latest/.
9. Accelerating the Super-Resolution Convolutional Neural Network. URL: mmlab.ie.cuhk.edu.hk/projects/FSRCNN.html.
10. LapSRN. URL: learnopencv.com/super-resolution-in-opencv/#sec6.

Дата поступления: 28.02.2024

Дата публикации: 1.04.2024