

Аверин Д.С. Исследование работы модели нейронной сети SRGAN для увеличения разрешающей способности изображения // Академия педагогических идей «Новация». Серия: Студенческий научный вестник. – 2019. – №6 (июнь). – АРТ 514-эл. – 0,2 п.л. - URL: <http://akademnova.ru/page/875550>

РУБРИКА: ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 004.8

Аверин Дмитрий Сергеевич

магистрант 2 курса, факультет «ИВТ»

Научный руководитель: Марков Е.М., к.т.н., доцент

ФГБОУ ВО «ИжГТУ им. М.Т. Калашникова»

г. Ижевск, Российская Федерация

e-mail: twosoul123@gmail.com

**ИССЛЕДОВАНИЕ РАБОТЫ МОДЕЛИ НЕЙРОННОЙ СЕТИ SRGAN
ДЛЯ УВЕЛИЧЕНИЯ РАЗРЕШАЮЩЕЙ СПОСОБНОСТИ
ИЗОБРАЖЕНИЯ**

Аннотация: В данной статье рассматривается генеративно-сопоставительная модель нейронной сети SRGAN для увеличения разрешения изображения. Производится обучение данной модели и сравнение результатов ее работы с одним из методов интерполяции.

Ключевые слова: суперразрешение, генеративно-сопоставительная нейронная сеть, машинное зрение.

Averin Dmitriy Sergeevich

2nd year master student, Faculty of Informatics and Computer Engineering

Supervisor: E. M. Markov, PhD, Associate Professor

FGBOU VO Kalashnikov Izhevsk State Technical University

Izhevsk, Russian Federation

RESEARCH OF NEURAL NETWORK SRGAN TO INCREASE IMAGE RESOLUTION

Abstract: This article discusses the generative-adversarial model of the neural network SRGAN to increase the image resolution. Training of this model and comparison of its results with one of the interpolation methods are carried out.

Key words: super resolution, generative-adversarial neural network, machine vision.

Генеративно-сопоставительные сети — это модели нейронных сетей (НС) с обучением без учителя, построенные на комбинации из двух нейронных сетей, как правило сверточных. Первая сеть (генератор) генерирует образцы. Вторая сеть (дискриминатор) принимает на вход результаты работы первой сети и правильные («подлинные», «оригинальные») данные, а затем пытается отличить оригинальные образцы от сгенерированных. [1].

В данной модели необходимо обучать сразу две НС и необходимо настроить баланс между ними. Генератор обучается создавать близкие к оригиналу данные, чтобы дискриминатор не мог их отличить, а дискриминатор обучается распознавать реальные данные от

сгенерированных. Если одна из сетей будет обучена лучше, чем другая, то модель в целом не будет работать корректно.

Генеративно-сопоставительной сети (Generative Adversarial Network, GAN) для повышения разрешения изображений (image super resolution). Входными данными служат изображения. Обозначим изображение, которое мы получаем в результате работы нейронной сети, как I^{SR} , изображение низкого разрешения, которое необходимо увеличить в n раз, как I^{LR} . Изображение высокого разрешения (без обработки) обозначим как I^{HR} . I^{LR} получается в результате уменьшения I^{HR} в n раз.

В результате обучения необходимо получить генерирующую функцию G . Генератор обучается как сверточная нейронная сеть прямого распространения с параметрами Θ_G :

$$\theta_G = \{ W_{1:L}; b_{1:L} \}, \quad (1)$$

где $W_{1:L}$ и $b_{1:L}$ — это, соответственно, веса и смещения нейронной сети глубиной в L слоев. Этот параметр вычисляется путем оптимизации составной функции потерь l_{SR} . Для обучения нейронной сети находим значение:

$$\theta_G = \min \frac{1}{N} l_{SR}(G(I_n^{LR}), I_n^{HR}) \quad (2)$$

Функция потерь l_{SR} состоит из следующих компонентов:

$$l_{SR} = l_{MSE} + 10^{-3} l_{GEN}, \quad (3)$$

где l_{MSE} — средняя квадратическая ошибка при сравнении сгенерированного изображения I^{SR} с его оригиналом I^{HR} ,

l_{GEN} — соревновательная функция потерь, которую можно представить следующей формулой:

$$I_{\text{Gen}} = \sum_{n=1}^N -\log(D(G(I^{\text{LR}}))), \quad (4)$$

где $D(G(I^{\text{LR}}))$ — это вероятность распознавания дискриминатором того, что созданное генератором изображение $G(I^{\text{LR}})$ является изображением высокого разрешения I^{HR} .

Стоит отметить, что авторы исследования предлагают использовать при сравнении изображений не MSE, а предобученную нейронную сеть VGG-19, т.к. результаты ее работы больше соответствуют тому, как человеческий глаз воспринимает разницу между двумя изображениями [2].

Авторы данной модели применяют нормализацию к входным данным нейронной сети, т.к. это позволяет увеличить эффективность работы алгоритма. Подробно о преимуществах нормализации рассказано в статье [3].

При обучении используется оптимизатор Adam (Adaptive Moment Estimation) [4].

Архитектура генеративной нейронной сети (G) состоит из следующих блоков: 2 сверточных слоя с ядром 3×3 , 64 карты признаков, за которыми следуют слои нормализации и два субпиксельных сверточных слоя.

Для ускорения работы нейронной сети используется метод батчнормализации (Batch Normalization), предложенный сотрудниками компании Google в 2015 году. [14] Иногда встречается перевод этого термина «пакетная нормализация». Этот метод предполагает добавление дополнительного слоя, который позволяет снизить искажение данных путем их нормализации. Этот метод позволяет сократить время обучения за счет использования более высокого значения параметра скорости обучения (learning rate), а также снизить переобучение.

Структура ИНС – Generator Network показана на рисунке 1.

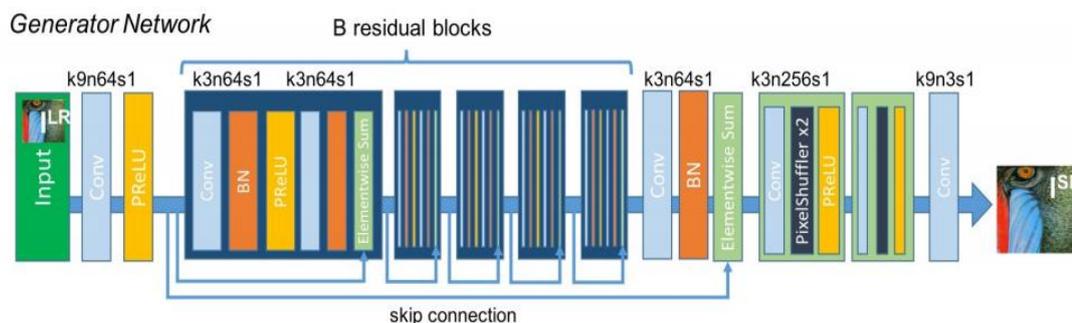


Рисунок 1 - Структура Generator Network

Input I^{LR} – изображение низкого разрешения, которое подается на вход генератору; Conv – сверточный слой с n картами признаков (feature maps) и шагом (stride), равным s ; ReLU – активационная функция Parametric ReLU; BN – Batch Normalization, батч-нормализация; skip connection – дополнительные связи; B residual blocks – остаточные блоки; Структура ИНС – Discriminator Network показана на рисунке 2.

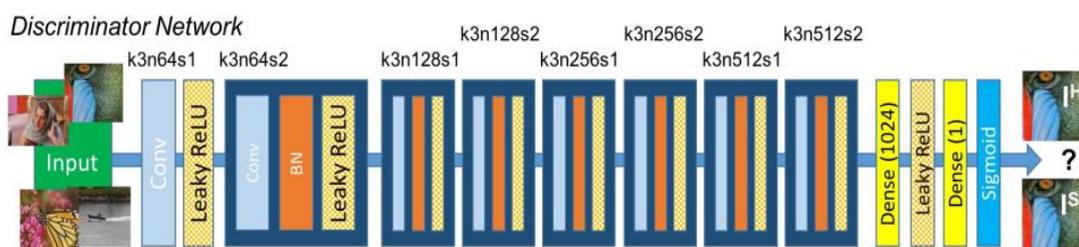


Рисунок 2 - Структура Discriminator Network

Input – входные данные, изображения высокого разрешения и изображения, созданные генератором; Conv – сверточный слой с n картами признаков (feature maps) и шагом (stride), равным s ; Leaky ReLU – активационная функция Leaky ReLU; BN – Batch Normalization, батч-

нормализация; Dense – полносвязный слой; Sigmoid – активационная функция сигмоида;

В данной статье была обучена SRGAN в течении 300000 итераций. Для сравнения результатов с бикубической интерполяцией использовались метрики PSNR и SSIM [5] на наборе данных Set14, состоящего из 14 изображений. Результаты показаны в таблице 1.

Таблица 1 – Сравнение полученных результатов на наборе данных Set14.

Метрика	Бикубическая интерполяция	SRGAN
PSNR	24,538	26,842
SSIM	0.680	0,750

На рисунке 3 представлен пример работы SRGAN и бикубической интерполяции.

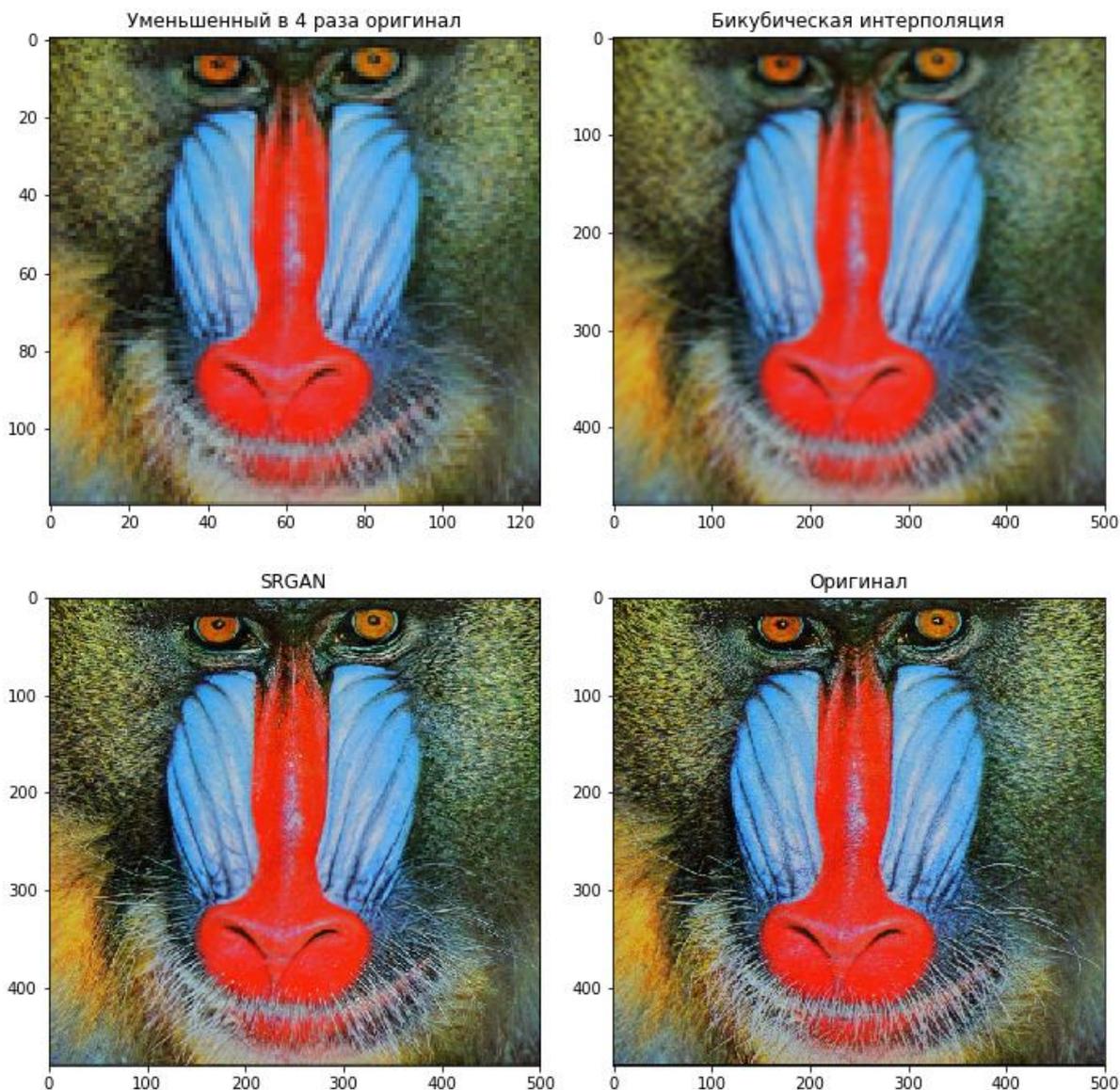


Рисунок 3 – Пример работы SRGAN и бикубической интерполяции

Заключение. В данной статье была рассмотрена и обучена генеративно-состязательная модель нейронной сети SRGAN. Метрики PSNR и SSIM показывают, что данная сеть работает лучше бикубической интерполяции. Так же визуально видно, что результат, полученный

нейронной сетью, гораздо ближе к оригиналу, его резкость выше, в отличие от результата бикубической интерполяции.

Список использованной литературы:

1. Дорогой Я. Ю. Архитектура обобщенных сверточных нейронных сетей // Вестник Национального технического университета Украины. — 2011. - №. 54.
2. Ledig C., Theis L., Huszar F., Caballero J., Cunningham A., Acosta A., Aitken A., Tejani A., Totz J, Wang Z., Shi W., Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, CVPR, 2016.
3. Ioffe, Sergey, and Christian Szegedy. «Batch normalization: Accelerating deep network training by reducing internal covariate shift.» arXiv preprint arXiv:1502.03167 (2015).
4. Diederik P. Kingma, Jimmy Lei Ba, Nadav Cohen. Adam: A Method for Stochastic Optimization. The Hebrew University of Jerusalem. Advanced Seminar in Deep Learning (#67679). October 18, 2015.
5. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4):600–612, 2004

Дата поступления в редакцию: 18.06.2019 г.

Опубликовано: 25.06.2019 г.

© Академия педагогических идей «Новация». Серия «Студенческий научный вестник», электронный журнал, 2019

© Аверин Д.С., 2019