

Адаптивный нейросетевой метод построения интерполяционной формулы для удвоения размера изображения

С.Е. Ваганов¹

¹ Ивановский государственный университет, Иваново, Россия

Аннотация

Предложена архитектура искусственной нейронной сети, решающей задачу построения интерполяционных формул для удвоения размера изображений. Обученная модель получает в качестве аргумента матрицу размера 4×4 , а результатом работы является интерполяционная формула, представленная в виде весового вектора для 4 точек.

Произведено сравнение основных оценок качества предложенного метода с некоторыми известными адаптивными подходами. Сравнительный анализ показал, что предложенный подход имеет лучшее качество интерполяции по сравнению с методами New Edge-Directed Interpolation и Directional Cubic Convolution Interpolation.

Ключевые слова: интерполяция, машинное обучение, искусственная нейронная сеть, градиентный спуск, качество изображения.

Цитирование: Ваганов, С.Е. Адаптивный нейросетевой метод построения интерполяционной формулы для удвоения размера изображения / С.Е. Ваганов // Компьютерная оптика. – 2019. – Т. 43, № 4. – С. 627-631. – DOI: 10.18287/2412-6179-2019-43-4-627-631.

Введение

Все алгоритмы масштабирования изображений можно разбить на два класса: адаптивные и неадаптивные.

Наиболее широко используемыми среди неадаптивных методов являются билинейная и бикубическая интерполяция [1]. Эти методы популярны в силу простоты реализации и эффективности, но имеют ряд недостатков с точки зрения качества получаемого изображения, в том числе эффект Гиббса, размытие и лестничный эффект. Данные дефекты возникают в силу недостаточно точной аппроксимации в точках, расположенных рядом или непосредственно на границах восстанавливаемого изображения.

Целью адаптивных методов является борьба с данными дефектами. Обычно суть адаптации состоит в выборе оптимальной интерполяционной формулы [2, 3] или в её построении [4] для каждой интерполируемой точки изображения.

В алгоритме Directional Cubic Convolution Interpolation (DCCI) [2] обученный классификатор по значению градиента выбирает одну из трёх интерполяционных формул, а именно: кубическую свёртку с одной из двух диагоналей (при наличии в интерполируемом окне граничных точек) либо взвешенную композицию пары кубических свёрток. Как правило, обобщения метода DCCI представляют собой его расширения на большее число интерполяционных формул [3].

Метод New Edge-Directed Interpolation (NEDI) [4] осуществляет построение интерполяционных формул, представленных в виде набора весовых коэффициентов. Интерполяция вычисляется по 4 точкам. Весовые коэффициенты в NEDI находятся посредством решения СЛАУ, содержащей K^2 уравнений от 4 неизвестных для прямоугольного локального окна размерами $K \times K$.

На сегодняшний день использование методов машинного обучения стало трендом во многих обла-

стях, в частности: в цифровой обработке изображений и компьютерном зрении. Стремительный рост числа исследований в данной области отражает актуальность применения нейросетевых подходов для решения различных прикладных задач, в том числе интерполяции изображений [5–7]. В частности, в настоящей работе показано, что использование нейросетевого подхода позволяет заметно улучшить качество интерполяции по сравнению с классическими адаптивными алгоритмами [2, 4].

В настоящей работе используется следующая схема [2–4] построения удвоенного изображения:

- 1) инициализация значений в чётных точках изображения высокого разрешения точками исходного изображения (чёрные круги на рис. 1);
- 2) интерполяция в нечётных точках (белые круги на рис. 1), по чётным из некоторой окрестности. Например, в DCCI значение в центральном круге (на рис. 1) интерполируется по 16 чёрным точкам;
- 3) интерполяция в оставшихся точках (серые круги на рис. 1) по значениям, полученным в пунктах 1–2. Данный шаг выполняется аналогично второму с учётом поворота интерполяционного окна на 45 градусов.

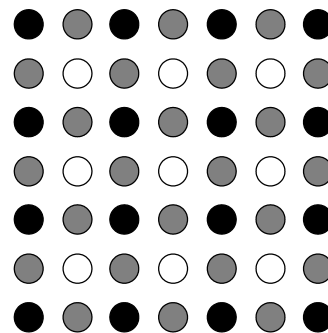


Рис. 1. Фрагмент сетки пикселей восстанавливаемого изображения

Посредством использования классической архитектуры регрессионной модели (выходной нейрон

которой возвращает результат аппроксимации) не удалось получить результат лучший, чем *NEDI* и *DCCI*. Поэтому в работе рассмотрен иной подход.

Цель работы заключалась в построении нейросетевой адаптивной интерполяционной формулы. На вход сети подаётся матрица размера 4×4 (чёрные круги на рис. 1), а результатом является набор из четырёх весовых коэффициентов. Аппроксимация строится посредством линейной комбинации полученных весов со значениями интенсивностей в ближайших четырёх чёрных точках к интерполируемой (центральной).

В параграфе 1 представлено описание архитектуры и методов обучения нейронной сети, решающей задачу построения интерполяционной формулы. В параграфе 2 приведён сравнительный анализ качества интерполяции полученной модели с некоторыми известными адаптивными методами.

1. Архитектура нейронной сети

На вход модели поступает матрица размера 4×4 (значения интенсивностей в чётных точках изображения). Результатом работы сети является вектор весовых коэффициентов $V = (v_1, v_2, v_3, v_4)$ для 4 ближайших точек относительно интерполируемой. Аппроксимация значения в точке (x, y) вычисляется посредством линейной комбинации следующего вида:

$$I(x, y) = \sum_{i=1}^4 v_i \cdot f_i, \quad (1)$$

где (f_1, f_2, f_3, f_4) – значения интенсивностей изображения в точках $(x-1, y-1)$, $(x+1, y-1)$, $(x-1, y+1)$ и $(x+1, y+1)$ соответственно.

Для решения задачи построения интерполяционной формулы экспериментальным путём была подобрана архитектура нейронной сети, включающая следующие три слоя:

1) Свёрточный слой $4 \times 4 \times N$ без смещения (эквивалентен полносвязному слою из N нейронов). Результат применения свёрток к входным матрицам 4×4 – вектор размерности N . Активационная функция – гиперболический тангенс:

$$th(t) = \frac{e^t - e^{-t}}{e^t + e^{-t}}. \quad (2)$$

2) Полносвязный слой из N -нейронов со смещением и активационной функцией *RELU*:

$$RELU(t) = \max(0, t). \quad (3)$$

3) Полносвязный выходной слой из 4 нейронов со смещением и нормализацией:

$$Norm(z_i) = z_i / \sum_{j=1}^4 z_j, \quad (4)$$

где z_i – результат работы i -го нейрона 3-го слоя.

В результате применения (4) будет получен вектор весовых коэффициентов V . Интерполяция вычисляется по формуле (1).

Экспериментально было установлено, что для фиксированной обучающей выборки, построенной по трём изображениям (рис. 2), изменение глубины сети (добавление полносвязных скрытых слоёв) негативно воздействует на качество интерполяции. В частности, было опробовано добавление скрытого слоя с активационной функцией *RELU* между 2-м и 3-м слоем, а также удаление 2-го слоя.



Рис. 2. Исходные изображения

Число параметров модели можно выразить следующим соотношением:

$$P(N) = N^2 + 21N + 4. \quad (5)$$

Далее приведена таблица, позволяющая оценить зависимость качества интерполяции от числа нейронов N в каждом слое (табл. 1).

Табл. 1. Усреднённые оценки качества интерполяции построенных моделей на наборе тестовых изображений [8]

N	4	8	16	24	32	64
<i>RMSE</i>	9,656	9,494	9,347	9,275	9,255	9,233
<i>PSNR</i>	29,40	29,58	29,75	29,84	29,85	29,88
<i>SSIM</i>	0,978	0,978	0,979	0,979	0,979	0,979
$P(N)$	104	236	596	1 084	1 700	5 444

В среднем (по табл. 1) *RMSE* (корень среднеквадратичного отклонения) при увеличении числа нейронов в каждом слое N с 32 до 64 уменьшился не более чем на 0,025, при этом сложность модели выросла более чем в 3 раза.

Таким образом, в противоположность изменению глубины сети увеличение ширины слоёв позволяет улучшить качество интерполяции. Однако число параметров при этом становится настолько велико, что это не всегда можно назвать обоснованным. С увеличением количества параметров растёт вычислительная сложность модели.

На рис. 3 представлены нормированные ядра свёрток первого слоя обученной модели при $N=16$. Значения коэффициентов лежат в интервале $(-3, 1; 3, 1)$. Чем светлее пиксель, тем выше величина параметра.

Более подробный сравнительный анализ обученных моделей будет представлен в параграфе 2.

Обучение нейронной сети

В настоящей работе для построения обучающей выборки были использованы три стандартных изображения, среди которых пара фотореалистичных (*fruits* и *monarch*) и одно построенное синтетическим путём (*watch*). Элементы обучающей выборки были сформированы из матриц интенсивностей размера 4×4 и соответствующих значений в интерполируемых точках.

Изображения, использованные для построения обучающей выборки, представлены на следующей иллюстрации.

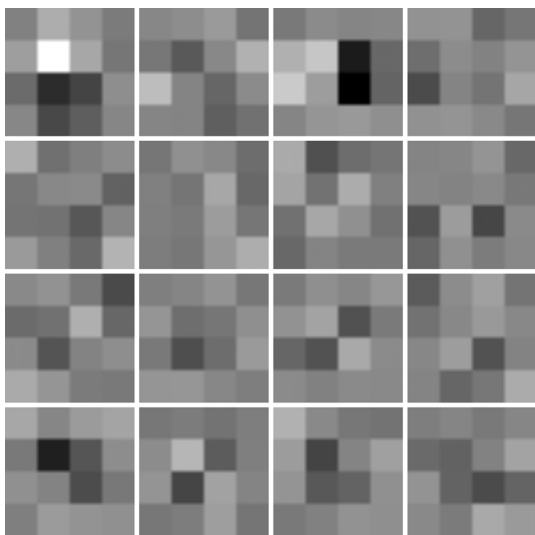


Рис. 3. Ядра свёрток первого слоя архитектуры $N=16$

Параметры моделей находились посредством минимизации функции потерь, которая представляла собой оценку средней абсолютной ошибки (*MAE*) истинных значений интенсивностей от аппроксимаций, найденных посредством использования полученной интерполяционной формулы:

$$Loss(Y, \hat{Y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (6)$$

где $Y, \hat{Y} \in R^n$ – вектора истинных значений из обучающей выборки и аппроксимаций, полученных посредством нейросетевой формулы.

Для решения задачи минимизации использовался метод *Adam* (одна из модификаций стохастического градиентного спуска) [9].

Процесс обучения осуществлялся посредством использования функционала библиотеки *TensorFlow* [10]. Обучение производилось по мини-батчам со следующими параметрами:

Размер мини-батча (<i>Batch size</i>)	100
Начальный шаг смещения в направлении антиградиента (<i>ILR, Initial Learning Rate</i>)	0,01
Количество эпох	32
Число элементов в обучающей выборке	1 417 324

Каждую эпоху параметр *LR* (*Learning Rate*) вычислялся как $LR = ILR / EN$, где *EN* соответствует номеру текущей эпохи.

В силу случайности выбора начального приближения, обучение каждой архитектуры было проведе-

но по 20 раз. Затем был осуществлён выбор наилучшей модели с точки зрения усреднённой величины *RMSE*, вычисленной для тестовой выборки.

2. Оценка качества интерполяционной формулы

Для оценки качества полученной интерполяционной формулы использовался набор стандартных фотореалистичных изображений [8]. Каждое изображение было уменьшено вдвое посредством прореживания (взятия значений в чётных точках).

Расчёт показателей качества для увеличенных изображений выполнялся только по нечётным точкам (белые круги на рис. 1) и с отступом в 10 пикселей от границ изображений. Оценка качества аппроксимации интенсивностей, получаемых на третьем шаге алгоритма, не использовалась ввиду накопления погрешности интерполяции со второго шага. Отступ необходим для сравнения с методом *NEDI* при различных размерах локального окна.

В следующих таблицах представлены значения оценок *PSNR* и *SSIM* исходных и интерполированных изображений, построенных различными адаптивными методами. Подходы *NEDI4*, *NEDI6* и *NEDI8* – интерполяция посредством *NEDI* с размером локального окна 4×4 , 6×6 и 8×8 соответственно. Алгоритм *EDIST* (*Edge-directional interpolation algorithm using structure tensor*) описан в работе [11]. Предложенный в работе метод обозначен как *NX*, где *X* – число нейронов в каждом слое рассматриваемой архитектуры (табл. 2, $N = 32$).

Табл. 2. Оценки *PSNR* для различных адаптивных методов [2, 4, 11]

Image	EDIST	DCCI	NEDI4	NEDI8	N32
Airplane	30,54	30,48	29,69	30,22	31,13
Car	27,35	27,17	26,53	27,27	27,52
Couple	29,37	29,52	28,75	29,23	29,75
Girl	31,23	31,38	30,74	31,40	32,06
House	28,41	28,75	27,97	28,60	29,34
Lake	25,91	25,68	25,38	25,89	26,03
Lena	30,47	30,70	30,29	30,76	31,07
Mandrill	19,99	20,06	19,93	20,73	20,26
Peppers	32,97	33,26	32,96	33,21	33,77
Splash	34,60	34,76	34,24	34,46	36,13
Tiffany	30,27	30,57	30,34	30,91	30,84
Tree	25,27	25,35	24,72	25,27	25,76
Woman	34,01	34,16	33,81	33,92	34,43
Mean	29,26	29,37	28,87	29,37	29,85

Табл. 3. Оценки *SSIM* для различных адаптивных методов [2, 4, 11]

Image	EDIST	DCCI	NEDI4	NEDI8	N32
Airplane	0,985	0,985	0,982	0,984	0,987
Car	0,979	0,979	0,975	0,979	0,980
Couple	0,962	0,963	0,956	0,960	0,965
Girl	0,991	0,991	0,990	0,991	0,993
House	0,982	0,984	0,980	0,983	0,986
Lake	0,982	0,981	0,980	0,982	0,983
Lena	0,992	0,992	0,991	0,992	0,993
Mandrill	0,895	0,897	0,894	0,910	0,901
Peppers	0,996	0,996	0,996	0,996	0,997
Splash	0,998	0,998	0,998	0,998	0,999
Tiffany	0,982	0,983	0,983	0,985	0,984
Tree	0,977	0,978	0,974	0,977	0,980
Woman	0,981	0,981	0,980	0,980	0,982
Mean	0,977	0,978	0,975	0,978	0,979

Жирным шрифтом отмечены лучшие показатели качества. Усреднённые оценки *PSNR* и *SSIM* метода *N32* имеют наилучшие результаты среди рассматриваемых подходов. Следовательно, полученная интерполяционная формула (в среднем) позволяет получить более точную аппроксимацию по сравнению с методами [2, 4, 11].

Далее представлена таблица усредненных оценок *RMSE*, *PSNR* и *SSIM* для рассматриваемых в работе подходов (табл. 4).

Табл. 4. Усреднённые оценки качества интерполяции для различных методов

Метод	<i>RMSE</i>	<i>PSNR</i>	<i>SSIM</i>	<i>P(N)</i>
<i>DCCI</i>	9,697	29,37	0,978	-
<i>EDIST</i>	9,786	29,26	0,977	-
<i>NEDI4</i>	10,230	28,87	0,975	-
<i>NEDI6</i>	9,738	29,24	0,978	-
<i>NEDI8</i>	9,579	29,37	0,978	-
<i>N4</i>	9,656	29,40	0,978	104
<i>N8</i>	9,494	29,58	0,978	236
<i>N16</i>	9,347	29,75	0,979	596
<i>N24</i>	9,275	29,84	0,979	1084
<i>N32</i>	9,255	29,85	0,979	1700
<i>N64</i>	9,233	29,88	0,979	5444

Согласно экспериментальным данным (табл. 4), предложенная нейросетевая модель (в среднем) позволяет получить лучшее качество среди рассмотренных методов уже при 8 нейронах в каждом слое. При этом число точек, используемых для построения интерполяционной формулы (16 точек), оказывается в несколько раз меньше, чем у *NEDI* с размером локального окна 8×8 (100 точек).

В табл. 2 и табл. 3 приведён метод *N32*, так как, по мнению автора, дальнейшее увеличение числа нейронов в предложенной модели не является обоснованным в силу её усложнения из-за роста числа параметров при незначительном улучшении качества (табл. 4).

Заключение

Предложенная в работе нейросетевая модель позволяет получить более точную аппроксимацию интенсивностей по сравнению с методами [2, 4, 11].

К сожалению, понять принцип работы алгоритма интерполяции по набору весовых параметров построенной нейронной сети автору не удалось. Тем не менее, предложенный метод демонстрирует возможность построения интерполяционной формулы по 16 точкам, дающей более высокую точность, чем методы *NEDI* и *DCCI*.

Переход к большему числу узлов, используемых для построения нейросетевой интерполяционной формулы, возможно, позволит увеличить точность аппроксимации, однако повлечёт за собой рост числа параметров и, как следствие, увеличение сложности модели.

Предложенный алгоритм может быть использован для решения задачи иерархического сжатия изображений [12].

Реализация метода удвоения доступна по адресу [13].

Литература

1. Ваганов, С.Е. Сравнение алгоритмов удвоения размера изображения / С.Е. Ваганов, С.И. Хашин // Моделирование и анализ информационных систем. – 2016. – Т. 23, № 4. – С. 389-400. – DOI: 10.18255/1818-1015-2016-4-389-400.
2. Zhou, D. Image zooming using directional cubic convolution interpolation / D. Zhou, X. Shen, W. Dong // IET Image Processing. – 2012. – Vol. 6, Issue 6. – P. 627-634. – DOI: 10.1049/iet-ipr.2011.0534.
3. Jing, L. Directional bicubic interpolation – a new method of image super-resolution / L. Jing, G. Zongliang, Z. Xiuchang. – In: Proceedings of 3rd International Conference on Multimedia Technology (ICMT-13). – Atlantis Press, 2013. – P. 470-477. – DOI: 10.2991/icmt-13.2013.57.
4. Li, X. New edge-directed interpolation / X. Li, M.T. Orchard // IEEE Transactions on Image Processing. – 2001. – Vol. 10, Issue 10. – P. 1521-1527. – DOI: 10.1109/83.951537.
5. Dong, C. Image super-resolution using deep convolutional networks / C. Dong, C.C. Loy, K. He, X. Tang // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2016. – Vol. 38, Issue 2. – P. 295-307. – DOI: 10.1109/TPAMI.2015.2439281.
6. Plaziac, N. Image interpolation using neural networks / N. Plaziac // IEEE Transactions on Image Processing. – 1999. – Vol. 8, Issue 11. – P. 1647-1651. – DOI: 10.1109/83.799893.
7. Hu, H. Image interpolation using classification-based neural networks / H. Hu, P.M. Holman, G. de Haan // IEEE International Symposium on Consumer Electronics. – 2004. – P. 133-137. – DOI: 10.1109/ISCE.2004.1375920.
8. Тестовые bmp-файлы [Электронный ресурс]. – URL: http://math.ivanovo.ac.ru/dalgebra/Khashin/bmp_ex/ (дата обращения 06.02.2019).
9. Kingma, D.P. Adam: A method for stochastic optimization [Electronical Resource] / D.P. Kingma, J. Ba // arXiv preprint arXiv:1412.6980. – 2017. – URL: <https://arxiv.org/abs/1412.6980> (request date 07.02.2019).
10. TensorFlow. Official site [Electronical Resource]. – URL: <http://www.tensorflow.org> (request date 07.02.2019).
11. Nasonov, A.V. Edge-directional interpolation algorithm using structure tensor / A.V. Nasonov, A.S. Krylov, X. Petrova, M.N. Rychagov. – In: Electronic Imaging, Image Processing: Algorithms and Systems XIV / ed. by S.S. Agaian, K.O. Egiazarian, A.P. Gotchev. – Ingenta, 2016. – P. 1-4. – DOI: 10.2352/ISSN.2470-1173.2016.15.IPAS-026.
12. Гашиников, М.В. Интерполяция на основе контекстного моделирования при иерархической компрессии многомерных сигналов / М.В. Гашиников // Компьютерная оптика. – 2018. – Т. 42, № 3. – С. 468-475. – DOI: 10.18287/2412-6179-2018-42-3-468-475.
13. Реализация метода удвоения размера изображения посредством нейросетевой адаптивной интерполяционной формулы [Электронный ресурс]. – URL: <http://math.ivanovo.ac.ru/dcompmath/Vaganov/Interp.html> (дата обращения 11.02.2019).

Сведения об авторе

Ваганов Сергей Евгеньевич, 1992 года рождения, в 2015 окончил магистратуру в Ивановском государственном университете по направлению «Математика и компьютерные науки». В настоящий момент (2019 год)

является аспирантом Ивановского государственного университета (ИвГУ) по профилю 01.06.01 «Математическая логика, алгебра и теория чисел», работает ведущим программистом факультета математики и компьютерных наук в ИвГУ. Область научных интересов: цифровая обработка изображений, распознавание образов, компьютерное зрение, машинное обучение и численные методы. E-mail: pro100-pioner@mail.ru, vaganovs@ivanovo.ac.ru.

ГРНТИ: 27.41.23

Поступила в редакцию 12 февраля 2019 г. Окончательный вариант – 25 июля 2019 г.

Adaptive ANN-based method of constructing an interpolation formula for doubling the image size

S.E. Vaganov¹

¹Ivanovo State University, Ivanovo, Russia

Abstract

The architecture of an artificial neural network that solves the problem of constructing interpolation formulas for doubling the size of images is proposed. The trained model receives a 4×4 matrix as an argument. The result is an interpolation formula represented as a weight vector for 4 points.

A comparison of the main quality assessments of the proposed method with some well-known adaptive approaches is made. The results of the comparative analysis show that the proposed approach has a better interpolation quality than *NEDI* and *DCCI* methods.

Keywords: interpolation, machine learning, artificial neural network, gradient descent, image quality.

Citation: Vaganov S. Adaptive ANN-based method of constructing an interpolation formula for doubling the image size. *Computer Optics* 2019, 43(4): 627-631. DOI: 10.18287/2412-6179-2019-43-4-627-631.

References

- [1] Vaganov SE, Khashin SI. Comparison of doubling the size of image algorithms [In Russian]. *Modelirovanie i Analiz Informacionnyh Sistem* 2016; 23(4): 382-400. DOI: 10.18255/1818-1015-2016-4-389-400.
- [2] Zhou D, Shen X, Dong W. Image zooming using directional cubic convolution interpolation. *IET Image Processing* 2012; 6(6): 627-634. DOI: 10.1049/iet-ipr.2011.0534.
- [3] Jing L, Zongliang G, Xiuchang Z. Directional bicubic interpolation – a new method of image super-resolution. In Book: *Proceedings of 3rd International Conference on Multimedia Technology (ICMT-13)*. Atlantis Press; 2013: 470-477. DOI:10.2991/icmt-13.2013.57.
- [4] Li X, Orchard MT. New edge-directed interpolation. *IEEE Transactions on Image Processing* 2001; 10(10): 1521-1527. DOI:10.1109/83.951537.
- [5] Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2016; 38(2): 295-307. DOI: 10.1109/TPAMI.2015.2439281.
- [6] Plaziac N. Image interpolation using neural networks. *IEEE Transactions on Image Processing* 1999; 8(11): 1647-1651. DOI: 10.1109/83.799893.
- [7] Hu H, Holman PM, de Haan G. Image interpolation using classification-based neural networks. *IEEE International Symposium on Consumer Electronics* 2004: 133-137. DOI: 10.1109/ISCE.2004.
- [8] Test bmp files [In Russian]. Source: http://math.ivanovo.ac.ru/dalgebra/Khashin/bmp_ex/.
- [9] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 2017. Source: <https://arxiv.org/abs/1412.6980>.
- [10] TensorFlow (official site). Source: <http://tensorflow.org>.
- [11] Nasonov AV, Krylov AS, Petrova X, Rychagov MN. Edge-directional interpolation algorithm using structure tensor. In Book: *Agaian SS, Egiazarian KO, Gotchev AP, eds. Electronic Imaging, Image Processing: Algorithms and Systems XIV*. Ingenta; 2016: 1-4. DOI: 10.2352/ISSN.2470-1173.2016.15.IPAS-026.
- [12] Gashnikov M.V. Interpolation based on context modeling for hierarchical compression of multidimensional signals. *Computer Optics* 2018; 42(3): 468-475. DOI: 10.18287/2412-6179-2018-42-3-468-475.
- [13] Implementation of an adaptive ANN-based image interpolation method. Source: <http://math.ivanovo.c.ru/dcomp-math/Vaganov/Interp.html>.

Author's information

Sergey Evgenevich Vaganov (b. 1992) graduated from Ivanovo State University in 2015 (Master's degree in Mathematics and Computer Science). Currently, he is a post-graduate student of Ivanovo State University majoring in «Mathematical Logic, Algebra and Number Theory». Also, he works as a leading programmer at the Mathematic and Computer Science department of Ivanovo State University. His main research interests are digital image processing, pattern recognition, computer vision, machine learning and numerical methods. E-mail: pro100-pioner@mail.ru, vaganovs@ivanovo.ac.ru.

Received February 12, 2019. The final version – July 25, 2019.