

УДК 004.932.1

DOI: 10.18799/29495407/2024/2/53

Шифр специальности ВАК: 1.2.1

Семантическая сегментация воздуха в кернах с шумоподавлением и постобработкой

М.И. Волков✉, П.С. Каргина

Национальный исследовательский Томский политехнический университет, Россия, г. Томск

✉miv5@tpu.ru

Аннотация. Представлены результаты по семантической сегментации воздуха для цифровых кернов из бетона. В качестве предобработки изображений использованы нормировка яркости и шумоподавление. Для увеличения эффективности сегментации предложен метод по сегментации в трех направлениях с последующим суммированием полученных масок и применением медианного фильтра. Полученные маски изображений обладают трехмерной структурой и могут использоваться для формирования обучающего набора данных для 3D-архитектур сверточных нейронных сетей. Также предложенный метод можно использовать для увеличения изображений в обучающей выборке при условии долгого времени разметки данных.

Ключевые слова: семантическая сегментация пор воздуха, шумоподавление, BM3D, постобработка, медианный фильтр

Для цитирования: Волков М.И., Каргина П.С. Семантическая сегментация воздуха в кернах с шумоподавлением и постобработкой // Известия Томского политехнического университета. Промышленная кибернетика. – 2024. – Т. 2. – № 2. – С. 39–44. DOI: 10.18799/29495407/2024/2/53

UDC 004.932.1

DOI: 10.18799/29495407/2024/2/53

Semantic air segmentation in core with noise reduction and post-processing

M.I. Volkov✉, P.S. Kargina

National Research Tomsk Polytechnic University, Tomsk, Russian Federation

✉miv5@tpu.ru

Abstract. This paper presents results on semantic air segmentation for digital concrete cores. Brightness normalization and noise reduction were used as image preprocessing. To increase the efficiency of segmentation, a method is proposed for segmentation in three directions, followed by summing the resulting masks and using a median filter. The resulting image masks have a three-dimensional structure and can be used to generate a training dataset for 3D convolutional neural network architectures. The method proposed can be used as well to enlarge images in the training set if data labeling takes a long time.

Keywords: semantic segmentation of air pores, noise reduction, BM3D, post-processing, median filter

For citation: Volkov M.I., Kargina P.S. Semantic air segmentation in core with noise reduction and post-processing. *Bulletin of the Tomsk Polytechnic University. Industrial Cybernetics*, 2024, vol. 2, no. 2, pp. 39–44. DOI: 10.18799/29495407/2024/2/53

Введение

Метод компьютерной томографии является одним из способов исследования внутренней структуры образца без его разрушения. Таким способом часто исследуется внутренняя структура образцов горных пород или бетонов. Исследование образцов рентгеновским излучением основывается на различии в прохождении рентгеновского излучения в зависимости от плотности объектов в породе (например, излучение спокойно проходит через воздушные поры и практически не проходит через очень плотные камни) [1].

После проведения компьютерной томографии керна на выходе получается цифровой kern – оцифрованная модель образца, которая представляет собой совокупность последовательно идущих изображений. Многообразие исследуемых образцов приводит к тому, что цифровые керны различаются степенью зашумленности изображений, контрастом, типом наблюдаемого шума и многими другими параметрами [2, 3]. Помимо этого, качество цифровых кернов зависит от используемого оборудования и программного обеспечения.

После томографии цифровой kern обрабатывается различными алгоритмами для получения информации о его внутренней структуре [4]. В данной работе представлены результаты по семантической сегментации пор воздуха в образце бетона для их анализа (в частности, подсчет числа пор в образце, определение расположения пор и их площади и т. д.). Помимо семантической сегментации в работе присутствует предобработка данных (шумоподавление и нормировка яркости изображений) и по-

стобработка. Актуальность данной работы заключается в нетривиальности проведения сегментации вследствие большого разнообразия структур исследуемых образцов и разной степени зашумленности исследуемых цифровых кернов [2].

Целью данной работы являлось написание части программного обеспечения для проведения семантической бинарной сегментации воздуха с точностью на тестовой выборке выше 90 % по метрике IOU (Intersection over Union) для дальнейшего анализа внутренней структуры пор воздуха на изображениях бетонов.

Основная часть

В качестве входных данных выступали цифровые керны, которые представляют собой последовательные срезы образца бетона. Для обучения и тестирования нейросети использованы 2048 изображений с разрешением 2048*2048.

Пример одного из срезов представлен на рис. 1.

Изображения, которые необходимо сегментировать, содержат четыре основных класса: песок, галька, воздух и цемент (он на рисунке не показан). Саму гальку можно разделить на несколько видов: «простая галька» (в данной работе сюда отнесли песок и однородную гальку), «пористая галька» и «сложная галька» (далее названия будут использоваться без кавычек). Данные определения придуманы автором и не являются общепринятыми. На рис. 4 в качестве примера разными цветами обозначены выделенные в этой работе виды гальки. В данной работе рассмотрена только задача сегментации воздуха.

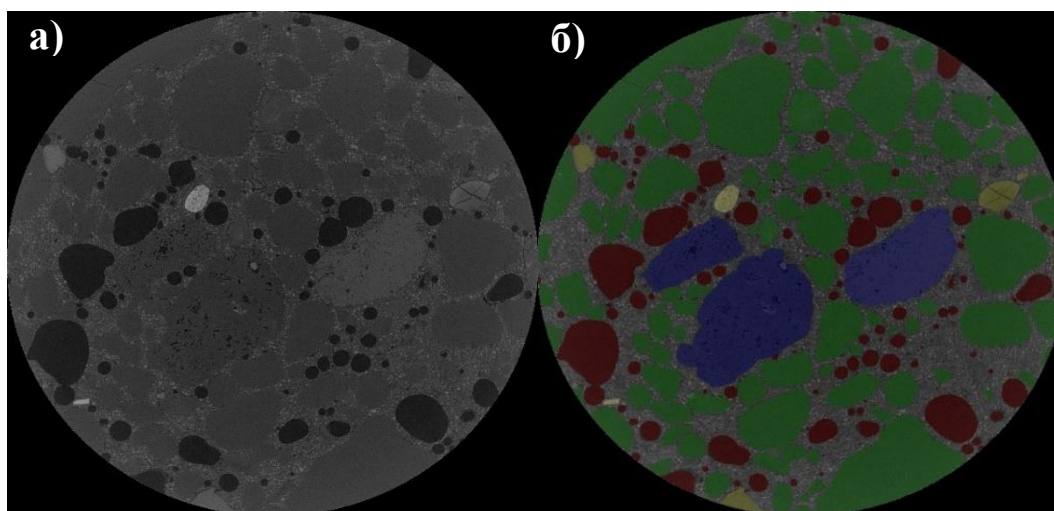


Рис. 1. Изображение слоя цифрового керна (а) и это же изображение с наложенной на него маской (б) (зеленым цветом отмечена «простая галька», синим – «пористая галька», желтым – «сложная галька», красным отмечен воздух)

Fig. 1. Image of a digital kern layer (a) and the same image with a mask superimposed on it (b) (green indicates “simple pebbles”, blue indicates “porous pebbles”, yellow indicates “complex pebbles”, red indicates air)

Поскольку срезы бетона последовательные, сами изображения имеют структуру вдоль трех осей, т. е. данные трехмерные. Маски должны также обладать трехмерной структурой, что накладывает на них ограничения непрерывности: объекты в маске не должны резко появляться или пропадать. Это условие необходимо для работы программы, которая считает количество пор, их размеры и другие параметры из трехмерной маски. Если объект воздуха будет резко появляться и пропадать вдоль одной из осей (то есть объекты воздуха на масках будут прерываться, например, из-за несовершенства работы нейросети), то программа по подсчету объектов воздуха будет давать завышенное число объектов воздуха для такой трёхмерной маски. Например, когда на n -ом изображении определенный объект воздуха, а на изображениях с номерами $n-1$, $n-2$, $n+1$, $n+2$ воздух сегментирован, то программа по подсчету количества объектов воздуха на изображении выдаст завышенные результаты: вместо одного объекта программа увидит два. В рамках этой работы данная проблема носит название «проблема исчезновения воздуха».

Шумоподавление на изображениях методом BM3D

Одним из методов по эффективности сегментации является шумоподавление [5, 6]. Подавление

шума может быть реализовано с помощью методов фильтрации [7], вариационных методов шумоподавления, нейронных сетей (например, CNN [5, 6, 8], GAN-архитектур). В данной работе в качестве инструмента для шумоподавления использован метод пространственной фильтрации BM3D, который базируется на методе Non-local means. Данный метод шумоподавления выбран вследствие простоты использования и высокой степени подавления шума [9].

Результаты шумоподавления для разных классов представлены на рис. 2.

Из полученных результатов видно, что после шумоподавления среднеквадратическое отклонение объекта воздуха уменьшилось практически в 10 раз. Ожидается, что это должно облегчить сегментацию воздуха.

Обучение и тестирование модели

В качестве предобработки изображений использованы шумоподавление, нормировка яркости и нормализация [10]. Нормировка яркости необходима для приведения всех гистограмм изображений к практически одному и тому же виду. В результате яркость и контрастность изображений будут приблизительно одинаковыми. Далее для подавления шума испиливался алгоритм BM3D.

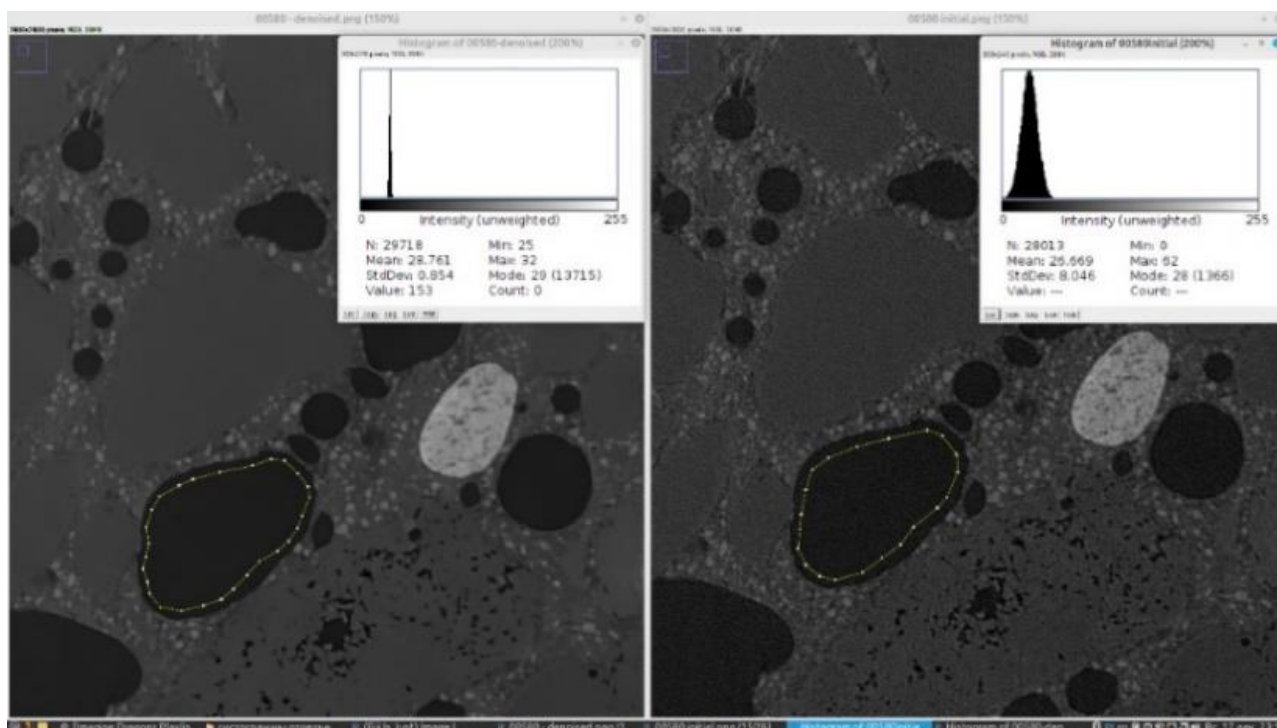


Рис. 2. Изображения с указанным объектом воздуха, гистограммы этих объектов до (справа) и после (слева) шумоподавления

Fig 2. Images with the specified air object, histograms of these objects before (right) and after (left) noise reduction

Объекты воздуха в исходном наборе данных имеют различную площадь на изображениях вплоть до нескольких пикселей. Помимо этого, объекты имеют расплывчатые границы, которые могут быть обнаружены только при большом приближении. В результате сжатия таких изображений довольно много информации будет потеряно. Исходя из этого, одним из критериев для данной задачи является высокое разрешение изображений. Помимо этого, тренировку нейросети необходимо было проводить на ПК (одно из условий в техзадании). Эти условия наложили ограничения на применение трехмерных архитектур нейросетей (не хватит памяти GPU чтобы вместить одно трехмерное изображение) и на применение сторонних сервисов (например, Google Colab).

Исходные изображения с размерами 2048×2048 сжимались до размеров 1024×1024 . Сегментация производилась с использованием библиотеки Pytorch. Для работы было размечено всего 18 изображений (малое число изображений в обучающей выборке является особенностью данной работы).

Параметры обучения нейросети представлены в таблице.

Таблица. Параметры обучения нейросети

Table. Neural network training parameters

Параметр/Parameter	Значение/Meaning
Количество классов/Number of classes	2 (воздух и фон) 2 (air and background)
Число обучающих изображений Number of training images	12
Число проверочных изображений Number of verification images	4
Число тестовых изображений Number of test images	2
Число аугментированных изображений Number of augmented images	400
Используемые методы в аугментации Methods used in augmentation	A.ImageCompression, A.GlassBlur, A.GaussNoise, A.RandomResizedCrop, A.RandomRotate90, A.SafeRotate, A.VerticalFlip, A.HorizontalFlip
Размер партии/Batch size	2
Скорость обучения/Learning rate	0,001
Сокращение веса/Weight decay	0,001
Число эпох обучения Number of training epochs	80
Используемая модель/Model used	smp.Unet(encoder_name='resnet18')
Функция потерь/Loss function	CrossEntropyLoss
Используемые метрики/Metrics used	IOU, accuracy

Зависимость метрик IOU и accuracy от числа эпох в процессе обучения приведена на рис. 3.

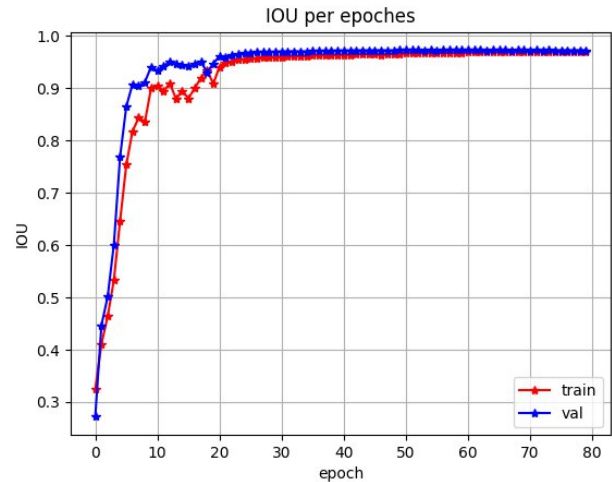


Рис. 3. Зависимость IOU на обучающей и проверочной выборках от числа эпох

Fig. 3. Dependence of IOU on the training and testing samples on the number of epochs

При дальнейшем увеличении числа эпох наблюдалось снижение эффективности на проверочной выборке и увеличение ее на обучающей (переобучение). Эффективность сегментации на тестовых данных составила 0,944.

После теста модель была использована для предсказания масок для неразмеченных изображений в цифровом керне. В качестве примера на рис. 4 представлены два исходных изображения (а, г) и соответствующие предсказанные маски (б, д), которые имеют «проблему исчезновения воздуха».

Из приведенных результатов видно, что нейросеть сегментирует практически весь воздух на изображении. Однако на масках все еще видна «проблема исчезновения воздуха». В данном случае эта проблема, вероятно, возникает по следующим причинам:

- 1) очень малое количество изображений в тренировочной выборке;
- 2) использование 2D-архитектуры вместо 3D.

В связи с недостатком времени на разметку данных эту проблему было принято решить с использованием метода постобработки, который состоит из двух этапов. Первый этап заключается в проведении сегментации в трёх плоскостях цифрового керна: ZOY, XOY и XOZ. Далее три полученные объемные маски складываются попиксельно по логике «2 из 3»: если данный пиксель отмечен классом «воздух» как минимум на двух масках, то на новой маске он будет отмечен также (если нет – пиксель с такими координатами будет помечен классом «фон»). Далее полученная объемная маска фильтруется медианным фильтром размерами (3, 3, 3). Применение фильтра обусловлено тем, что маски в остальных двух проекциях сегментируются достаточно «рвано» и полученная сложением по логике «2 из 3» маска получается сильно зашумленной.

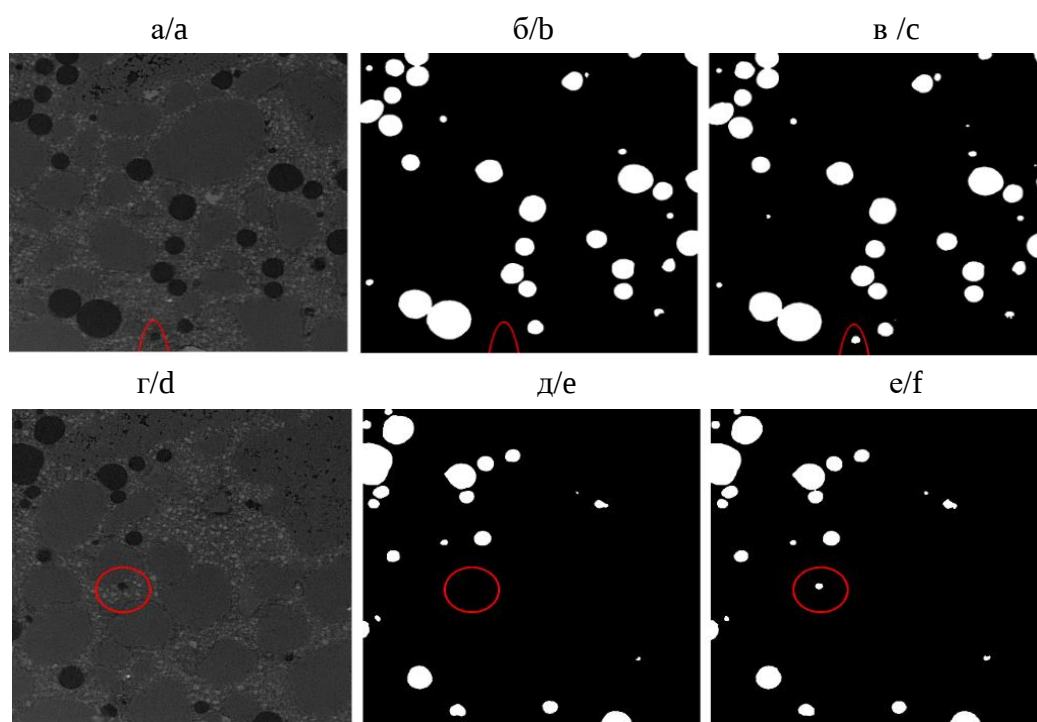


Рис. 4. Примеры сегментации двух разных изображений (а, г), их масок без применения постобработки (б, д) и масок с применением постобработки (в, е). Красным цветом выделен воздух, который отсутствовал на масках без применения постобработки

Fig. 4. Examples of segmentation of two different images (a, d), their masks without post-processing (b, d) and masks with post-processing (c, f). Air that was not present on the masks without post-processing is highlighted in red

В результате применения такой постобработки удалось практически полностью избавиться от «проблемы исчезновения воздуха». Из исходного набора данных для предсказания выбраны 25 изображений, в которых после сегментации без постобработки выявлена «проблема исчезновения воздуха». Далее эти же изображения были проанализированы после применения постобработки. В результате из 25 изображений воздух пропадал только на двух масках после применения постобработки. Примеры результатов постобработки представлены на рис. 4 справа.

К минусам такого алгоритма можно отнести:

- 1) долгое время вычислений. Требуется в три раза больше времени относительно времени сегментации изображений в наборе данных (поскольку производится сегментация по осям ox , oy , oz). Помимо этого, фильтрация полученных масок медианным фильтром занимает более часа для трехмерной маски размерами $1024 \times 1024 \times 1024$;
- 2) использование медианного фильтра будет искажать размеры объектов воздуха на масках, что даст погрешность при расчете их характеристик.

В дальнейшем планируется использование других алгоритмов постобработки и сравнение полученных результатов с используемым в данной работе методом.

Заключение

В ходе проделанной работы реализована часть программного обеспечения, которая ответственна за сегментацию воздуха в цифровых кернах с точностью по метрике $IOU=94,4\%$ (на тестовой выборке). В качестве предобработки выступал метод ВМЗД. Возникшая «проблема исчезновения воздуха» была решена с использованием сегментации в трех направлениях с последующим суммированием полученных масок по логике «2 из 3» и дальнейшей трехмерной фильтрацией объемной маски. Данная объемная маска может быть использована для дальнейшего анализа внутренней структуры пор воздуха на изображениях бетонов.

Используемый метод для постобработки не может быть использован в программном обеспечении из-за слишком большого времени обработки. Однако данный метод может быть использован для увеличения количества изображений в обучающей выборке при условии сложности разметки, поскольку он увеличивает эффективность работы нейросети. Помимо этого, использование данного метода позволит подготовить набор данных для обучения 3D-архитектур сверточных нейронных сетей, поскольку полученные маски обладают трехмерной структурой.

СПИСОК ЛИТЕРАТУРЫ

1. Иванов М.К., Бурлин Ю.К., Калмыков Г.А. Петрофизические методы исследования кернового материала. – М.: Изд-во Моск. ун-та, 2008. – 112 с.
2. Browse Projects // Digital Rocks. – 2024. URL: www.digitalrockportal.org/projects/ (дата обращения: 15.02.2024).
3. Зиньков А.В., Макишин В.Н. Цифровизация керна. – Владивосток: Изд-во Дальневост. федерал. ун-та, 2023. – 73 с.
4. Porespy // Quantitative Image Analysis of Porous Materials. – 2024. URL: <https://porespy.org/> (дата обращения: 15.02.2024).
5. Ilesanmi A.E., Ilesanmi T.O. Methods for image denoising using convolutional neural network: a review // Nigeria. – 2021. – № 7. – P. 2179–2199.
6. Tian C., Xu Y., Fei L. Enhanced CNN for image denoising // China. – 2019. – № 4. – P. 17–23.
7. Гонсалес Р., Вудс Р., Вудс Р.Е. Цифровая обработка изображений. – М.: Изд-во «Техносфера», 2012. – 1104 с.
8. CNN с комплексным знаком для шумоподавления медицинских изображений // MLCENTRE. – 2024. URL: <https://mlcentre.ru/articles/629239/> (дата обращения: 15.02.2024).
9. Описание проекта // bm3d 4.0.1. – 2024. URL: <https://pypi.org/project/bm3d/> (дата обращения: 15.02.2024).
10. Histogram matching with OpenCV, scikit-image, and Python // GeeksforGeeks. – 2024. URL: <https://www.geeksforgeeks.org/histogram-matching-with-opencv-scikit-image-and-python/> (дата обращения: 15.02.2024).

Информация об авторах

Михаил Игоревич Волков, инженер, Научно-образовательный центр перспективных исследований, Национальный исследовательский Томский политехнический университет, Россия, 634050, г. Томск, пр. Ленина, 30, miv5@tpu.ru

Полина Сергеевна Каргина, инженер, Исследовательская школа физики высокоэнергетических процессов, Национальный исследовательский Томский политехнический университет, Россия, 634050, г. Томск, пр. Ленина, 30, psk22@tpu.ru

Поступила: 20.04.2024

Принята: 20.06.2024

Опубликована: 30.06.2024

REFERENCES

1. Ivanov M.K., Burlin Yu.K., Kalmykov G.A. *Petrophysical methods for studying core material*. Moscow, Moscow University Publ. house, 2008. 112 p. (In Russ.)
2. Browse Projects. *Digital Rocks*. 2024. Available at: www.digitalrockportal.org/projects/ (accessed: 15 February 2024).
3. Zinkov A.V., Makishin V.N. *Core digitalization*. Vladivostok, Far East Federal University Publ. House, 2023. 73 p. (In Russ.)
4. Porespy. *Quantitative Image Analysis of Porous Materials*. 2024. Available at: <https://porespy.org/> (accessed: 15 February 2024).
5. Ilesanmi A.E., Ilesanmi T.O. Methods for image denoising using convolutional neural network: a review. *Nigeria*, 2021, no. 7, pp. 2179–2199.
6. Tian C., Xu Y., Fei L. Enhanced CNN for image denoising. *China*, 2019, no. 4, pp. 17–23.
7. Gonzalez R., Woods R., Woods R.E. *Digital image processing*. Moscow, Tekhnosphere Publ. house, 2012. 1104 p. (In Russ.)
8. Complex Signed CNN for Medical Image Denoising. *MLCENTRE*. 2024. Available at: <https://mlcentre.ru/articles/629239/> (accessed: 15 February 2024).
9. Project description. *bm3d 4.0.1*. 2024. Available at: <https://pypi.org/project/bm3d/> (accessed: 15 February 2024).
10. Histogram matching with OpenCV, scikit-image, and Python. *GeeksforGeeks*. 2024. Available at: <https://www.geeksforgeeks.org/histogram-matching-with-opencv-scikit-image-and-python/> (accessed: 15 February 2024).

Information about the authors

Mikhail I. Volkov, Engineer, National Research Tomsk Polytechnic University, 30, Lenin avenue, Tomsk, 634050, Russian Federation, miv5@tpu.ru

Polina S. Kargina, Engineer, National Research Tomsk Polytechnic University, 30, Lenin avenue, Tomsk, 634050, Russian Federation, psk22@tpu.ru

Received: 20.04.2024

Revised: 20.06.2024

Accepted: 30.06.2024