DOI: https://doi.org/10.17816/vto626361

Neural network model development for detecting atypical mitoses in histological slides

Gennadiy N. Berchenko, Nina V. Fedosova, Mikhail G. Kochan, Dmitriy V. Mashoshin

N.N. Priorov National Medical Research Center of Traumatology and Orthopedics, Moscow, Russia

ABSTRACT

BACKGROUND: Modern computer systems allow digitizing and examining images of histological preparations, which led the authors to the idea of using Machine Learning (hereafter — ML) tools usage in digital pathohistology. The ability of neural networks to find sub-visual image features in digitized histological preparations provides the basis for better qualitative and quantitative image analysis. Existing machine learning methods provide good accuracy and speed in recognizing various images, which gives hope for their wide application, including in oncologic diagnostics.

AIM: Use methods of mathematical modeling to identify pathological mitoses in histological preparations as the main sign of the difference between malignant and benign tumor growth.

MATERIALS AND METHODS: Histological images of the N.N. Priorov National Medical Research Center of Traumatology and Orthopedics were used as a data set for the neural network model. The model was tested using 188 histologic slides from 67 patients treated at the institute. Histological preparations were scanned on a Leica Aperio CS2 microscope with a ×400 resolution and converted into JPEG format with further processing. Next, the test images were analyzed in streaming mode using the created neural network model in order to obtain the coordinates of the desired diagnostic object — pathological mitosis and the probability with which the model found the object of this category. The obtained images were analyzed by a pathologist to determine whether the detected object corresponded to pathological mitosis.

RESULTS: The authors have chosen an architecture, developed a methodology for training a neural network, and created a model that can be used to detect pathologic mitoses in histologic preparations. The authors do not attempt to replace the physician, but show the possibility of an integrated approach to data analysis by a computer system and a pathologist.

CONCLUSIONS: The developed mathematical model of neural network used as a part of technological solution for recognizing pathological mitoses in scanned histological preparations can be used as a tool to reduce the time of research and increase the accuracy of diagnosis by a pathologist.

Keywords: neural network; mathematical model; artificial intelligence; tumor; pathological mitosis; machine learning; bone pathology.

To cite this article:
Berchenko GN, Fedosova NV, Kochan MG, Mashoshin DV. Neural network model development for detecting atypical mitoses in histological slides. *N.N. Priorov Journal of Traumatology and Orthopedics*. 2024;31(3):337–349. DOI: https://doi.org/10.17816/vto626361

Received: 01.02.2024 **Accepted:** 13.05.2024 **Published online:** 02.09.2024

DOI: https://doi.org/10.17816/vto626361

Разработка модели нейронной сети для выявления патологических митозов в гистологических препаратах

Г.Н. Берченко, Н.В. Федосова, М.Г. Кочан, Д.В. Машошин

Национальный медицинский исследовательский центр травматологии и ортопедии им. Н.Н. Приорова, Москва, Россия

АННОТАЦИЯ

Обоснование. Современные компьютерные системы позволяют оцифровывать и исследовать изображения гистологических препаратов, что натолкнуло авторов на идею использования инструментов машинного обучения в цифровой патогистологии. Возможности нейронных сетей находить субвизуальные особенности изображения на оцифрованных гистологических препаратах создают основу для лучшего качественного и количественного анализа изображений. Существующие методы машинного обучения дают хорошие показатели по точности и скорости при распознавании различных изображений, что позволяет надеяться на их широкое применение, в том числе и в онкологической диагностике.

Цель. Использовать методы математического моделирования для выявления патологических митозов в гистологических препаратах как основного признака различия злокачественного и доброкачественного опухолевого процесса.

Материалы и методы. В качестве набора данных для модели нейронной сети применялись гистологические изображения НМИЦ травматологии и ортопедии им. Н.Н. Приорова. Тестирование модели выполнено с помощью 188 гистологических стёкол 67 пациентов, проходивших лечение в институте. Гистологические препараты были отсканированы на микроскопе Leica Aperio CS2 с разрешением ×400 и преобразованы в формат JPEG с последующей обработкой. Далее в потоковом режиме был выполнен анализ тестовых изображений с использованием созданной модели нейронной сети с целью получения координат искомого объекта диагностики — патологического митоза и вероятности, с которой модель находила объект данной категории. Полученные изображения были проанализированы врачом-патологоанатомом на предмет соответствия выявленного объекта патологическому митозу.

Результаты. Авторы выбрали архитектуру, разработали методологию обучения нейронной сети и создали модель, которую можно использовать для обнаружения патологических митозов в гистологических препаратах. Авторы не пытаются заменить врача, а показывают возможность комплексного подхода к анализу данных компьютерной системой и врачом-патологоанатомом.

Заключение. Разработанная математическая модель нейронной сети, используемая в составе технологического решения для распознавания патологических митозов в отсканированных гистологических препаратах, может применяться как инструмент для сокращения времени исследования и повышения точности диагностики врача-патологоанатома.

Ключевые слова: нейронная сеть; математическая модель; искусственный интеллект; опухоль; патологический митоз; машинное обучение; костная патология.

Как цитировать:

Берченко Г.Н., Федосова Н.В., Кочан М.Г., Машошин Д.В. Разработка модели нейронной сети для выявления патологических митозов в гистологических препаратах // Вестник травматологии и ортопедии им. Н.Н. Приорова. 2024. Т. 31, № 3. С. 337–349. DOI: https://doi.org/10.17816/vto626361

Рукопись получена: 01.02.2024 **Рукопись одобрена:** 13.05.2024 **Опубликована online:** 02.09.2024

BACKGROUND

The morphological research method has significantly expanded the diagnostic capabilities of clinicians, making the pathologist an integral participant in the diagnostic process. The volume of work involving biopsy studies is continuously growing, reflecting the increasing importance of analyzing biopsy materials in diagnostics and treatment. In modern pathological anatomy, clinical pathology is becoming increasingly significant; this concept is closely related to surgical pathology in English-language literature, denoting a branch of science that deals with intravital diagnostics based on the results of studying the material obtained from biopsies.

The specific aims of pathomorphological studies include clarification of clinical diagnoses or their establishment in unclear cases; identification of the initial stages of the disease; recognition of inflammatory, hyperplastic, and tumor processes; determination of the degree of malignancy of neoplasms; and the dynamics of changes under the influence of the treatment. Together with clinicians, pathologists participate in determining the scope of surgical intervention, establishing the activity of the process and the severity of the lesion, assessing the effectiveness of treatment, and so on.

Morphological examination is mandatory for diagnosing tumors and tumor-like diseases of the musculoskeletal system. Bone tumor diagnostics is one of the most complex areas in clinical oncology, which can be explained by the rarity of such tumors, their diversity, and the pronounced heterogeneity of individual nosological forms. Thus, it is necessary to adhere to a certain algorithm in conducting a pathological anatomical study, which implies a sequence of certain actions to make a diagnosis.

In the pathomorphological differential diagnostics of benign processes of the musculoskeletal system (i.e., benign tumors and tumor-like diseases) and malignant tumors, one of the most objective criteria for their distinction is the determination of pathological mitoses [1]. It is particularly difficult to differentiate benign processes from low-grade sarcomas, such as osteosarcoma, chondrosarcoma, fibrosarcoma, and angiosarcoma. Based on morphological features, mitoses are conventionally classified into normal (i.e., typical) and atypical (i.e., pathological). The biological significance of typical mitosis consists in a strictly identical distribution of chromosomes between the daughter nuclei of a dividing cell, which ensures the formation of genetically identical daughter cells and maintains continuity in a series of cellular generations. Mitotic division ensures the growth of multicellular eukaryotes by increasing the population of tissue cells. The presence of pathological mitoses indicates a disruption in the normal course of mitotic division and often leads to the emergence of cells with unbalanced karyotypes, which leads to the development of mutations and aneuploidy. Pathological mitoses are often registered in carcinogenesis, radiation disease, cancer, and precancerous hyperplasia.

In today's world, neural networks are widely used in image recognition tasks. However, there is currently no single industrial solution for recognizing morphological or histological images. The use of neural networks and machine learning, which allow automating the process of recognizing characteristic patterns of cells and the intercellular matrix, as well as providing a specialist with the analysis results as auxiliary information for making an informed decision when making a diagnosis, will significantly reduce the time for data processing, improve the quality of diagnostics, and reduce the physical and mental load on the pathologist-expert.

The problems of this study are associated with the following difficulties:

1. First, there is no methodology for selecting and refining the architecture of a neural network for analyzing histological images, and as a result, there are no pretrained models created on data sets from histological images, which slows down research.

2. Second, there are no high-quality prepared data sets from histological images in the public domain.

3. Third, the development of artificial intelligence (AI) models for the pathohistological diagnostics of tumors, if they exist, is not presented in a sufficient number of publications from both a medical and technical point of view, which was the reason for using trial and error in this study.

In addition, the size of atypical mitoses is very small, and the forms are diverse. All these points are aggravated by disagreements among experts in recognizing atypical mitoses.

In this paper, we demonstrate a new approach and methodology for developing a mathematical AI model for detecting specific objects in histological images characteristic of oncological diseases of the musculoskeletal system. This study is based on the idea of the presence of pathological mitoses as one of the main criteria for the malignancy of a tumor process in bone tissue.

This research aims to evaluate the possibility of searching for pathological mitoses using an artificial neural network and to create a prototype of a software solution to assist a doctor in the differential diagnostics of benign processes (i.e., benign tumors and tumor-like diseases) and malignant tumors of the musculoskeletal system.

The following problems were solved within the research and development work:

- assistance to a pathologist in the pathomorphological diagnostics of oncological diseases;
- additional training of the AI model to minimize the number of false responses;
- automated stream processing of an array of histological images.

The study objectives were the following:

- selection of the neural network architecture;
- formulation of nonfunctional requirements for the data set and its creation in accordance with these requirements;

- development of a mathematical model for the analysis of images of digitized histological preparations of patients with tumors and tumor-like diseases of the musculoskeletal system in streaming mode without the participation of a pathologist.

Based on the results of this work, we can state the development of an AI model, as well as software, for stream processing of digitized histological preparations using this model. Despite the successful and promising initial results, further testing and refinement of this tool in an oncology institution is required, which will allow its use in everyday clinical practice.

MATERIALS AND METHODS

A collection of histological images, collected by Professor G.N. Berchenko for over more than 30 years of his work as the head of the pathological anatomy department of the N.N. Priorov National Medical Research Center of Traumatology and Orthopedics, was used as a data set for training the neural network model.

Model creation and training

The first rather serious problem that we faced was the determination of neural network architecture for creating the model. Therefore, existing reviews were used to select a suitable architecture [2–4].

We analyzed the architecture of the Faster R-CNN neural network [2–4] for object detection [5–8] as having the best indicators of precision and recall, according to the results of tests conducted at the beginning of the study (May 2018).

Initially, only images taken with a photo attachment to an Olympus BX51 microscope under immersion at a magnification of ×1000 were used to train the model. This approach enabled the creation of a model that, starting with 300 images in the training data set, could identify the desired objects with high accuracy. However, attempts to search for objects in images taken at a magnification of ×400 were unsuccessful. This fact made it impossible to use the model to recognize scanned histological preparations on a Leica SC2 scanning microscope, which has a maximum magnification of ×400, and to perform stream processing of histological preparations. We were compelled to search for ways to solve this problem. After several unsuccessful attempts to refine the model, We empirically came to the solution to select images taken with an attachment to an Olympus BX51 microscope in the amount of 1,200 pieces (i.e., 600 pairs) of photographs of the same pathological mitosis at a magnification of ×1000 and ×400.

A separate task was the formulation of nonfunctional requirements for images. We had to determine what image size was required to train the model. On the one hand, the image size should be comparable to the microscope's field of view; on the other hand, it is necessary to calculate the minimum number of image pixels sufficient to create the model. If, at a magnification of ×1000, the image had a size of 1632×1229 pixels (Fig. 1), the size of the selected element was 90×87 pixels (Fig. 2). At a magnification of ×400 (Fig. 3), the size of the selected fragment was only 30×40 pixels (Fig. 4). Therefore, first, it was necessary to determine the requirements for the image size that should be used in this model.

In addition, it was necessary to change the neural network architecture to create the model. Initially, the Faster R-CNN architecture was used. However, the created model showed the following results based on functional testing. The original files with histological images were "cut" into sections of 1024×1024 pixels and streamed through the model. We

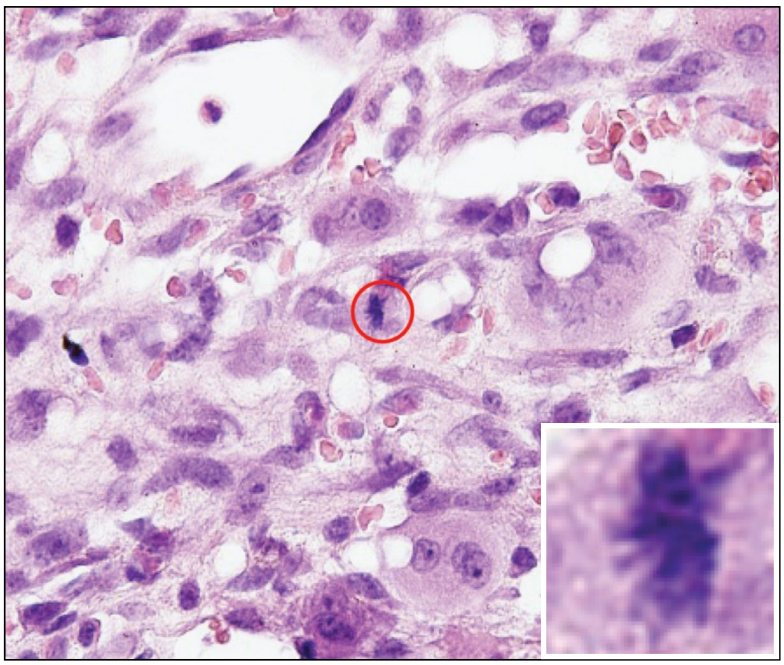

Fig. 1. Photograph of a histological image of a fragment of a malignant tumor with the presence of pathological mitosis, taken at a magnification of ×1000, resolution 1632×1229 pixels.

Fig. 2. A fragment of the image "pathological mitosis", selected for recognition by the model at a magnification of ×1000, resolution 90×87 pixels.

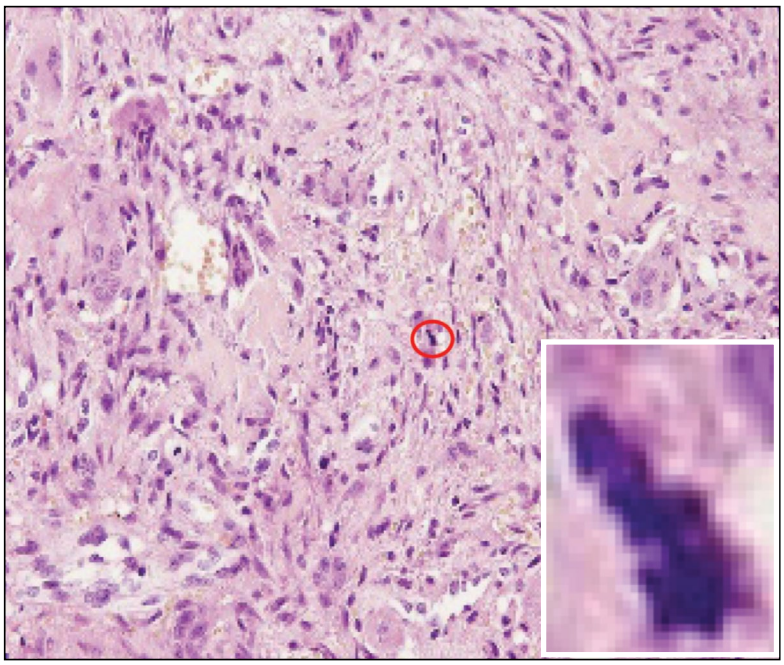

Fig. 3. Photograph of pathological mitosis at a magnification of ×400, resolution 1632×1229 pixels.

Fig. 4. A fragment of the image "pathological mitosis", selected for recognition by the model at a magnification of ×1000, resolution 30×40 pixels.

obtained 3,000 to 10,000 files with a file size of 600 to 1500 MB from just one digitized image. Even with 1% of false responses, they obtained up to 100 false positive images that needed to be rechecked by a pathologist, which did not correspond to the aim of this study, which was to simplify and speed up the doctor's work during the primary processing of histological images. Increasing the number of images that were used to create the model also did not help to reduce the percentage of false responses. Therefore, it was decided to change the architecture of the neural network. We considered a neural network with the Mask R-CNN architecture developed by Facebook Research. One of the features of the network with the Mask R-CNN architecture is its ability to highlight the contours (or "masks") of instances of different objects in photographs, even if there are several such instances and they have different sizes and partially overlap. Since the key problem for this study was the ability to train the model to recognize images that have different sizes (i.e., magnification), we hoped that this architecture would help solve this problem.

In addition, the Mask R-CNN architecture was open source [9], which was important for this study. The model was retrained using the open-source Detectron architecture (Mask R-CNN) [10]. Detectron2 enables working with both Model Zoo (i.e., a set of pretrained models) and CNN-Zoo (i.e., the most famous neural network architectures, including Faster R-CNN, Mask R-CNN, RetinaNet, DensePose, Cascade R-CNN, Panoptic Feature Pyramid Network [FPN], and TensorMask). This architecture enables using synchronous batch normalization and new data sets for object recognition. It has integrated modules that enable modifying the model architecture. In addition, Detectron2 enables using various data set formats, which makes it easy to train the models additionally using the automatic labeling function.

Pretrained models usage significantly reduces the additional training time for the model, since the main weights of the model have already been calculated. The additional training time is reduced from several months to several hours, and the computer power required for additional training can be an Nvidia video card that supports CUDA software. All these conditions were critical in this study, since technical resources are very expensive, and their use would significantly increase the cost of the study.

In this study, we used the Nvidia RTX1070 video display card and the CUDA 11.3 software package. All the Detectron2 capabilities, as well as the open-source code, enabled using it as the main tool in this study.

One of the main quality criteria for the created mathematical model is the minimum number of false responses. However, in many cases, for diagnostics, a pathologist needs to see not only pathological mitoses, of which the number can be very small, but also all sorts of premitosis. Therefore, an expert physician studied all false responses of the model to determine whether a certain finding of the model can be considered false (therefore, the model needs to be further trained) or this identified object should be left in the resulting sample, and only the physician can decide whether this finding is evidence of malignancy or not. For this purpose, based on the results of the initial testing of the model on histological preparations, an expert pathologist analyzed the false positive findings of the mathematical model. Based on the conclusion, further additional training of the model was performed.

For this purpose, individual categories of similar images, from the model's point of view, were allocated to separate subcategories, and additional images of these subcategories were added to the data set so that each subcategory had an equal number of images.

{"webAPIId":"fb765913-6cfb-4fac-91ca-d48319aba49c", "imageUrl":"http://mitoz:9080/mitoz-api/fb765913-6cfb-4fac-91ca-d48319aba49c/2e70df53-7084-4c3a-b981-5801382ac890.jpg","imageMd5":"b879472e619e839b6d2e0df8d5335b4b","classified":[{"confidence":0.974871015548706,"ymax":2121,"label":"mitoz","xmax":1875,"xmin":1673,"ymin":1929,"polygons":[[[1673,1929],[1673,2119],[1695,2143],[1860,2143],[1875,2127],[1875,1929]]]}],"result":"success"}

Fig. 5. The model returned a response indicating that it successfully found the "mitosis" object.

Model testing

The model was tested in several stages. Initially, testing was conducted on new images taken at ×400 and ×1000 magnification using a photo attachment to the Olympus BX51 microscope. The completed control testing of the mathematical model of the neural network showed its high accuracy and the possibility of obtaining good results.

Functional testing of the model was performed using 188 histological slides of 67 patients who sought consultation at the N.N. Priorov National Medical Research Center of Traumatology and Orthopedics. Histological preparations were scanned on a Leica Aperio CS2 scanning microscope with a resolution of ×400 and converted to JPEG format. Then, the JPEG image was "cut" using a computer program into square sections comparable to the field of view of a microscope of 1024×1024 pixels. Then, all the obtained squares of the scanned images were analyzed in streaming mode using the developed neural network model. The neural network model presented the result of the image analysis as a JSON file containing the coordinates of the identified diagnostic object (i.e., mitosis) and the probability with which the model classified this object in this category (Fig. 5).

After that, the identified "mitosis" object was outlined in the image with a square with a diagonal of x_{min}, y_{min} / x_{max}, y_{max} using a script, where x_{min}, y_{min} are the minimum coordinates of the object, and x_{max}, y_{max} are the maximum coordinates of the object. In addition, the probability of the object belonging to this category was revealed (Fig. 6). The resulting images were analyzed by a pathologist to confirm or refute the resemblance of the identified object to pathological mitosis.

In addition to searching, the program counted the number of objects detected. If the sides of the squares that outlined the images intersected (Fig. 7), the objects were considered identical, and the total number of pathological mitoses revealed did not increase.

RESULTS

Clinical testing of the neural network model that determines the search object "pathological mitosis" in scanned images of histological preparations is illustrated by the following clinical examples.

During the study, 188 histological slides of 67 patients with both benign tumors and tumor-like processes and

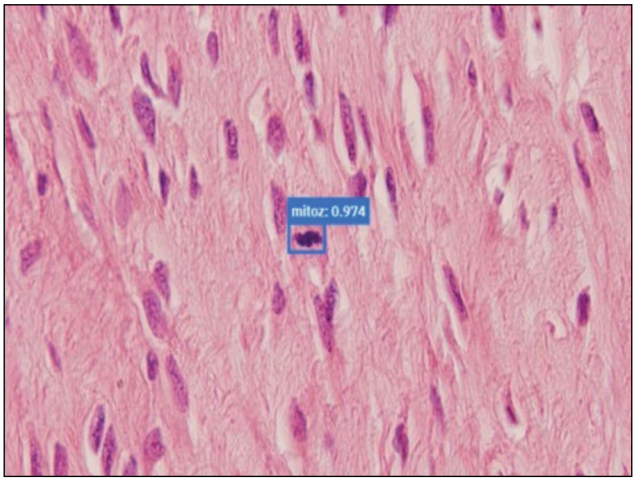

Fig. 6. The result of a program for drawing an object on an image. The model detected pathological mitosis with a probability of 97.4%.

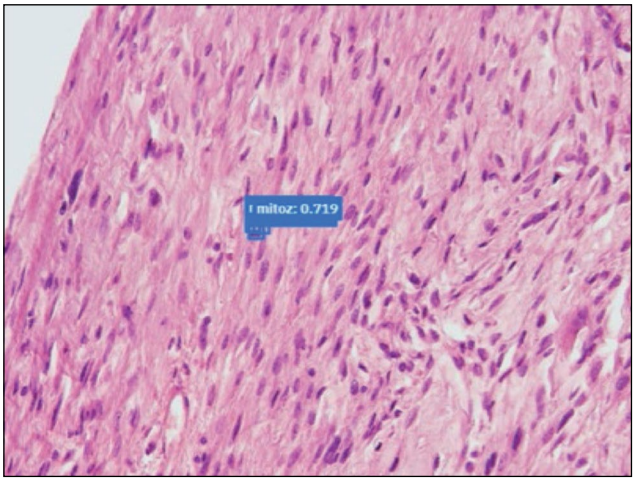

Fig. 7. The model defined one object as several.

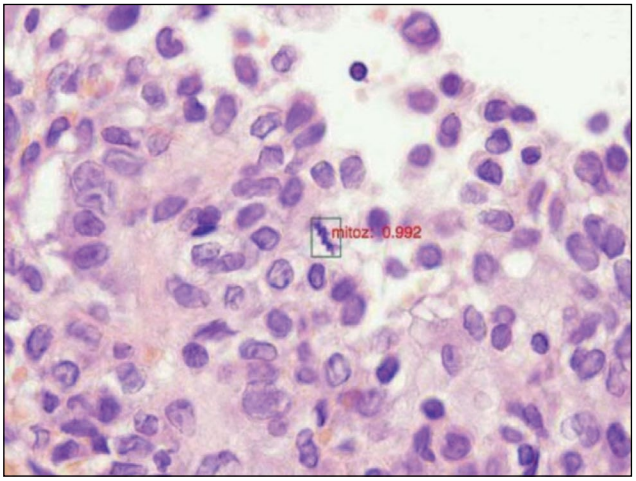

Fig. 8. The model found the object "pathological mitosis" with a probability of 99.2%.

malignant tumors of the musculoskeletal system were scanned and processed using a neural network model. The results of the morphological analysis were presented as a set of images, in which the sought-after objects (i.e., pathological mitoses) and the probability with which the model classified the revealed object into this category were highlighted in a square frame (Fig. 8).

The images in which the model found the "pathological mitosis" objects were presented to an expert pathologist. After analyzing all the images, the expert confirmed that the pathological mitoses revealed on the histological preparations of patients with a probability higher than 70% (Fig. 8) indicated the presence of a malignant tumor of the musculoskeletal system. The "pathological mitosis" objects, detected by the model with a probability lower than 60%, were not pathological mitoses. The expert also confirmed that histological preparations on which the model did not reveal "pathological mitosis" objects with a probability of 70% or higher can be diagnosed as a benign tumor or tumor-like process.

Metrics and characteristics of the created model

Having received a histological image, the model can give two responses:

- Positive indicates the presence of the "pathological mitosis" object in the image;
- Negative indicates the absence of the "pathological mitosis" object in the image.

In this case, during the training process, intermediate control testing of the trained mathematical model of the neural network is performed with the calculation of the following parameters:

- True Positive (TP) indicates the number of correct responses about the presence of pathological mitosis;
- False Positive (FP) indicates the number of false responses about the presence of pathological mitosis;
- True Negative (TN) indicates the number of correct responses about the absence of pathological mitosis;
- False Negative (FN) indicates the number of false responses about the absence of pathological mitosis.

The main metric of the model is the accuracy parameter. It indicates the proportion of correct responses of the algorithm. This parameter measures the percentage of correctly classified images and is calculated by the equation (TP+TN)/(TP+TN+FP+FN). It is noteworthy that this metric is useless in problems with unequal classes, so it is necessary to ensure that the same number of objects is maintained in each category.

For additional assessment of the algorithm quality, the precision and recall metrics are introduced. Precision determines the extent that one can trust the model. It shows the percentage of images that are identified correctly. In other words, when the model searches for an object, precision indicates the frequency it does so correctly, which is calculated as TP/(TP+FP).

Recall determines the number of violations that the model finds. It shows the proportion of objects out of all objects of a given class found by the algorithm, that is, how often it is identified when an image is assigned to a given category. Recall is calculated as TP/(TP+FN).

At the same time, both metrics characterize different aspects of the quality of the trained mathematical model, namely, the higher the precision, the fewer false responses, the higher the recall, the fewer false omissions. Precision and recall, unlike accuracy, do not depend on the ratio of classes and are therefore applicable in conditions of unbalanced samples.

The harmonic mean (or F-measure) was used as the final measure of quality assessment:

$$F=2\times precision\times recall/(precision+recall).$$

The F-measure is the standard in machine learning for obtaining the average accuracy. It reaches its maximum when recall and precision are equal to one and is close to zero if one of the arguments is close to zero.

Characteristics of the obtained model:

- • precision = 0.99834;
- • recall = 1;
- • accuracy = 0.958.

We calculate the F-measure as F = (2×0.99834×1)/(0.99834+1) = 0.99917.

Similarly, from the equation recall = TP/(TP+FN) we obtain:

TP/(TP+FN) = 1 => TP = (TP+FN) => FN = 0; TP = 1.

Substituting the value revealed into the equation, we obtain:

Precision = TP/(TP+FP) = 0.99834 => FP = 0.0016.

Thus, the probability of obtaining false positives, according to the internal characteristics of the model, is less than two tenths of a percent (i.e., from 4 to 10 false responses per histological slide). This paper provides the following case as a clinical example.

Patient E, born in 1996, at the age of 17 noted the appearance of a tumor-like formation in the distal part of the right femur. In 2016, surgery to remove the neoplasm was performed at the healthcare facility at the place of residence; the histological conclusion was osteochondral exostosis. Subsequently, the patient noted a relapse of the neoplasm. In 2019–2022, she was examined at the healthcare facility at the place of residence; continued growth and deterioration were detected, and therefore, it was recommended to contact a federal center. The patient contacted the N.N. Priorov National Medical Research Center of Traumatology and Orthopedics. The presented CT scans from 2019 to 2022 and magnetic resonance imaging from 2022 revealed a massive parosteal neoplasm that had spread into the soft tissues of the thigh and lower leg, with involvement of the femoral vascular bundle, sciatic nerve, and popliteal vascular-nerve bundle in the tumor process. When reviewing histological preparations from 2016, a bone neoplasm was detected, which was represented by bundles of fibroblasts and osteoblast-like cells randomly located in the collagen stroma with mild nuclear atypia, single pathological mitoses, signs of atypical osteogenesis, namely, afunctionally located bone trabeculae of varying degrees of maturity, and infiltrative tumor growth into the adjacent striated muscle fibers. Histological conclusion included the histological presentation, taking into account the data of radiation diagnostic methods, corresponding to parosteal osteosarcoma of the GI of the distal right femur.

Within the software testing, histological preparations were scanned and examined using the developed software. Single figures of pathological mitoses were revealed in the histological preparations, which is consistent with the histological conclusion of this study.

DISCUSSION

In most of the analyzed studies, the pathologist's responses are compared with the result of the mathematical model, which, in our opinion, is a conceptual error, since pathologists themselves do not always agree on the diagnosis, and the diagnosis made by the majority of votes is not always correct. From the standpoint of this study, it is advisable to use AI models as a tool for the primary processing of the incoming flow of graphic information received during scanning of histological preparations.

In addition, in the process of analyzing publications on similar topics, we noticed that scientific studies in the field of pathohistological diagnostics use neural networks not to detect specific histological patterns, but to classify or, at best, segment histological images, which significantly reduces the accuracy of the work and, as a result, the reliability of the study. In addition, such an approach is impossible in streaming mode and requires additional time and knowledge in the field of computer technology from the doctor, which can complicate the work of the pathologist. Therefore, we believe such tools most probably will not be in demand among specialists.

We are aware of the study by Pantanowitz et al. [11], which reported the implementation of an AI-based algorithm for the automatic detection of prostate cancer, but the magnification of ×6 and ×200 used by that study on histological preparations when training the model casts doubt on the possibility of using this model in clinical practice. The image classification method used in that article [11] does not search for objects in the images, as shown by the illustrations. Using the classification method, it is possible only to determine the category to which the image belongs, the probability with which the model assigns the image to this category, and a heat map. Due to the low accuracy of the classification method, high requirements are imposed on the images, in terms of size, magnification, and preprocessing. All image fragments, both used to create the model and processed by the model, must be of the same size and be performed at the same magnification. The classification method cannot process all histological slides and identify fragments on them that belong to a particular category. In addition, the work under consideration does not present either the issues of choosing the neural network architecture for creating the model, or the requirements for the images.

The main problem with using the classification method is the need to use a large number of prepared images to create the model. The study by Pantanowitz et al. reported 1,357,480 labeled image areas [11]. However, as it is known, models constructed using the classification method require preliminary preparation of the image data set, while the topic of image preprocessing is not mentioned in this article.

The study by Pantanowitz et al. [11] focused on the development of a deep learning algorithm to improve the assessment of prostate cancer according to the Gleason scale. However, the article did not present the most important aspect, namely, what pathohistological features were trained in the "machine vision," on the basis of which differential diagnostics was performed between adenocarcinomas with different degrees of malignancy. In addition, the classification method was also used in that study. If the images in the data set used to train and test the AI model have significant differences, a large number of omissions and false responses can occur, which makes it impossible to use the model in clinical practice. Therefore, we question both the possibility of using the classification method for histological diagnostics in clinical practice and the practical value of the AI model created in the said study [11].

The first rather serious problem that we faced was the choice of the neural network architecture. Theoretically, the more layers a neural network has, the better the result it can show. However, as the layers of a neural network increase, its accuracy sharply decreases, which is caused by the disappearance of the gradient (i.e., backpropagations). This happens because the backpropagation process finds the derivatives of the entire network, moving from the last layer

to the first layer. According to the chain rule for calculating derivatives, the derivatives of each layer are multiplied by each other to calculate the derivatives of the input layers. The repeated multiplication process makes the derivatives and, therefore, the weights infinitely small. Therefore, the thresholds of the input layers are not updated during the training process. Since these input layers are critical for recognizing key elements of the input data, this results in inaccuracy of the entire network and slow learning speed.

Therefore, we analyzed the architecture of the neural ResNet, which uses residual blocks that bypass one or more layers. The residual block trains the residual function, and adding residual blocks allowed preserving large gradients to the original layers, mitigating the vanishing gradient problem.

To create a neural network model, we used the open-source framework Detectron2, released by Facebook AI Research. Detectron2 enabled working with both Model Zoo (a set of pretrained models) and CNN-Zoo (the most famous neural network architectures). In addition, Detectron2 enabled using various data set formats, which made it easy to train the models additionally using the automatic labeling function.

Knowing the specifics of the problem being solved, namely, searching for objects measuring 30×40 pixels in an image measuring 4K pixels, it was necessary to take these features into account when choosing the architecture implementation.

We compared the architectures of the ResNet network with a depth of 50 layers. ResNet extracted the signs from the last convolutional layer of the stage 4, which is called C4. In Detectron2, the architecture of this network was presented as ResNet-50-C4.

We tested another version of the architecture using the FPN technology, which uses a descending architecture with lateral connections to construct a pyramid of functions in the network from single-scale input data. Despite the relevance of the problem, the we were able to find only a few scientific studies on the creation of a mathematical AI model for detecting pathological mitoses, cells, and elements of the intercellular matrix, presented in Russian or international publications. We compared their results with existing scientific publications [12]. Similar to other studies mentioned, they tested the model on image fragments, as well as data sets in the public domain (MITOS2012), and compared the results with that of the developed model.

The ResNet-101 architecture, which has 101 layers, was not previously tested due to the large number of layers and, as a result, the high probability of a large number of false responses. Using the pretrained model from Model Zoo Detectron2, we additionally trained the model using the aforementioned architecture. However, the testing results were unsatisfactory. The model metrics became worse, and functional testing showed poor results.

Another serious problem that we had to solve was the processing of large-sized images. The size of a scanned histological image ranges from 600 MB to 1.5 GB, whereas the image of pathological mitosis, which the model must find, does not exceed approximately 400 bytes (i.e., 30×40 pixels). In the process of creating and training the model, we faced the problem of detecting small objects in large areas, which is discussed in the article [13]. Due to the high quality of the original images, we were able to divide the image into segments of approximately 400 KB. The algorithm for constructing the model is similar to that proposed in the work of Simonyan and Zisserman [14]. Images taken at different magnifications were used to train the model.

As noted earlier, despite the relevance of the problem, we were unable to find scientific studies on the creation of a mathematical AI model for detecting pathological mitoses, cells, and elements of the intercellular matrix, presented in Russian or international publications. Most of the studies found were of a review or descriptive nature [15–17].

As far as we know, this is the first report on the creation of an AI-based algorithm that allows for the streaming detection of pathological mitoses characteristic of malignant bone tumors in digitized histological preparations, as well as the first example of the clinical use of an AI-based algorithm in bone and joint pathology.

We also reviewed other studies on the use of mathematical models for cancer diagnostics. In studies on the development of a deep learning algorithm to improve prostate cancer Gleason grading [11, 12], they were unable to find out what pathohistological signs the mathematical model was trained on. In addition, these studies did not indicate anything about the architecture used and the creation of the data set, and the choice of model architecture took up a significant part of this study. The quality and size of images that were used to train and test the AI model affected its metrics and the possibilities of using it in diagnostics. We question the applicability of the classification method for histological diagnostics due to its inaccuracy, as well as the practical value of the AI models created in such studies [11, 12, 17, 18] due to the lack of information about the neural network architectures used. In addition, we [11, 12, 17, 18] did not mention the problems we encountered in the process of creating the model, for example, increasing the accuracy of the model and eliminating false responses.

According to this paper, the main problem that still prevents the use of AI models in clinical practice is the lack or insufficient understanding of the basics of pathomorphology on the part of specialists who create AI models, which, as a result, ignores the experience of a pathologist. Mathematicians engaged in data analysis are too keen on comparing neural network architectures and do not try to find practical applications for them. We did not find a single work where histological image data sets were used to test the MASK R-CNN architecture. Studies containing information on testing and comparing neural

network architectures were performed using the Pascal VOC or COCO data sets [6–13].

In addition, specialists who create terabyte databases of digital images often have neither basic knowledge nor experience in creating neural networks and, as a result, do not understand at which magnification these images should be taken and what quality of digitalized histological preparations are suitable for creating AI models. At the same time, expert pathologists are skeptical about new technologies and are in no hurry to share either their accumulated experience and knowledge, or their own data sets of histological images captured over many years of work.

Therefore, in existing approaches to the use of AI in histological diagnostics, there is a significant gap between the mathematical and medical components of this study. All the articles that were studied had a clear bias in one direction or the other, missing the very idea of creating an approach to diagnostics that is understandable to both a doctor and a mathematician, namely, the "task–tool" link.

This paper reports on the development of an AI model for stream processing of digitalized histological preparations and the introduction of this tool into everyday clinical practice. The mathematical model was developed by a team of mathematicians and programmers using the object detection method and the open architecture of the Detectron neural network (Mask R-NN). The created model was tested on an external data set by an expert pathologist with over 30 years of experience in this field to identify pathological mitoses in malignant tumors of both low and high grades of malignancy of the bone and joint system.

As a result of the conducted research, we recommend adhering to certain rules when setting a problem, which should be formulated by specialized medical personnel, as well as when constructing a mathematical model. The basic rules for constructing a model are presented as follows:

- selection of neural network architecture based on the analysis of work on testing this architecture and its suitability for the task;
- formulation of nonfunctional requirements for data set images for both creating and using the model;
- creation of a data set in accordance with the nonfunctional requirements for the architecture selected;
- labeling the data set in accordance with the criteria of expert pathologists;
- use of more than a thousand image fragments for training one category;
- use of three or more categories for recognition;
- absence of jumps or attenuation of metrics during training;
- repeated functional testing on a data set of more than 100 images.

CONCLUSION

This paper has developed a mathematical model of the neural network, used as part of the hardware and software complex for the continuous recognition of pathological mitoses in scanned histological preparations, which can be used as a tool to help a pathologist conduct diagnostics. According to this paper, further elaboration of the model by adding modern research in the field of mathematical methods to the CUDA library of mathematical functions, as well as the development and implementation of data set optimization methods when training the model, are the most promising areas of AI development that must be used for processing and analyzing histological images. This technology will reduce the time and improve the quality of the study, thereby increasing the chances of early diagnostics of the disease and, as a result, the success of patient treatment.

ADDITIONAL INFO

Autor contribution. All authors confirm that their authorship meets the international ICMJE criteria (all authors have made a significant contribution to the development of the concept, research and preparation of the article, read and approved the final version before publication).

Funding source. The work was carried out within the framework of realization of the research work under the state assignment "Development of a mathematical model of neural network for the system of support of decision making by a pathologist in the diagnosis of diseases of the musculoskeletal system" (RK № 124040100041-5).

Competing interests. The authors declare that they have no competing interests.

Consent for publication. Written consent was obtained from the patient for publication of relevant medical information and all of accompanying images within the manuscript.

ДОПОЛНИТЕЛЬНО

Вклад авторов. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, проведение исследования и подготовку статьи, прочли и одобрили финальную версию перед публикацией).

Источник финансирования. Работа выполнена в рамках реализации научно-исследовательской работы по государственному заданию, РК № 124040100041-5, «Разработка математической модели нейронной сети для системы поддержки принятия решения врачом-патологоанатомом в диагностике заболеваний опорно-двигательного аппарата».

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с проведённым исследованием и публикацией настоящей статьи.

Информированное согласие на публикацию. Авторы получили письменное согласие законных представителей пациента на публикацию медицинских данных и фотографий.

REFERENCES

1. Dorfman HD, Czerniak B. *Bone tumors.* 2nd edition. St. Louis: Mosby; 2015. 1261 p.
2. Girshick R. Fast R-CNN. In: *IEEE International Conference on Computer Vision (ICCV)*; 2015.
3. Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans Pattern Anal Mach Intell.* 2017;9(6):1137–1149. doi: 10.1109/TPAMI.2016.2577031
4. Ren S, He K, Girshick R, Zhang X, Sun J. Object Detection Networks on Convolutional Feature Maps. *IEEE Trans Pattern Anal Mach Intell.* 2017;39(7):1476–1481. doi: 10.1109/TPAMI.2016.2601099
5. Lin T-Y, Dollár P, Girshick R, He K, Hariharan B, Belongie S. *Feature Pyramid Networks for Object Detection, Computer Science > Computer Vision and Pattern Recognition* [Submitted on 9 Dec 2016 (v1), last revised 19 Apr 2017 (this version, v2)]. Available from: https://arxiv.org/abs/1612.03144
6. Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *CVPR*; 2014. arXiv: 1311.2524.
7. Girshick R, Donahue J, Darrell T, Malik J. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Trans Pattern Anal Mach Intell.* 2016;38(1):142–58. doi: 10.1109/TPAMI.2015.2437384
8. Uijlings J, van de Sande K, Gevers T, Smeulders A. Selective search for object recognition. *International Journal of Computer Vision.* 2013;104(2):154–171. doi: 10.1007/s11263-013-0620-5
9. He K, Gkioxari G, Dollar P, Girshick R. Mask R-CNN. *IEEE Trans Pattern Anal Mach Intell.* 2020;42(2):386–397. doi: 10.1109/TPAMI.2018.2844175
10. *Detectron* [Internet]. Available from: https://github.com/facebookresearch/Detectron
11. Pantanowitz L, Quiroga-Garza GM, BienRonen L, et al. An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study. *Lancet Digital Health.* 2020;2(8):e407–e416. doi: 10.1016/S2589-7500(20)30159-X
12. Pantanowitz L, Hartman D, Yan Qi, Eun Yoon Cho, et al. Accuracy and efficiency of an artificial intelligence tool when counting breast mitoses. *Diagn Pathol.* 2020;15(1):80. doi: 10.1186/s13000-020-00995-z
13. Van EttenIn A. *Satellite Imagery Multiscale Rapid Detection with Windowed Networks.* arXiv: 1809.09978v1. Available from: https://arxiv.org/pdf/1809.09978.pdf
14. Simonyan K, Zisserman A. *Very deep convolutional networks for large-scale image recognition.* 10 Apr 2015. arXiv: 1409.1556. Available from: https://arxiv.org/pdf/1409.1556.pdf
15. Barisonia L, Hodgin JB. Digital pathology in nephrology clinical trials, research, and pathology practice. *Curr Opin Nephrol Hypertens.* 2017;26(6):450–459. doi: 10.1097/MNH.0000000000000360
16. Burt JR, Torosdagli N, Khosravan N, et al. Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks. *Br J Radiol.* 2018;91(1089):20170545. doi: 10.1259/bjr.20170545
17. Mermel C, Kunal Nagpal MS. Using AI to identify the aggressiveness of prostate cancer. *Google Health.* 2020. Available from: https://blog.google/technology/health/using-ai-identify-aggressiveness-prostate-cancer
18. Nagpal K, Foote D, Tan F, et al. Development and Validation of a Deep Learning Algorithm for Gleason Grading of Prostate Cancer from Biopsy Specimens. *JAMA Oncol.* 2020;6(9):1–9. doi: 10.1001/jamaoncol.2020.2485

СПИСОК ЛИТЕРАТУРЫ

1. Dorfman H.D., Czerniak B. Bone tumors. 2nd edition. St. Louis: Mosby, 2015. 1261 p.
2. Girshick R. Fast R-CNN. In: IEEE International Conference on Computer Vision (ICCV), 2015.
3. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // IEEE Trans Pattern Anal Mach Intell. 2017. Vol. 39, № 6. P. 1137–1149. doi: 10.1109/TPAMI.2016.2577031
4. Ren S., He K., Girshick R., Zhang X., Sun J. Object Detection Networks on Convolutional Feature Maps // IEEE Trans Pattern Anal Mach Intell. 2017. Vol. 39, № 7. P. 1476–1481. doi: 10.1109/TPAMI.2016.2601099
5. Lin T-Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. Feature Pyramid Networks for Object Detection, Computer Science > Computer Vision and Pattern Recognition [Submitted on 9 Dec 2016 (v1), last revised 19 Apr 2017 (this version, v2)]. Режим доступа: https://arxiv.org/abs/1612.03144
6. Girshick R., Donahue J., Darrell T., Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: CVPR, 2014. arXiv: 1311.2524.
7. Girshick R., Donahue J., Darrell T., Malik J. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation // IEEE Trans Pattern Anal Mach Intell. 2016. Vol. 38, № 1. P. 142–58. doi: 10.1109/TPAMI.2015.2437384
8. Uijlings J., van de Sande K., Gevers T., Smeulders A. Selective search for object recognition // International Journal of Computer Vision. 2013. Vol. 104, № 2. P. 154–171. doi: 10.1007/s11263-013-0620-5
9. He K., Gkioxari G., Dollar P., Girshick R., Mask R-CNN // IEEE Trans Pattern Anal Mach Intell. 2020. Vol. 42, № 2. P. 386–397. doi: 10.1109/TPAMI.2018.2844175
10. Detectron [Интернет]. Режим доступа: https://github.com/facebookresearch/Detectron
11. Pantanowitz L., Quiroga-Garza G.M., BienRonen L., et al. An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study // Lancet Digital Health. 2020. Vol. 2, № 8. P. e407–e416. doi: 10.1016/S2589-7500(20)30159-X
12. Pantanowitz L., Hartman D., Yan Qi, Eun Yoon Cho, et al. Accuracy and efficiency of an artificial intelligence tool when counting breast mitoses // Diagn Pathol. 2020. Vol. 15, № 1. P. 80. doi: 10.1186/s13000-020-00995-z
13. Van EttenIn A. Satellite Imagery Multiscale Rapid Detection with Windowed Networks. arXiv: 1809.09978v1. Режим доступа: https://arxiv.org/pdf/1809.09978.pdf
14. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. 10 Apr 2015. arXiv: 1409.1556. Режим доступа: https://arxiv.org/pdf/1409.1556.pdf

15. Barisonia L., Hodgin J.B. Digital pathology in nephrology clinical trials, research, and pathology practice // Curr Opin Nephrol Hypertens. 2017. Vol. 26, № 6. P. 450–459. doi: 10.1097/MNH.0000000000000360

16. Burt J.R., Torosdagli N., Khosravan N., et al. Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks // Br J Radiol. 2018. Vol. 91, № 1089. P. 20170545. doi: 10.1259/bjr.20170545

17. Mermel C., Kunal Nagpal M.S. Using AI to identify the aggressiveness of prostate cancer // Google Health. 2020. Режим доступа: https://blog.google/technology/health/using-ai-identify-aggressiveness-prostate-cancer

18. Nagpal K., Foote D., Tan F., et al. Development and Validation of a Deep Learning Algorithm for Gleason Grading of Prostate Cancer from Biopsy Specimens // JAMA Oncol. 2020. Vol. 6, № 9. P. 1–9. doi: 10.1001/jamaoncol.2020.2485

AUTHORS' INFO

*** Gennadiy N. Berchenko,** MD, Dr. Sci. (Medicine), professor;
address: 10 Priorova str., 127299 Moscow, Russia;
ORCID: 0000-0002-7920-0552;
eLibrary SPIN: 3367-2493;
e-mail: berchenko@cito-bone.ru

Nina V. Fedosova, MS;
ORCID: 0000-0002-0829-9188;
eLibrary SPIN: 5380-3194;
e-mail: hard_sign@mail.ru

Mikhail G. Kochan;
ORCID: 0009-0002-0699-1370;
e-mail: mk_system@mail.ru

Dmitriy V. Mashoshin;
ORCID: 0009-0003-5442-5055;
eLibrary SPIN: 5981-4084;
e-mail: dima_mash@mail.ru

ОБ АВТОРАХ

*** Берченко Геннадий Николаевич,** д-р мед. наук, профессор;
адрес: Россия, 127299, Москва, ул. Приорова, 10;
ORCID: 0000-0002-7920-0552;
eLibrary SPIN: 3367-2493;
e-mail: berchenko@cito-bone.ru

Федосова Нина Вениаминовна;
ORCID: 0000-0002-0829-9188;
eLibrary SPIN: 5380-3194;
e-mail: hard_sign@mail.ru

Кочан Михаил Геннадьевич;
ORCID: 0009-0002-0699-1370;
e-mail: mk_system@mail.ru

Машошин Дмитрий Викторович;
ORCID: 0009-0003-5442-5055;
eLibrary SPIN: 5981-4084;
e-mail: dima_mash@mail.ru

* Corresponding author / Автор, ответственный за переписку