Development and application of artificial intelligence for the detection of the lymph nodes metastases in colorectal cancer

Olga A. Mainovskaya, Sergey I. Achkasov, Alexey V. Devyatkin, Elena V. Serykh, Vladimir V. Rybakov, Talgat Kh. Makambaev, Daria I. Suslova, Maria A. Ryakhovskaya

[1]Ryzhikh National Medical Research Centre for Coloproctology (Salyama Adilya str., 2, Moscow, 123423, Russia)
[2]Wave Access Service (Kamennoostrovsky str., 11, bld. 2, St. Petersburg, 197046, Russia)
[3]JSC "National Center of Service integration" (Usacheva str., 33, bld. 2, Moscow, 121059, Russia)

ABSTRACT   *AIM: to create a marked data set (histoscans of lymph nodes) for use in the development of medical decision support systems (based on machine learning) in pathomorphology, which will allow determining the presence of metastatic lymph node lesions in CRC.*
RESULTS: the dataset included 432 files with digital images and markings of 1019 lymph nodes, including lymph nodes with and without metastases. Based on the marked-up data, a neural network model was trained to determine the probability of metastatic lesion for each pixel in the area of interest — the lymph node (Dice 0.863 for the replaced tissue, Dice macro 0.923). In addition, pre- and postprocessing methods were implemented to represent input data in a form acceptable for machine learning and to represent the AI model's response in a form convenient for user perception. Additionally, a neural network model has been developed that predicts the probability of finding artifacts in digital images of lymph nodes with the possibility of forming an artifact probability map (Dice macro0.776; Dice for artifacts 0.552; IoU macro 0.725 and IoU for artifacts 0.451).
CONCLUSION: the developed model is a good basis for the implementation of a full-fledged solution, on the basis of which a system can be developed to assist doctors in finding and evaluating the replacement of tissue structures and determining metastatic lymph node lesions, detecting artifacts and evaluating the quality of digital images.

KEYWORDS: Computational pathology (CPATH), artificial intelligence (AI), MachineLearning (ML), colorectal cancer (CRC), staging, lymph node metastasis (LNM)

CONFLICT OF INTEREST: the authors declare no conflict of interest

ADDRESS FOR CORRESPONDENCE: Mainovskaya O.A., Ryzhikh National Medical Research Centre for Coloproctology, Salyama Adilya str., 2, Moscow, 123423, Russia; e-mail:info@gnck.ru

The accuracy of diagnosis and staging of tumors of various sites determines the prognosis and the choice of treatment approach, largely ensuring its success. In diagnostic disciplines such as radiation diagnostics and pathology (pathomorphology), the diagnosis of the tumor is based on image analysis by recognizing certain patterns and interpreting them. One of the main problems in diagnostics is the reproducibility of diagnostic signs and patterns among specialists, which in conditions of increasing demands of precision medicine may exceed the possibilities of visual assessment by a person. In addition, the problem of shortage of personnel in diagnostic disciplines, especially pathomorphology, leads to an increase in the burden on doctors who have to view and analyze a large number of images, which may affect the quality of diagnostics.

49

Разработка и применение искусственного интеллекта
для обнаружения метастазов в лимфатических
узлах при колоректальном раке   Development and application of artificial intelligence for the
detection of the lymph nodes metastases in colorectal cancer

The development of computer technologies in pathomorphology, the advent of equipment for scanning histological specimens with obtaining a digital image of the entire histological section /specimen with high resolution — whole slide images (WSI) technology and the possibility of its study using computer image analysis programs significantly expanded the possibility of morphometric studies and determining prognostic factors for tumors of various localizations.

The use of computer technologies for the analysis of digital images of histological sections (Computational pathology, CPATH) using machine learning algorithms (ML) and artificial intelligence (AI) for the analysis of digital images of histological specimens has revealed new possibilities of CPATH for diagnostics in pathology. In recent years, there has been a transition to more advanced AI models with the solution of more complex diagnostic tasks [1–4]. An example of the transition from the analysis of limited areas to the study of the entire digital image of the specimen for the purpose of diagnosis can be the CAMELYON 16 and CAMELYON 17 studies, in which the task of detecting breast cancer metastases in lymph nodes using CPATH programs was solved. The main purpose of the study was to reproduce routine diagnostics for the detection of macro- and micrometastases in lymph nodes by means of developed computer algorithms. According to the results of the study, the best tested algorithms worked comparably with the best pathologists in determining macrometastases, while when detecting micrometastases, the most effective algorithms demonstrated the best average AUC (0.885 [95% CI 0.812–0.997]) compared with the AUC of the best pathologists (0.808 [95% CI, 0.704–0.908]). CAMELYON16 was one of the first studies to demonstrate that the interpretation of digital images by deep learning algorithms can be performed with an accuracy that can rival human capabilities [6, 7]. The clinical focus of CAMELYON16 was a powerful incentive that promotes the development of AI in this direction. It should be noted that morphological examination of lymph nodes in the removed specimen for tumors of any site is a time-consuming process, including the detection of lymph nodes in the specimen and their further evaluation after histological processing. At the same time, it is the number of lymph nodes examined that reflects the quality of the morphological examination of the removed specimen and allows for the most accurate staging of the tumor, determining the pN category [8, 9]. Thus, the pathologist is faced with the task of examining a large number of lymph nodes by light microscopy from each patient in order to detect metastases. Automation of repetitive and time-consuming tasks in the context of clinical practice can have a great impact on optimizing the workload of pathologists, which in conditions of shortage of personnel will ensure the necessary quality of histological studies.

Colorectal cancer (CRC) is one of the most common oncological diseases in the world, ranking second in terms of morbidity and mortality in Russia [10]. Improving the results of treatment and reducing mortality from colorectal cancer is a socially significant task, one of the tools for solving which is to increase the accuracy of postoperative staging of CRC by optimizing and improving the quality of pathomorphological examination of the surgical specimen. The stage of CRC based on the results of a pathomorphological study is currently the main prognostic factor determining further treatment tactics and prognosis. In the current Clinical Guidelines (2020), adjuvant chemotherapy is recommended for all patients with colorectal cancer at the stage of pT1-4N1-2M0. The status of lymph nodes, namely their metastatic lesion, is the main indicator determining the appointment of adjuvant chemotherapy [11, 12].

In this regard, the clinical objective of the study was to improve the quality of pathomorphological examination of lymph nodes and, consequently, the staging of colon cancer by creating a promising information product that automatically recognizes metastatically affected lymph nodes on digital images of lymph nodes (histoscans) using artificial intelligence (AI).

MATERIALS AND METHODS

The study was carried out using digital images of lymph nodes obtained using an Aperio T2 Leica scanning microscope (at ×20 magnification) using WSI technology in SVS format with manual marking of normal lymphoid tissue, metastases and artifacts (using the software package for digital image analysis Aperio Imagescope V12.4.0, Aperio Tehnologies). To create the database, archival material of the Pathological Anatomy Unit of the RNMRC of Coloproctology of the Health Ministry of Russia was used.

The study used images larger than 5,000 pixels on each side, stored in RGB8 format. The initial resolution was on average at the level of 0.5 microns/pc. In total, 432 files were used for the work, which represents about 148 GB of data containing digital images of 1,000 lymph nodes. The markup of the files was text data in XML format.

When manually marking up (annotation) on the digital image of lymph nodes, zones of three main classes required for solving the task were identified: "Healthy tissue", "Substituted tissue" and "Artifacts".

An example of a markup image and description management is shown in Figure 1. Since the distribution of the area and number of artifacts and replacement tissue differed significantly, it was concluded that it was necessary to create a separate artifact detection system to eliminate the imbalance in machine learning.

The data in the SVS format are high-resolution image files. This kind of data contains a lot of useful information, which makes it impossible to compress the image to feed to the model, whereas in the training of deep neural networks, a strong limitation is the amount of video memory on the graphics adapter. One of the simplest and, at the same time, effective ways to solve this problem is to divide a large-sized digital image on a grid into cells (crops), which allows to process an image in high resolution, while occupying an adequate amount of video memory for machine learning. The separation of the image into crops is implemented on the basis of a regular grid with a specific cell size (a 1024 × 1024-pixel cell is taken as a basis) at a specific resolution (a resolution of 2 microns / pixel is taken as a basis).

A schematic representation of this process is shown in Figure 2.

In the course of machine learning, data augmentation was carried out — a method of changing data within acceptable limits to create greater variability, which is usually used to reduce the chance of retraining the model, as well as to increase the stability of the model to various kinds of perturbations.
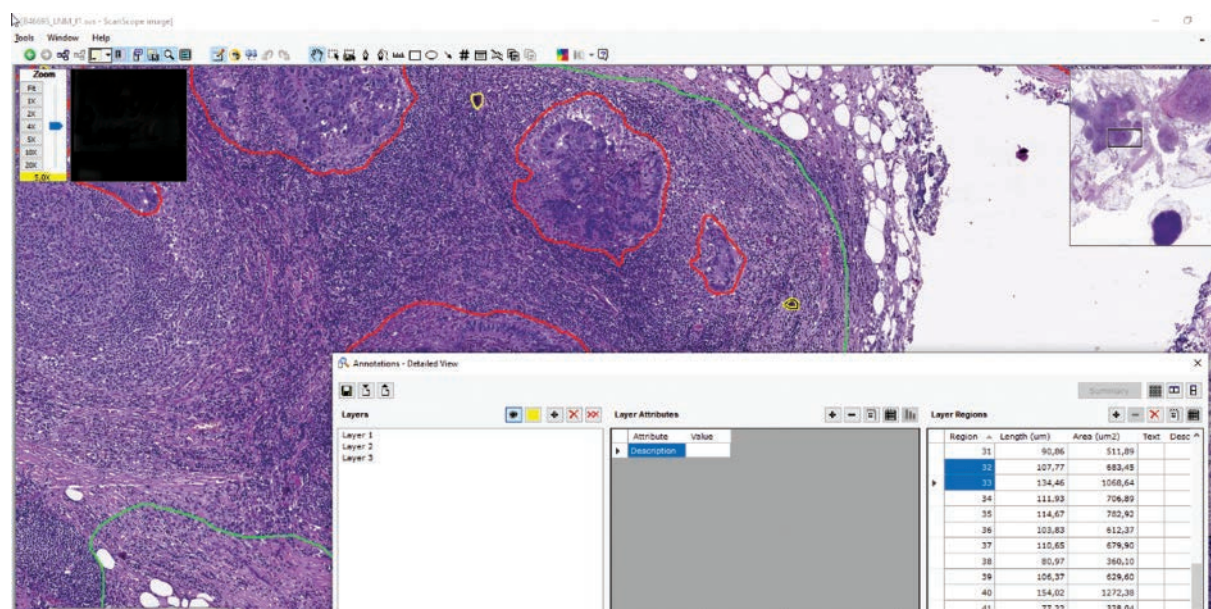

Figure 1. *Example of histoscan and annotation table display in Aperio Image Scope*

51

Разработка и применение искусственного интеллекта для обнаружения метастазов в лимфатических узлах при колоректальном раке

Development and application of artificial intelligence for the detection of the lymph nodes metastases in colorectal cancer

During the training of the model, the following augmentations were used:
- Conversion to gray image (chance = 30%)
- Vertical reflection (chance = 50%)
- Horizontal reflection (chance = 50%)
- Random 90-degree rotation (chance = 50%)
- One of (chance = 40%)
- Gaussian noise (chance = 50%)
- Multiplicative noise on each channel (chance = 30%)
- Image compression within 90% quality (chance = 50%)
- One of (chance = 20%)
- Offset blur (chance = 20%)
- Median blur (chance = 20%)
- Gaussian blur (chance = 20%)
- Random blur (chance = 20%)
- Random normalization/brightness change (chance = 30%)
- Random change of image components in HSV space (chance = 30%)
- Random change of scale (10%), offset (10%), rotation (45 degrees) with black/white background (chance = 50%).

These augmentations allow to change the image during training to reduce the likelihood of retraining and create artificial diversity of data.

The DeepLabV3 architecture was used as the basis for training the tissue segmentation model (https://arxiv.org/abs/1706.05587).

To search for artifacts, the Unet architecture was taken as a basis (https://arxiv.org/abs/1505.04597).

The models were trained using the following parameters and algorithms:
- Adam Optimizer
- Initial learning rate (LR) = 1e-4
- LR Manager Reduce On Plateau
- Number of waiting periods = 10
- Reduction factor LR = 0.5
- Minimum LR = 1e-7
- Number of periods = 200
- Batch size — depending on the image size and available video memory from 8 to 40.

A video card with 16GB of memory (Tesla T4) allows to download a package of images with a resolution of 1024 × 1024. The samples for training and validation were formed as a set of image-mask pairs in the "PNG" format. Images and masks represent areas that have been cut from a specific position of the image (Fig. 3).

Results of Model Training

As the main indicator for evaluating the results of training models, the DICE indicator (Sorensen
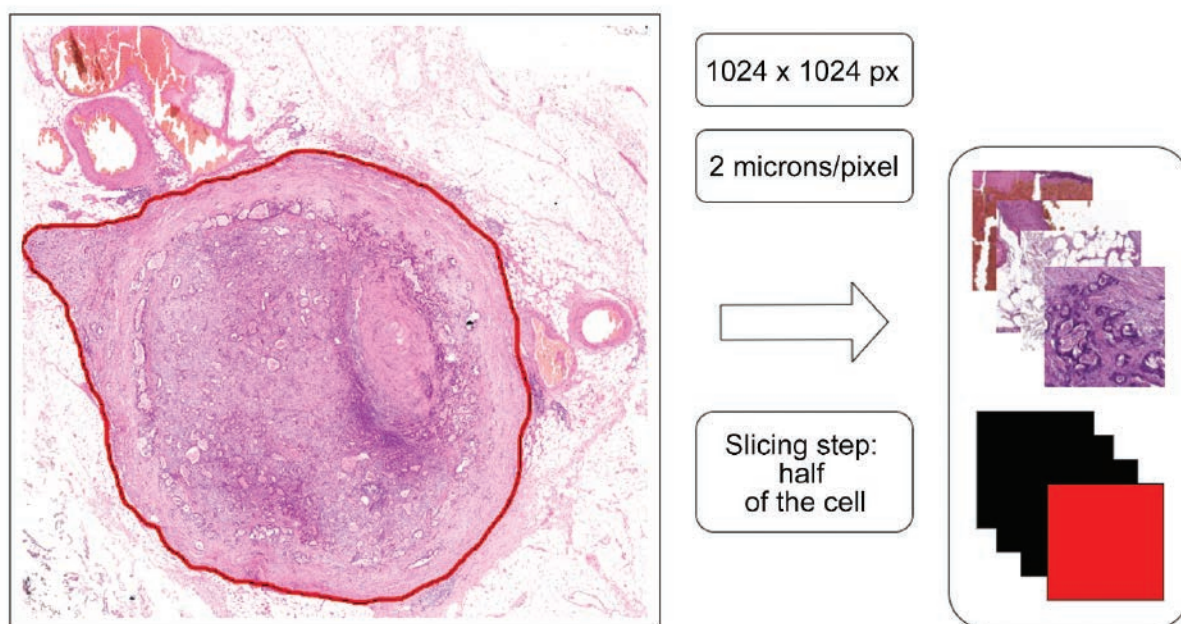

Figure 2. *Diagram of the method of separation («slicing») of the image*

Measure) was used, reflecting the quality of overlap of the predicted mask over the truth. This metric in the context of the segmentation problem is the ratio of the doubled intersection of the true markup and the predicted one to the sum of their areas. With an ideal overlay of the predicted mask on the markup, DICE is equal to one; with a complete discrepancy between the prediction and the mask, DICE is equal to 0. An alternative (more formal) way to determine this measure can be a formula where True Positive means the number of pixels with a prediction that matches the markup, False Positive — pixels where the mask was predicted, but there was no markup, and False Negative — pixels where there was markup, but the model did not predict the presence of a mask.
(https://en.wikipedia.org/wiki/S%C3%B8rensen%E2%80%93Dice_coefficient)
The tables with metrics indicate macro-metrics that are calculated through the calculation of the measure for each of the classes, after which their unweighted average is taken, as well as the metric for the target class separately.

RESULTS

As a result of training the model on the latest data, the final metrics of tissue segmentation had the following indicators:

Indicator	Value
Dice macro	0.923
Dice for substituted tissue	0.863

From the point of view of visual evaluation of the work, the developed model showed very meaningful results. In some cases, there is a preference for specificity, in other words, in a number of images there was a selection of a larger zone than noted in the markup (Fig. 4).

In general, the accuracy of tissue segmentation using the developed model made it possible to make a fairly accurate assessment of lymph node lesions. The algorithm for determining lymph nodes and classifying the type of tissue (healthy/substituted) works on the basis of highlighting large areas (areas of interest) with contours and marking them with descriptive rectangles with further estimation of the percentage of replaced tissue relative to the entire found lymphoid tissue in this zone (Fig. 5).

The accuracy of digital image analysis largely depends on its quality, which in turn is determined by the quality of the original histological specimen. If, when a pathologist views a histological specimen or histoscan, any artifact (a defect of a histological specimen or its digital image that occurred during histological processing of a tissue sample or during scanning) can be easily identified and taken into account when evaluating it, then the presence of an artifact in the computer analysis of digital images can lead
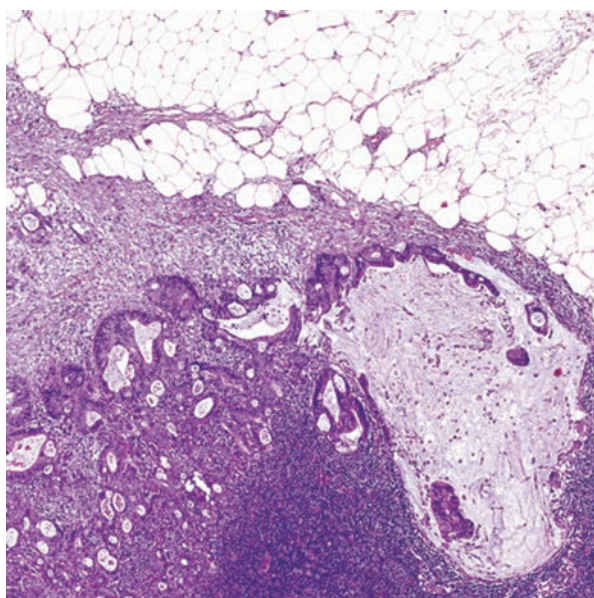

Figure 3. *Example of a crop-mask pair*

53

Разработка и применение искусственного интеллекта
для обнаружения метастазов в лимфатических
узлах при колоректальном раке   Development and application of artificial intelligence for the
detection of the lymph nodes metastases in colorectal cancer

to an incorrect assessment and a significant decrease in the accuracy of the analysis (Fig. 6). Therefore, in parallel with machine learning for detecting metastases in lymph nodes, an algorithm for recognizing artifacts in areas of interest (lymph node tissue) was created. The prediction of artifacts occurred according to a similar logic with tissue segmentation — the model should predict the probability of finding artifacts in places on the image so that a probability map of artifacts can be formed based on the result of predictions.

When solving this problem, the Intersection over Union (IoU) metric was added. It is a simple ratio of the intersection of the predicted and marked area to their union. In many ways, it is similar to the previously described DICE, but it is more penalized for individual poorly predicted examples, which is important to take into account in the context of the task of determining artifacts.

https://en.wikipedia.org/wiki/Jaccard_index

As well as for DICE, below are the values of macro metrics and metrics for the target class, which were obtained as a result of training the artifact segmentation model:

Indicator	Value
Dice macro	0.776
Dice for artifacts	0.552
IoU macro	0.725
IoU for artifacts	0.451

In the work of the artifact segmentation model, there was a tendency to predict small artifact zones well (Fig. 7), but there were shortcomings in the ability of the model to predict large artifacts (Fig. 8).

DISCUSSION

The process of developing a model of computer analysis of digital images in pathology has identified a number of problems that are associated with solving the task. First of all, it is the creation of a database with a large number of full-size digital images (WSI), which is necessary in the development of CPATH algorithms. This task can be performed in pathology laboratories equipped with modern high-performance scanners and possessing a sufficient amount of initial histological material. The volume of studies of lymph nodes in CRC, conducted in the Pathomorphology Unit of the RNMRC of Coloproctology, allowed to create a significant database in terms of volume to obtain the desired results.

The collection of WSI annotations was created by manually marking the studied areas of the image (highlighting different types of tissue or the location of certain types of cells), which was a very time-consuming process requiring appropriate knowledge and qualifications from the pathologist. However, for the machine learning (ML) process in order to create effective analysis algorithms, a large number of WSI with manual annotation is a prerequisite for the CPATH model to learn how to perform the
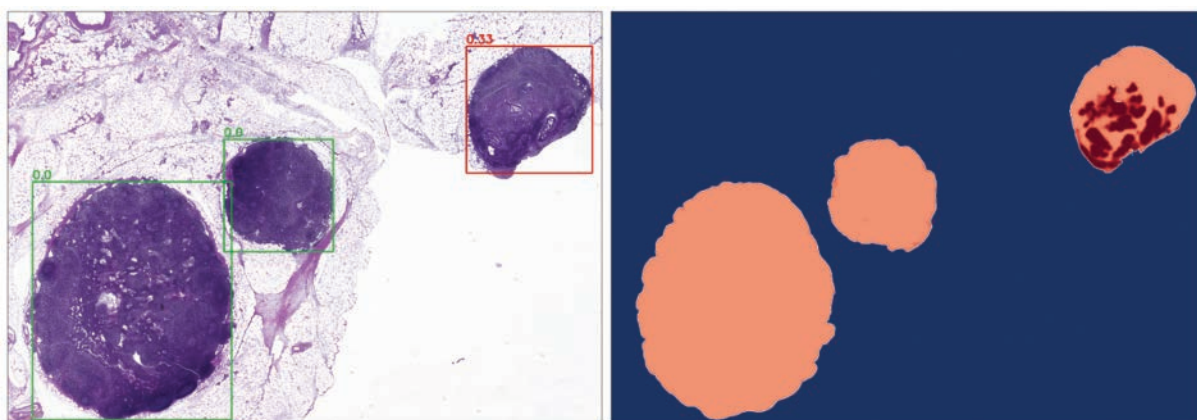

Figure 4. *An example of using the method for determining normal lymph nodes and lymph nodes metastases/replaced tissue (on the left is a markup, on the right is a heat map of prediction probabilities)*
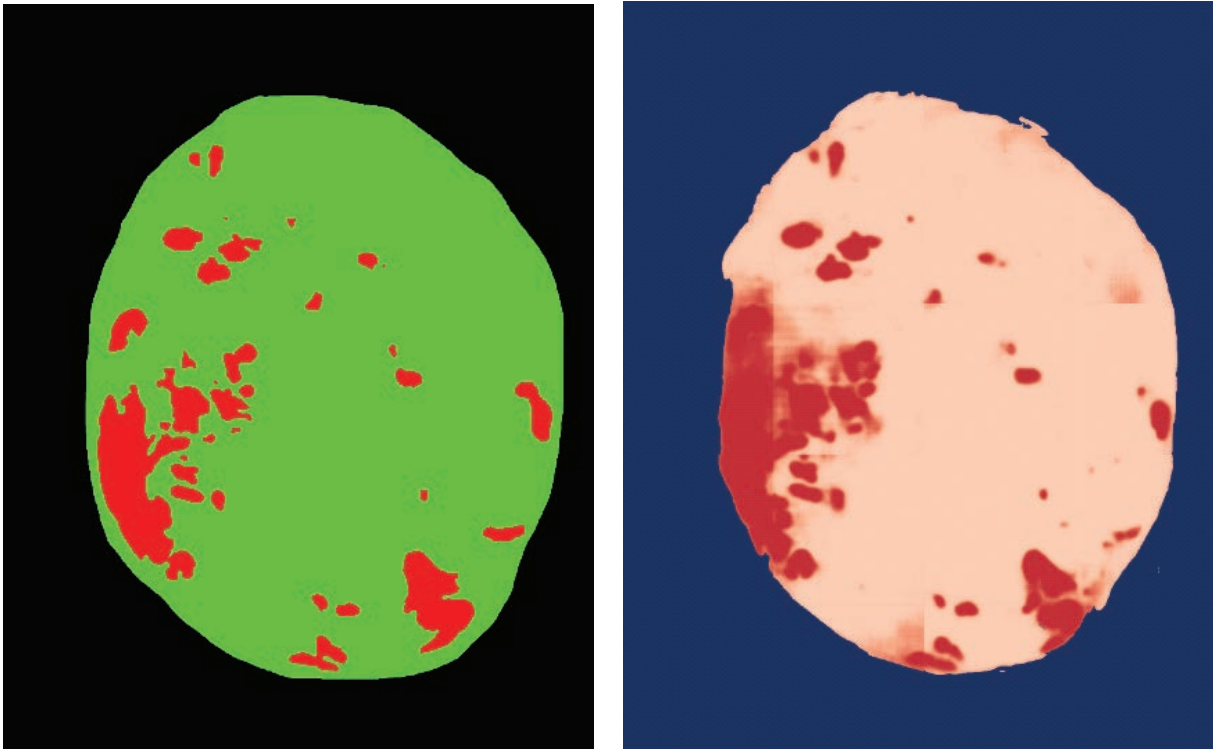
54

Figure 5. *An example of the specificity of the model (on the left — the markup, on the right — the heat map of prediction probabilities)*
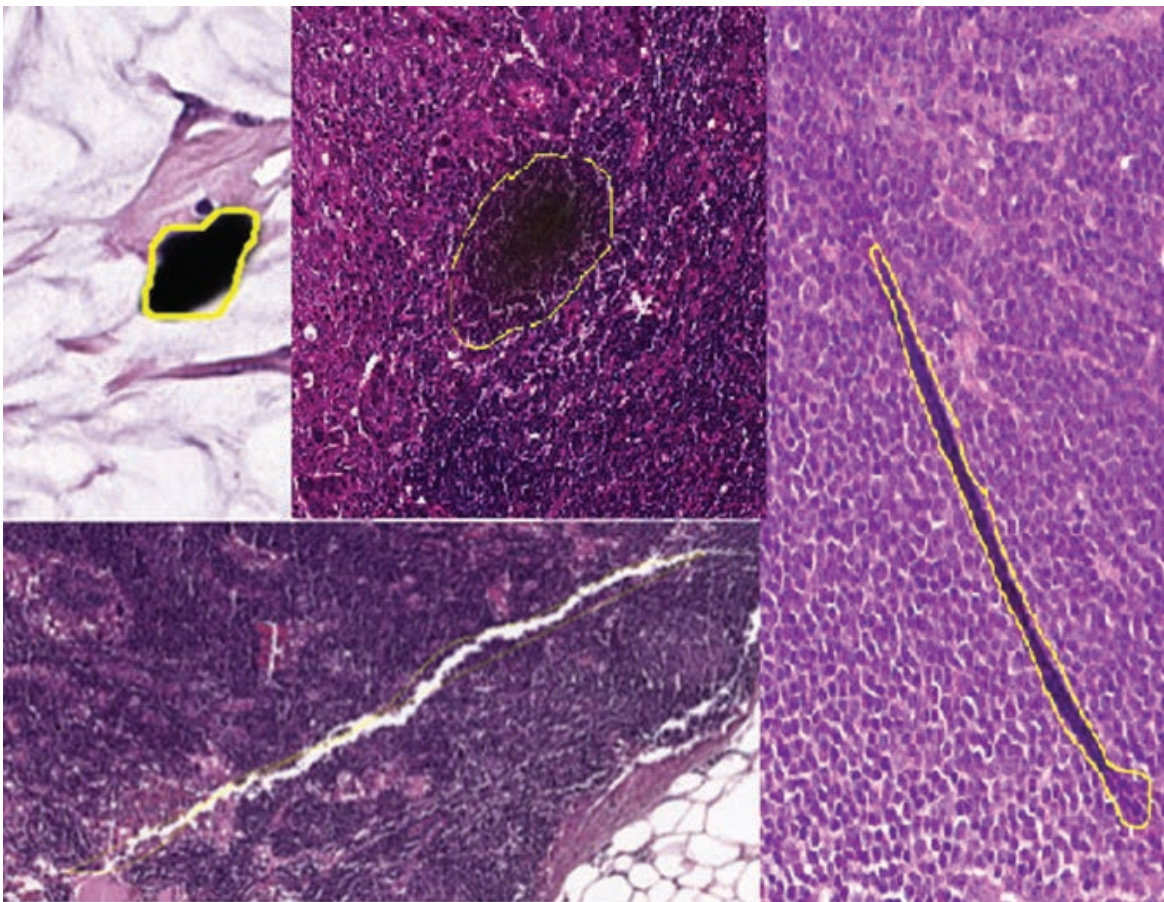
Figure 6. *Examples of detectable artifacts in histoscans*

55

Разработка и применение искусственного интеллекта
для обнаружения метастазов в лимфатических
узлах при колоректальном раке

Development and application of artificial intelligence for the
detection of the lymph nodes metastases in colorectal cancer

task. This dictates the need to develop methods that facilitate the effective creation of these annotations.

In the conducted study, a database with manual markup (annotation) of about 432 files containing WSI 1,000 lymph nodes was created, which made it possible to train an AI model to solve the problem of segmentation of both tissue type and identification of zones of possible artifacts in a digital image.

The tissue type segmentation model developed using the DeepLab V3 algorithm was selected based on experiments based on variation in resolution, model type and image size. According to the presented indicators, the tissue type segmentation model has a high accuracy in determining substitutions in lymph node tissue with a DICE value of about 0.9: macro DICE — 0.923 and substitution DICE — 0.863.

Visual assessment also confirms that the model makes adequate decisions regarding the determination of substitutions/metastases in lymph nodes.

The artifact segmentation model showed significantly lower accuracy (DICE and IoU metrics were 0.776 and 0.725, respectively, for macro; 0.552 and 0.451, respectively, for substitution), which is explained by the complexity of solving the problem due to the obviously more diverse size distribution of target objects and variability of artifact representations. The test results showed that the model is good at detecting small artifacts in digital images of lymph nodes, but at the same time confuses larger objects, especially ruptures, with background zones and light inclusions in lymphoid tissue. The conducted research has shown that the solution of the problem of artifact recognition is an urgent problem, since it allows to avoid errors in diagnosis and, at the same time, to control the quality of digital images, which is a necessary procedure in digital pathology.
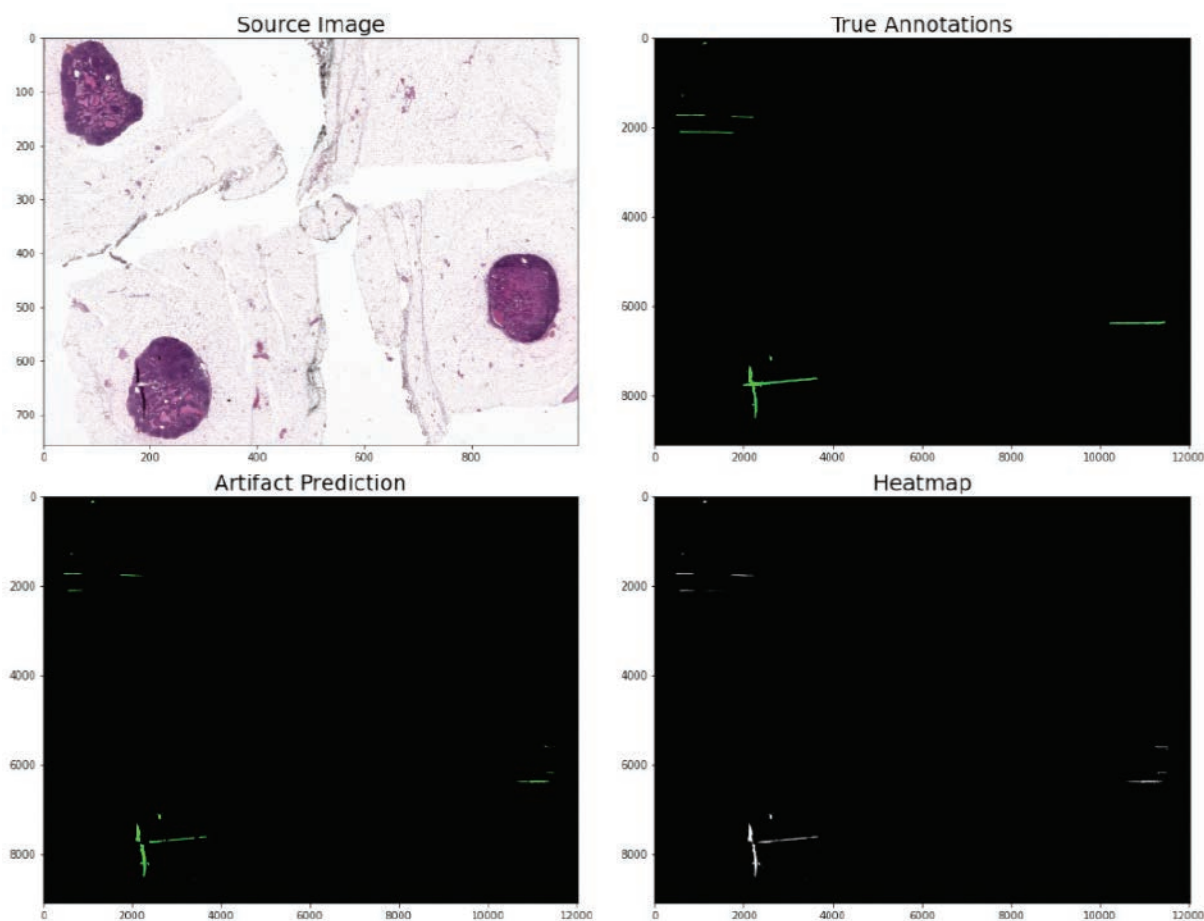

Figure 7. *A good prediction of small artifacts*

56

It should be noted that in the course of the study, a special pre-processing of digital images was developed to implement the training of the model, which allows processing huge WSI without large expenditures of both RAM and video memory.

Testing of the created model was carried out on a group of selected WSI (test sample) from the database we created. To create the final product, it is necessary to test the resulting model on digital images from other sources, since previous studies have shown that CPATH algorithms work optimally with data from the sources on which they were trained, but work worse (sometimes significantly) with data from other datasets [4]. This will make CPATH algorithms as reliable as possible in response to variations that may occur in real practice. An additional approach to the operation of CPATH algorithms with differences between digital images from different data sources is the normalization of images according to a common standard, which, in turn, is one of the main problems in digital pathology.

Nevertheless, the results of the research conducted with the development of CPATH algorithms, which have shown an efficiency comparable to the work of pathologists in performing a number of specific diagnostic tasks, as well as the results of our research can be considered promising. Currently, many technical problems in the development and application of AI in pathology have already been overcome, but its role in practical work has not yet been clearly defined [3, 4, 7]. When discussing the role of AI in pathology, it is important to understand that the work of a pathologist is not only to examine tissue fragments under a microscope. It includes the integration of the diagnostic process, clinical data and the specific circumstances of the patient, followed by the formulation of its conclusion on the results of morphological
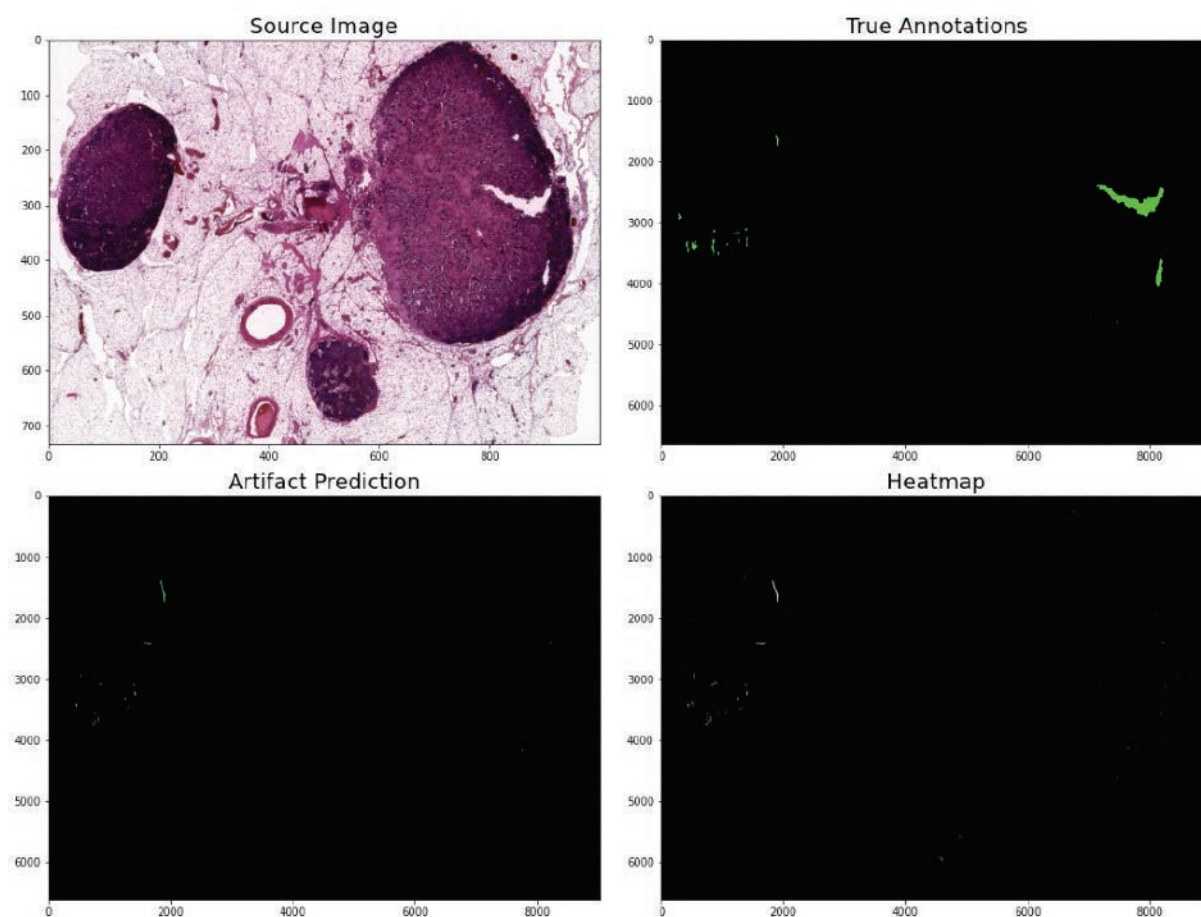
Figure 8. *An example of an error in a large rupture of a lymph node slice*

57

Разработка и применение искусственного интеллекта
для обнаружения метастазов в лимфатических
узлах при колоректальном раке   Development and application of artificial intelligence for the
detection of the lymph nodes metastases in colorectal cancer

examination for both clinicians and the patient. Therefore, AI algorithms are unlikely to be considered as an independent diagnostic tool in the near future. What can be achieved in the near future is the creation of AI algorithms that work together with pathologists, helping to reduce the amount of time-consuming monotonous, but necessary work, in particular, such as the detection of metastases in lymph nodes. The obtained results of our study allow us to speak with some confidence about the possibility of using AI to solve such diagnostic problems in pathology.

CONCLUSION

The results obtained in the study suggest that the developed model is a good basis for the implementation of a full-fledged solution, on the basis of which a system can be developed to assist doctors in finding and evaluating the replacement of tissue structures, detecting artifacts and evaluating the quality of digital images. In addition to the task of collecting sufficiently large sets of annotated WSI, it is necessary to conduct prospective studies to show the importance of AI in pathomorphological diagnostics.

AUTHORS CONTRIBUTION

Concept and design of the study: *Alexey V. Devyatkin, Vladimir V. Rybakov,*
Collection and processing of materials: *Maria A. Ryakhovskaya, Talgat Kh. Makambaev, Daria I. Suslova, Olga A. Maynovskaya*
Statistical processing: *Elena V. Serykh*
Text writing: *Olga A. Maynovskaya, Alexey V. Devyatkin*
Editing: *Sergey I. Achkasov*

INFORMATION ABOUT THE AUTHORS (ORCID)
Olga A. Maynovskaya — 0000-0001-8189-3071
Sergey I. Achkasov — 0000-0001-9294-5447
Alexey V. Devyatkin — 0000-0002-5712-1626
Elena V. Serykh
Vladimir V. Rybakov
Talgat Kh. Makambaev
Daria I. Suslova — 0000-0003-1106-5486
Maria A. Ryakhovskaya — 0000-0002-2243-1317

REFERENCES

1. Ciresan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. *Med Image Comput Comput Assist Interv*. 2013;8150:**4**11–418.

2. Cruz-Roa A, et al. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. *In Proc. SPIE Medical Imaging*. 2014;9041:904103.

3. Chetan L. Srinidhi et al. Deep neural network models for computational histopathology: A survey. *Medical Image Analysis*. 2020. www.elsevier.com/locate/media

4. Jeroen van der Laak et al. Deep learning in histopathology: the path to the clinic. *Nature Medicine*. May 2021;27:775–784.

5. Babak EB, Mitko V, Johannes van Diest P, et al. and the CAMELYON16 Consortium. *JAMA*. 2017;318(22):2199–2210.

6. Litjens G, et al. 1399 h&e-stained sentinel lymph node sections of breast cancer patients: the camelyon dataset. *Giga Science*. 2018; 7, giy065.

7. Bandi P, et al. From detection of individual metastases to classification of lymph node status at the patient level: the camelyon17 challenge. *IEEE Trans Med Imaging*. 2018;*3*8:550–560.

8. Shchaeva S.N. Survival analysis in complicated colorectal cancer. *Oncological coloproctology*. 2017;7(2):20–29. (in Russ.). DOI: 10.17650/2220-3478-2017-7-2-20-29.

9. Fortea-Sanchis C, Martínez-Ramos D, Escrig-Sos J. CUSUM charts in the quality control of colon cancer lymph node analysis: a population-registry study. *World J Surg Oncol*. 2018 Nov 30;16(1):230.

10. Malignant neoplasms in Russia in 2019 (morbidity and mortality). Edited by Kaprin A.D., Starinsky V.V., Shakhzadova A.O./M., Herzen

Moscow State Medical Research Institute — branch of the Federal State Budgetary Institution "NMIC of Radiology" of the Ministry of Health of Russia. 2020; 252 p. (in Russ.).

11. Clinical recommendations "Malignant neoplasms of the colon and rectosigmoid department". 2020. https://cr.minzdrav.gov.ru/schema/396_1 (in Russ.).

12. Clinical recommendations "Rectal cancer". 2020. https://cr.minzdrav.gov.ru/schema/554_2 (in Russ.).

Разработка и применение искусственного интеллекта
для обнаружения метастазов в лимфатических
узлах при колоректальном раке

Development and application of artificial intelligence for the
detection of the lymph nodes metastases in colorectal cancer