Analysis of Thyroid Gland Cytological Images Using Computer Vision

Ilya Lozhkin, Andrey Mironov, Konstantin Zaytsev, Aleksander Garmash,
Boris Shifman, Fatima Abdulkhabirova, Lilia Urusova, Nadezhda Platonova

***Abstract*—The aim of this article is to explore the capabilities of modern neural networks in analyzing cytological whole slide images in svs format in accordance with the Bethesda categorization system. The article presents the results of the proposed neural network models for solving tasks of automatic segmentation and classification of both various types of individual cells and their clusters. For the diagnosis of thyroid cancer using computer vision, the following cell types are identified: Hurthle cells, cells with pseudoinclusions, C-cells, and clusters of cells forming papillary structures, shapeless structures with ordered and unordered cell arrangements. The proposed models have demonstrated their effectiveness in solving tasks related to the intelligent analysis of both whole-slide cytological images and tiled image segmentation. The results obtained for the segmentation of individual cells are as follows: mean Dice coefficient (DC) = 90.9% for pseudoinclusions, DC = 86.2% for Hurthle cells, DC = 90.2% for C-cells. For cell cluster segmentation, the mean Intersection over Union (IoU) is 84%, and DC is 91%. The classification accuracy of cell clusters into three classes is 77.9%.**

***Keywords* – computer vision, deep learning, whole-slide image, thyroid cytology**

I. INTRODUCTION

Processing large images is a complex task compared to manipulating images of conventional sizes, as operations such as analysis, transmission, display, and others require significantly more time. This is particularly evident in fields where working with digitized images occurs in real time, such as medical diagnostics using instrumental methods. In such cases, it is often necessary to handle images or image series of varying volumes. Working with large images is especially relevant in cytological and histological analysis, where the size of a single file in svs format can reach several gigabytes.

Thyroid nodules (TN), being an extremely common medical issue, require further examination upon detection to rule out malignancy [1], [2]. Ultrasound (US) is the primary stage of diagnosis, as it provides initial data on the risk of malignancy and determines the indications for the subsequent step, i.e. fine-needle aspiration (FNA), followed by cytological examination [1], [3], [4].

Cytopathology reports should be presented according to the Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) system, which stratifies specimens into six diagnostic categories, each corresponding to a specific risk of malignancy [5]. Although the currently accepted diagnostic algorithm has proven to be an effective method for clarifying the nature of TN, it is important to note that US and FNA are operator-dependent methods. The quality and speed of these examinations are highly influenced by the experience and qualifications of the specialist performing them. This also makes both examinations labor-intensive, time-consuming, and, in some cases, prone to significant discrepancies when the same lesions are evaluated by different experts. While these issues are partially mitigated through the development of standardized assessment and reporting algorithms by professional communities, integrating software solutions utilizing computer vision for intelligent analysis at these diagnostic stages could help accelerate diagnostics and improve accuracy.

Such a software product extracts informative features and provides the physician with information about significant areas of the image, indicating the predicted disease class and forming a “second opinion” based on diagnostic models.

This study is dedicated to modeling automated intelligent analysis of cytological examination images. The number of publications in this field is significantly lower compared to cardiology, pulmonology, or dermatology, and they are generally unsystematic, addressing only individual features [6]-[10]. This is due to several challenges. Firstly, disease classification requires identifying more than ten informative objects at the cellular level, which is significantly more complex than analyzing a single thyroid nodule in ultrasound imaging [11]. Secondly, working with large images ranging from 3 to 7 GB greatly complicates all data processing operations.

II. LITERATURE REVIEW

Today, various software tools are available that allow users, including those without experience in intelligent image processing, to work with large whole-slide medical images. Some of the most popular ones include QuPath [12]-[15]. These systems provide functions for viewing and annotating images (including whole-slide images) using built-in machine learning models. They enable cell detection and classification, quantitative analysis, pixel classification, integration of a limited set of models, and the creation of custom scripts. Additionally, each system offers unique extra features.

Nevertheless, all the listed tools, as well as others not mentioned here, are designed to address typical tasks common to most histopathological images. However, the capabilities provided by these tools are insufficient for a comprehensive diagnosis of thyroid cytology using computer vision methods. Existing systems do not support the automation of thyroid cytological image classification according to the Bethesda system while accounting for the

identification of multiple informative features within the images.

III. MATERIALS AND METHODS

The objective of this study can be formulated as follows: "The input data for the research includes digitized thyroid cytology images (whole slide images), deep neural network architectures, machine learning models, and specialized software for working with visual data. The task is to conduct research aimed at developing a methodology for the automated extraction of significant features from thyroid cytology images and categorizing these images using computer vision in accordance with the international Bethesda criteria".

3.1. Data preparation

Preprocessing Features

The computer processing of ultrasound and cytological images differs significantly [11], [16], [17].

Ultrasound images and cine loops, ranging in size from several tens of kilobytes to megabytes, are processed using standard methods similar to conventional images. Cine loops containing several dozen frames are split into individual frames, followed by preprocessing, dataset formation in a format suitable for neural network training, and subsequently, network training and testing.

For digitizing cytological slides, whose image sizes can reach several gigabytes in svs format, the Whole Slide Imaging (WSI) technology is used. This method enables the acquisition of high-resolution images with exceptional clarity, which can be viewed at different magnification levels. The dimensions of such images can reach tens of thousands of pixels in both length and width, making their processing fundamentally different from conventional image processing. Whole-slide image analysis requires substantial time and computational resources. Additionally, to improve result accuracy, careful image preprocessing is essential [18].

Figure 1 shows an example of a thyroid cytology image at low magnification, while Figure 2 presents a fragment of the same image at high magnification.

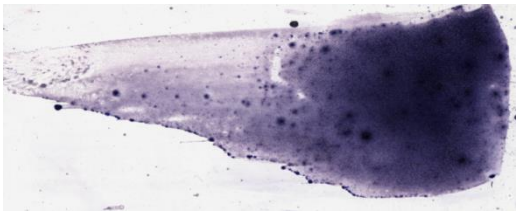

Fig. 1 – Example of a thyroid cytology image at low magnification

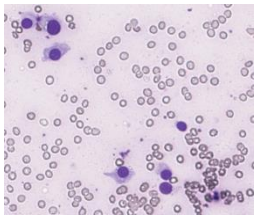

Fig. 2 – Example of a fragment of a cytology image at high magnification

Working with WSI cytology images can be done in two ways: by using specialized tools to process the entire image as a whole or by dividing it into smaller fragments (tiles).

Before splitting a whole-slide image into tiles, the magnification level must be determined. The higher the magnification, the longer the process takes and the more tiles are generated. The obtained tiles undergo preprocessing and are then analyzed by neural network models, which also requires significant time. To speed up processing, a lower magnification level can be chosen, but this may reduce prediction accuracy.

Working with small tiles is similar to processing ultrasound images. The preprocessing of tiles obtained from the original WSI aims to improve algorithm accuracy and includes removing irrelevant areas, resizing, color thresholding, normalization, and other methods, including image augmentation. After preprocessing, tiles that do not contain areas of interest are filtered out, which accelerates intelligent analysis.

Processing WSI without splitting it into tiles is used when the analysis focuses not on individual cells but on their clusters, for example, for counting these clusters. This approach is implemented using tools such as QuPath. In this case, preprocessing includes selecting the image type and configuring stain vector assessment parameters.

These differences highlight that processing small ultrasound images and large thyroid cytology images differs significantly.

When working with QuPath, preprocessing involved selecting the image type (Brightfield H&E) and configuring stain vectors to enhance analysis quality. During the WSI tiling process, some tiles were excluded due to the absence of significant areas, allowing data volume reduction without loss of informativeness. Preprocessing of each tile included resizing, normalization, and preparation for subsequent analysis steps.

For augmenting training sets of tiles, proven methods were used, which had previously demonstrated high efficiency in experiments on training models for thyroid ultrasound image segmentation and classification. These approaches improved model quality and their generalization ability.

Original Dataset and Annotation

The original dataset included digitized images of thyroid cytology studies and class labels. A detailed description of the dataset used in this study is presented in [19], which this article continues.

Ethical Review. The study protocol was reviewed and approved by the local ethics committee of the National Medical Research Center for Endocrinology of the Ministry of Health of Russia (protocol No. 14 dated 25.07.2023).

Based on a list of objects for detection agreed upon with a cytologist, image annotation was performed. Due to the large size of the original whole-slide images, manual annotation of all objects was time-consuming, so the functional capabilities of existing tools were utilized.

Among the objects requiring detection were clusters of cells of various categories. On thyroid cytology images, the number of such clusters can reach several hundred, necessitating their counting. Therefore, QuPath's capabilities were initially tested for detecting clusters on the original WSI images, with the experimental results detailed

in [19]. Based on the outcomes of QuPath's methods for detecting cell clusters, the best-performing method was chosen for annotation. After training and applying the pixel classifier, identified regions were transformed into individual objects with subsequent filtering by area. To export ROIs in PNG format along with corresponding masks, a Groovy script was developed using the built-in editor. The resulting image tiles were then manually categorized (with some being discarded due to irrelevance) and sent for expert review, where the annotations were adjusted if necessary by relocating images to the correct folders.

To accelerate the annotation process and improve its accuracy, annotation functions from the CVAT tool [20] were used for creating masks for high-magnification image tiles. A semi-automatic annotation approach was applied using the pre-trained SAM model [21]: clicking on an area of interest generated predictions, which could be manually adjusted if needed.

QuPath was also used for annotating individual cells. Initially, several images with different cell types were manually annotated. The resulting images, along with their masks, were divided into tiles, exported, and organized into folders before being validated by a cytologist.

At the initial stage, traditional computer vision algorithms utilizing the OpenCV library were employed. Image preprocessing included histogram equalization and Gaussian blur. Then, thresholding was applied to separate cells from the background, followed by contour detection on the resulting masks. However, due to high data noise and color variability in the images, the algorithm's accuracy was low. Further experiments were conducted using QuPath tools, where built-in algorithms such as Random Forest, Linear Regression, K-Means, and MLP were tested for semantic segmentation. While a neural network provided the best results, its accuracy was still insufficient for solving the task.

3.2. Creation of a model

The step-by-step representation of the intelligent analysis of thyroid cytology images is shown in Fig. 3.

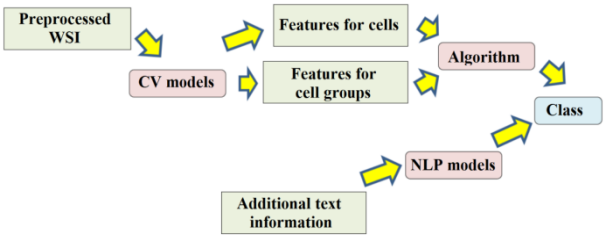

Fig. 3 – Intelligent analysis of thyroid cytology images

The preprocessed whole-slide image of thyroid cytology is fed into several deep learning models that extract various informative features. Based on the predicted features for individual cells and their groups, the digital slide is classified according to algorithms developed and approved in collaboration with an expert, which depend on the type of image. It is important to note that the algorithm for categorizing WSIs according to the Bethesda system using computer vision, developed and agreed with the expert, requires the expert to identify values for certain parameters that influence the final categorization result. Therefore, in the course of practical experiments, it is necessary to assess the applicability of existing approaches for classifying whole-slide images using a single neural network model.

For the localization and classification of cell clusters, an approach was developed that includes the following sequence of automated steps: splitting the whole-slide image into image tiles; semantic segmentation of image tiles; merging segmentation predictions into a whole-slide mask; applying the QuPath pixel classifier for detecting clusters on the whole-slide mask; exporting ROIs (regions with predicted clusters); classification of ROIs. The number of steps is due to the need for precise cluster counting, as segmentation of image tiles alone does not allow this due to the possibility of the same cluster appearing in different tiles.

To extract features related to individual cells, a similar approach was developed: splitting the WSI into image tiles, semantic segmentation of images, contour detection on the resulting masks, and merging all contours of the image into one, considering their position on the original image. In this case, the tile size is much larger than the size of the features, and the number of detected cells is usually large, which led to the decision not to proceed with further steps.

3.3 Applying trained models

To assess the performance of the trained models for semantic segmentation of images, IoU and DC metrics were applied [22]:

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

$$DC = \frac{2|A \cap B|}{|A| + |B|}, \text{ where} \quad (2)$$

The average values of IoU and DC were calculated for datasets consisting of several images.

To assess the performance of the trained models for image classification, the metrics accuracy, precision, recall, and f1-score were applied [23]:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$precision = \frac{TP}{TP + FP} \quad (4)$$

$$recall = \frac{TP}{TP + FN} \quad (5)$$

$$f1\ score = \frac{2 * recall * precision}{recall + precision} \quad (6)$$

The ability of existing approaches to classify whole-slide images (WSIs) at once using a single neural network model was explored in experiments with fine-tuning the CLAM model [24], and detailed results for classifying WSIs into two classes are presented in [19]. The results of training the DeepLabV3+ model [25] with the EfficientNet-B6 encoder [26] for semantic segmentation of tile images to detect cell clusters are also presented there.

The predicted regions of interest containing cell clusters need to be classified into categories. A series of model trainings from the YOLO family [27], [28] were conducted for classification. The experiments began with a balanced training dataset: shapeless structures with ordered cell arrangement (290 images) + shapeless structures with

unordered cell arrangement (290 images) + papillary structures (281 images) = 861 images (72%). The test dataset consisted of 102 + 134 + 95 = 331 images (28%). The best results on the test set for the YOLO models trained on the balanced training dataset are presented in Table 1.

Table 1 — Best results on the test set for YOLO models trained on the balanced training dataset for classifying cell clusters into 3 classes

	yolov8n-cls		yolov8s-cls		yolov8m-cls	
Input image width	224	640	224	640	224	640
$accuracy_{test}$	73,7%	70,4%	73,1%	65,9%	71,9%	66,5%

With such a small training dataset, better results were achieved by models with fewer layers (parameters).

Since there were also labeled and verified images from the first two classes, data was incrementally added to the training set, and the models were fine-tuned until the metrics on the test set improved. The results are presented in Table 2. The size of the maximum side of the input image, to which the images were resized, was 224 pixels.

Table 2 — Best test results of YOLO models trained on datasets with different class balance for classifying cell clusters into 3 classes

Training dataset size, images	Class imbalance	Accuracy on the model test	
		yolov8n-cls	yolov8s-cls
290 + 290 + 281 = 861	Balance	73,7%	73,1%
400 + 400 + 281 = 1081	400 / 281 ≈ 1,4	**76,1%**	73,7%
600 + 600 + 281 = 1481	600 / 281 ≈ 2,1	**76,1%**	74,6%
750 + 750 + 281 = 1781	750 / 281 ≈ 2,7	71,6%	**77,9%**
900 + 900 + 281 = 2081	900 / 281 ≈ 3,2	71,3%	73,1%

Based on the obtained results, it can be concluded that with an increase in the training dataset, the model with a larger number of layers (parameters) starts to perform better.

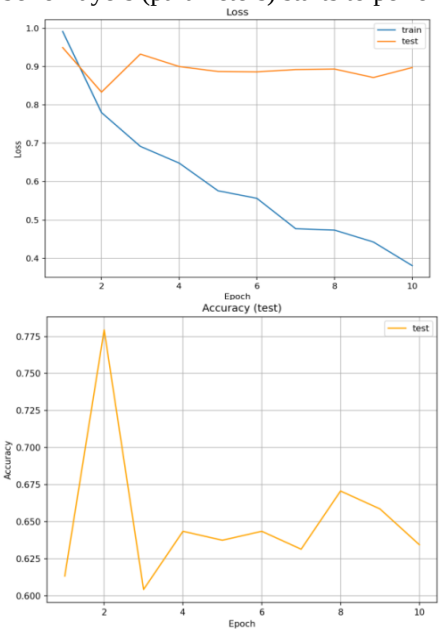

Fig. 4 — Loss function and accuracy metric graphs during fine-tuning and test accuracy of yolov8s-cls for classification of cell clusters into 3 classes

However, when there is a class imbalance of more than 3 times, the model's performance deteriorates. The fine-tuning graphs for the best-performing model, yolov8s-cls, at the moment are shown in Fig. 4. The best test performance when classifying into 3 classes is: accuracy = 77.9%.

After annotating new images with cell clusters from classes with fewer examples, the current model will be fine-tuned, and models with more layers will be trained.

Examples of predictions of cell clusters by the trained models are shown in Fig. 5.

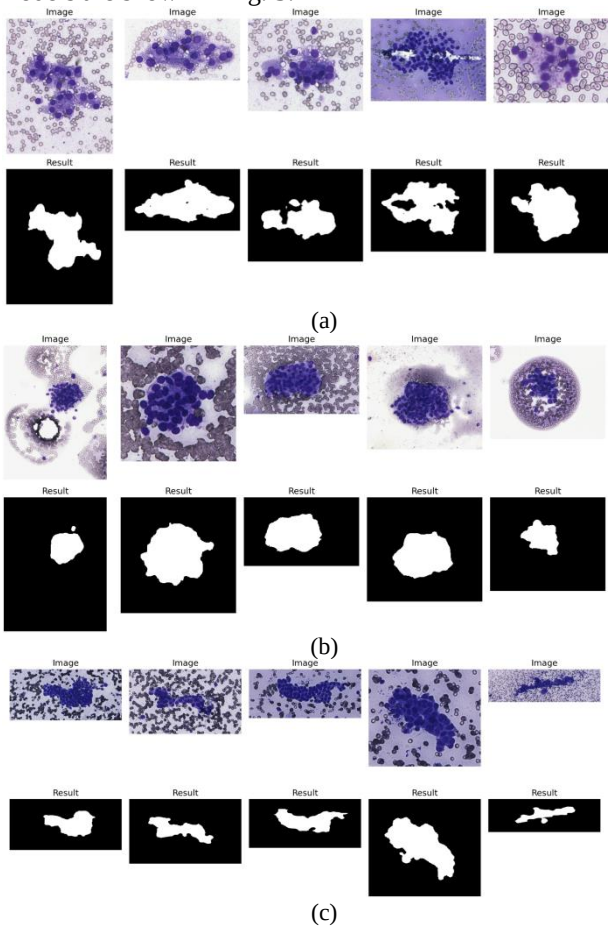

Fig. 5 — Examples of predictions by the trained models of cell clusters: (a) - monolayered sheets of evenly spaced follicular cells; (b) - crowded groups of follicular cells; (c) - papillary structures

Experiments with various algorithms were conducted for semantic segmentation of individual cells. Initially, different computer vision methods were applied to search for informative features, and parameters were selected on several images for more accurate segmentation. However, due to the high variability of the images, the algorithm performed poorly on the validation set. Additionally, the images often contained various types of noise, which the algorithm tried to highlight.

Subsequent experiments were carried out using a pre-trained SAM model. Since it was not trained on medical images, the accuracy was low in automatic mode; the model only detected nuclei and missed other parts of the cells.

Furthermore, experiments were conducted with lighter models specializing in working with medical images. Networks such as Unet, Unet++, LinkNet, and FPN were trained with different encoders: resnet101, resnet152, mobileone_s1, mobilenet_v2, efficientnet-b0, timm-

tf_efficientnet_lite4, timm-efficientnet-b8 [29]. Parameters such as weight_decay, max_lr, epoch, loss function, and optimizer were tuned to achieve the highest accuracy according to the DC metric. The training time per epoch for various models and encoders was also measured to find the optimal balance between "runtime - accuracy."

Based on the experiment results, the Unet++ model with the mobilenet_v2 encoder, JaccardLoss as the loss function, max_lr = 1e-4, epoch = 50, and weight_decay = 1e-3 was selected. The training time was 42 seconds per epoch, and the accuracy was DC_avg = 90.9% for a dataset consisting of 1500 pairs of "original image with pseudo inclusions/without pseudo inclusions — its binary mask".

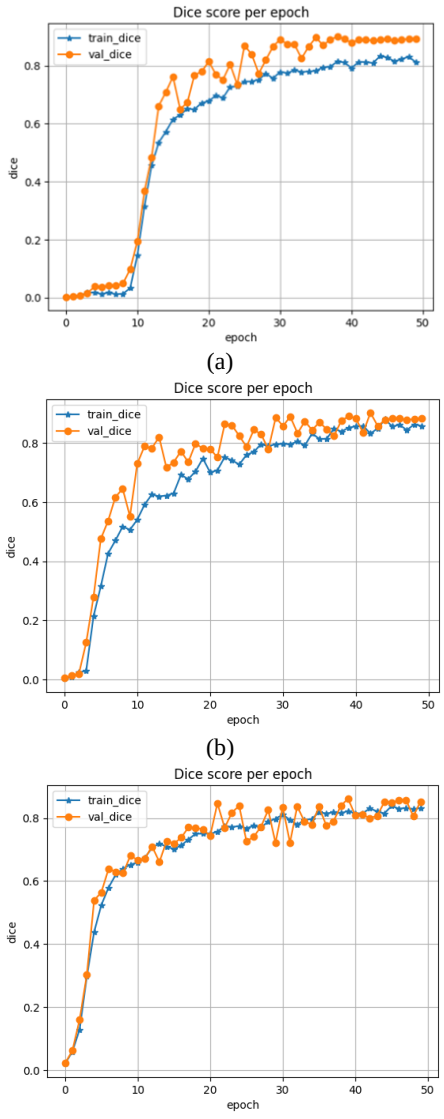

Fig. 6 — Training graphs of Unet++ with mobilenet_v2 encoder for localization (a) - pseudo-inclusions, (b) - C-cells, (c) - Hurthle cells

For the dataset with C-cells, the accuracy was DC_avg = 90.2%. The Hurthle cells were segmented with an accuracy of DC_avg = 86.2%. The metric graphs during training are shown in Fig. 6.

Examples of predictions made by trained models for individual cells when analyzing WSI are presented in Fig. 7.

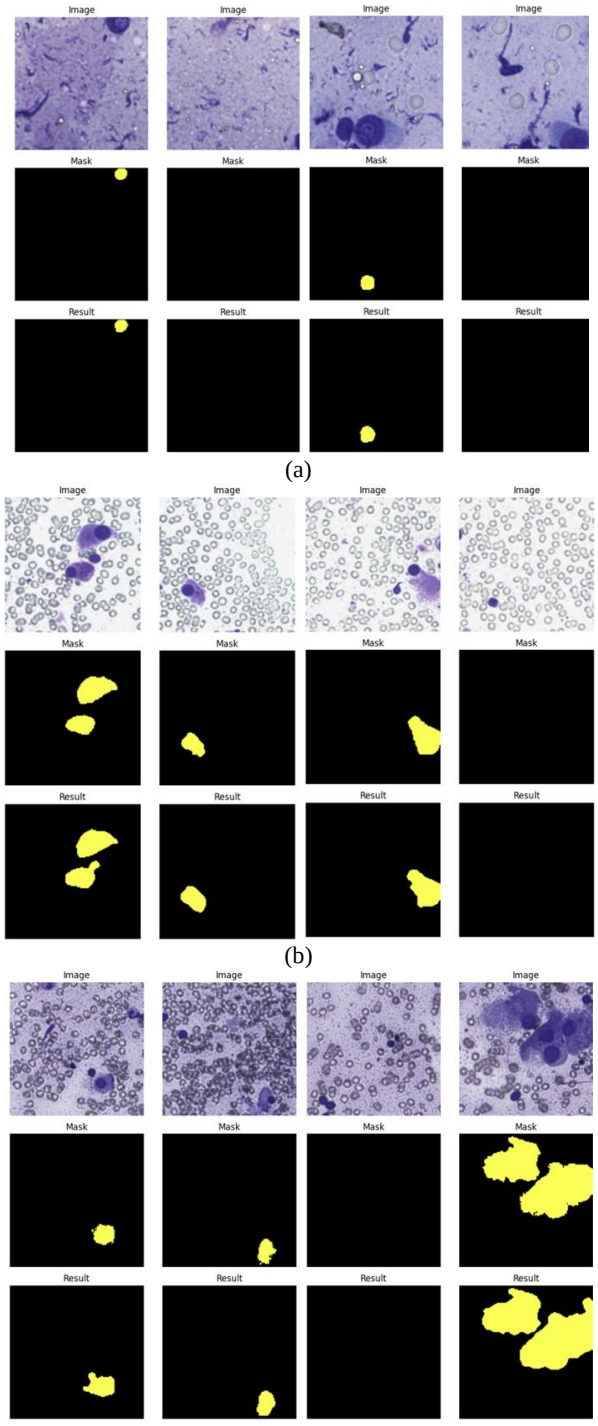

Fig. 7 — Predictions of the model on tiles: (a) — with pseudoinclusions, (b) — with C-cells, (c) — with Hurthle cells. Labels: Image - original image, Mask - annotated mask, Result - model output

IV. RESULTS

In the conducted study, approaches using neural network models for the automatic processing, segmentation, and classification of cytological images of the thyroid gland in accordance with the features necessary for classification according to the Bethesda system were developed. The main results include:

1. Models for analysis at the level of clusters and cells:
   - The average DC for cell segmentation with pseudoinclusions was 90.9%.

- The accuracy on the dataset with C-cells was 90.2%.
- The average DC for Hurthle cells reached 86.2%.
- Cell clusters were segmented with an IoU average of 84% and DC average of 91%.
- Classification of cell clusters into three categories reached an accuracy of 77.9%.

2. Models for analysis at the level of whole slides:

- The ability of existing approaches to classify whole slide images using a single neural network model was assessed: the accuracy of WSI classification into two classes (Bethesda-2 and Bethesda-5—Bethesda-6) was 90%.

3. Effectiveness of working with tiles:

- Dividing whole-slide images into tiles significantly reduced computational costs while maintaining a high level of analysis accuracy.
- Based on preprocessing and augmentation of the tiles, the models improved segmentation and classification results.

4. Practical significance:

- The system built on the proposed models allows for significant reduction of cytologist workload by highlighting informative areas on the images and automatically classifying them.
- The obtained results enable the integration of such systems into clinical practice, improving the quality and speed of thyroid disease diagnosis.

The data obtained confirm the promising application of artificial intelligence for the automatic processing and analysis of cytological thyroid gland images in medical practice.

V. DISCUSSION

The results of the study confirm that the application of modern deep neural network methods and machine learning for the analysis of cytological images of the thyroid gland has significant potential in clinical practice. However, during the work, some limitations were identified, as well as directions for further development.

1. Analysis of Segmentation and Classification Accuracy

- High accuracy metrics (Dice Coefficient and IoU) demonstrate the effectiveness of the proposed approach. However, the classification accuracy of cell clusters (77.9%) requires further improvement. This is due to the variability in the clusters' structures. Therefore, expanding the training dataset with new data is necessary.

2. Features of Whole Slide Images

- Despite the success in tile-based image processing, analyzing whole slide images without prior tiling demonstrated lower accuracy, particularly in classifying small features. This indicates the need for a further optimized approach for working with WSI. A possible solution could be the use of hybrid methods combining tile analysis and full-image context analysis.

3. Data Augmentation

- The use of data augmentation improved the results of segmentation and classification. However, employing more complex approaches, such as Generative Adversarial Networks, could provide more diverse data that accounts for real clinical scenarios.

4. Integration into Clinical Practice

- The developed system already shows potential in reducing the time spent on routine tasks for cytologists. However, for full-scale implementation, further clinical trials are necessary to assess how well the model performs with images from different medical institutions.
- An important issue is still the interpretability of neural network predictions. Explaining model results and integrating them into understandable reports for clinicians is a key aspect for gaining trust in such systems.

5. Ethical and Practical Aspects

- When developing such systems, it is important to consider potential errors that could lead to incorrect diagnoses. Therefore, such a system should be used as a support tool rather than replacing the physician.
- Additionally, patient data protection must be ensured, as the processing of medical images requires strict adherence to privacy legislation.

6. Directions for Future Research

- Investigation of alternative neural network architectures.
- Implementation of multimodal approaches (for example, analyzing text reports alongside image analysis) to improve diagnostic accuracy.
- Studying the possibility of automatic classification of all levels of the Bethesda system, considering finer subcategories.

VI. CONCLUSIONS

The paper presents the results of a study on the capabilities of modern neural networks for analyzing cytological images in the svs format according to the Bethesda categorization system. The characteristics of processing and intelligent analysis of whole-slide cytological images from thyroid fine needle aspiration are outlined. The paper discusses challenges that may arise when working with digitized cytological images, from the annotation process to automated intelligent analysis, as well as solutions that can address these identified issues. A methodology for the automated extraction of significant features from cytological thyroid images and their categorization using computer vision in accordance with the international Bethesda criteria is developed, and training and testing of the trained models are carried out.

REFERENCES

[1] Haugen, B.R., Alexander, E.K., Bible, K.C., Doherty, G.M., Mandel, S.J., Nikiforov, Y.E., Pacini, F., Randolph, G.W., Sawka, A.M., Schlumberger, M., Schuff, K.G., Sherman, S.I., Sosa, J.A., Steward, D.L., Tuttle, R.M., Wartofsky, L., 2016. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid 26, 1–133. https://doi.org/10.1089/thy.2015.0020.

[2] Fridrihsone, I., Strumfa, I., Strumfs, B., Vanags, A., Balodis, D., Jakovlevs, A., Abolins, A., Gardovskis, J., 2018. Thyroid Nodules in Diagnostic Pathology: From Classic Concepts to Innovations, in: Histopathology - An Update. InTech. Available: https://doi.org/10.5772/intechopen.77117.

[3] Russ, G., Bonnema, S.J., Erdogan, M.F., Durante, C., Ngu, R., Leenhardt, L., 2017. European Thyroid Association Guidelines for Ultrasound Malignancy Risk Stratification of Thyroid Nodules in Adults: The EU-TIRADS. Eur. Thyroid J. Available: https://doi.org/10.1159/000478927.

[4] Durante, C., Hegedüs, L., Czarniecka, A., Paschke, R., Russ, G., Schmitt, F., Soares, P., Solymosi, T., Papini, E., 2023. 2023 European Thyroid Association Clinical Practice Guidelines for thyroid nodule management. Eur. Thyroid J. 12. Available: https://doi.org/10.1530/ETJ-23-0067.

[5] The Bethesda System for Reporting Thyroid Cytopathology, 2023. Available: https://doi.org/10.1007/978-3-031-28046-7.

[6] Maleki, S., Zandvakili, A., Gera, S., Khutti, S.D., Gersten, A., Khader, S.N., 2019. Differentiating Noninvasive Follicular Thyroid Neoplasm with Papillary-Like Nuclear Features from Classic Papillary Thyroid Carcinoma: Analysis of Cytomorphologic Descriptions Using a Novel Machine-Learning Approach. J. Pathol. Inform. 10, 29. Available: https://doi.org/10.4103/jpi.jpi_25_19.

[7] Sanyal, P., Mukherjee, T., Barui, S., Das, A., Gangopadhyay, P., 2018. Artificial Intelligence in Cytopathology: A Neural Network to Identify Papillary Carcinoma on Thyroid Fine-Needle Aspiration Cytology Smears. J. Pathol. Inform. 9, 43. Available: https://doi.org/10.4103/jpi.jpi_43_18.

[8] Guan, Q., Wang, Y., Ping, B., Li, D., Du, J., Qin, Y., Lu, H., Wan, X., Xiang, J., 2019. Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study. J. Cancer 10, 4876–4882. Available: https://doi.org/10.7150/jca.28769.

[9] Elliott Range, D.D., Dov, D., Kovalsky, S.Z., Henao, R., Carin, L., Cohen, J., 2020. Application of a machine learning algorithm to predict malignancy in thyroid cytopathology. Cancer Cytopathol. 128, 287–295. Available: https://doi.org/10.1002/cncy.22238.

[10] Kezlarian, B., Lin, O., 2021. Artificial Intelligence in Thyroid Fine Needle Aspiration Biopsies. Acta Cytol. 65, 324–329. Available: https://doi.org/10.1159/000512097.

[11] Lozhkin I.A., Mironov A.M., Dunaev M.E., Zaytsev K.S., 2023. Features of image processing of cytological studies of thyroid gland with the use of computer vision. Procedings of of the XIX International Scientific and Practical Conference "Electronic means and control systems". pp. 305–306.

[12] Bankhead, P., Loughrey, M.B., Fernández, J.A., Dombrowski, Y., McArt, D.G., Dunne, P.D., McQuaid, S., Gray, R.T., Murray, L.J., Coleman, H.G., James, J.A., Salto-Tellez, M., Hamilton, P.W., 2017. QuPath: Open source software for digital pathology image analysis. Sci. Rep. 7, 16878. Available: https://doi.org/10.1038/s41598-017-17204-5.

[13] QuPath: Open source software for digital pathology image analysis. https://qupath.github.io/ (accessed 25 March 2025).

[14] Digital Pathology. Available: https://dpathology.ru/ (accessed 25 March 2025).

[15] Leica Biosystems: Aperio digital pathology software. Available: https://www.leicabiosystems.com/digital-pathology/manage/ (accessed 25 March 2025).

[16] I. Lozhkin, A. Mironov, A.G., 2023. Features of intelligent analysis of images of ultrasound and cytological studies of the thyroid gland. Procedings of of the 8-th International Symposium and Schools for Young Scientists "Physics, Engineering and Technologies for Biomedicine, November 11-15, 2023: Program. pp. 60–61.

[17] Lozhkin I.A., Mironov A.M., Pavlov D.V., Dunaev M.E., 2024. Digital transformation of comprehensive image analysis using computer vision in the diagnosis of thyroid diseases. Proceedings of the International Scientific and Practical Conference "Digital Transformation of Social and Economic Systems". pp. 242–250.

[18] Smith, B., Hermsen, M., Lesser, E., Ravichandar, D., Kremers, W., 2021. Developing image analysis pipelines of whole-slide images: Pre- and post-processing. J. Clin. Transl. Sci. 5, e38. https://doi.org/10.1017/cts.2020.531.

[19] Lozhkin I.A., Zaytsev K.S., Dunaev M.E., Shifman B.M., Abdulkhabirova F.M., 2024. Features of intelligent processing of cytological whole slide images. Int. J. Open Inf. Technol. 12, 84–93.

[20] CVAT: Leading Image & Video Data Annotation Platform. Available: https://www.cvat.ai/ (accessed 25 March 2025).

[21] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment Anything, in: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, pp. 3992–4003. Available: https://doi.org/10.1109/ICCV51070.2023.00371.

[22] Ekin Tiu, 2019. Metrics to Evaluate your Semantic Segmentation Model. Towar. Data Sci. Available: https://medium.com/data-science/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2 (accessed 25 March 2025).

[23] Classification and regression metrics. Available: https://education.yandex.ru/handbook/ml/article/metriki-klassifikacii-i-regressii (accessed 25 March 2025).

[24] Lu, M.Y., Williamson, D.F.K., Chen, T.Y., Chen, R.J., Barbieri, M., Mahmood, F., 2021. Data-efficient and weakly supervised computational pathology on whole-slide images. Nat. Biomed. Eng. 5, 555–570. Available: https://doi.org/10.1038/s41551-020-00682-w.

[25] Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 833–851. Available: https://doi.org/10.1007/978-3-030-01234-2_49.

[26] Tan, M., Le, Q. V., 2019. EfficientNet: Rethinking model scaling for convolutional neural networks, in: 36th International Conference on Machine Learning, ICML 2019.

[27] Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 779–788. Available: https://doi.org/10.1109/CVPR.2016.91.

[28] Terven, J., Córdova-Esparza, D.-M., Romero-González, J.-A., 2023. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. Mach. Learn. Knowl. Extr. 5, 1680–1716. Available: https://doi.org/10.3390/make5040083.

[29] Feng, S.J., Feng, Y., Zhang, X.L., Chen, Y.H., 2023. Deep learning with visual explanations for leakage defect segmentation of metro shield tunnel. Tunn. Undergr. Sp. Technol. 136, 105107. Available: https://doi.org/10.1016/j.tust.2023.105107.

.Article received August 10 2025.

Ilya Lozhkin, National Research Nuclear University MEPhI, lozkin.ilya@gmail.com
Andrey Mironov, National Research Nuclear University MEPhI, andrey_mironov02@mail.ru
Konstantin Zaytsev, National Research Nuclear University MEPhI, KSZaytsev@mephi.ru
Aleksander Garmash, National Research Nuclear University MEPhI, AAGarmash@mephi.ru
Boris Shifman, National Medical Research Centre for Endocrinology, boris-11@mail.ru
Fatima Abdulkhabirova, National Medical Research Centre for Endocrinology, a-fatima@yandex.ru
Lilia Urusova, National Medical Research Centre for Endocrinology, liselivanova89@yandex.ru
Nadezhda Platonova, National Medical Research Centre for Endocrinology, doc-platonova@inbox.ru