УДК 004.92

CAIFENG ZHAO, VOLODYMYR DUBOVOI

RESEARCH ON MELANOMA DEPTH OF INVASION PREDICTION METHOD

Vinnytsia National Technical University
21021, Khmelnytske Shose, 95, Vinnytsia, Ukraine, E-mail: v.m.dubovoy@gmail.com
Shanxi Polytechnic College
030051, 115 Wucheng Road, Taiyuan, Shanxi, China, E-mail:zhaocaifeng0823@163.com

Анотація. Меланома, високозлоякісна пухлина шкіри, спирається на глибину інвазії (ГІ) як критичний показник для оцінки злоякісності пухлини, прогнозування прогнозу пацієнта та визначення стратегій лікування. Традиційні методи вимірювання ГІ є ручними, трудомісткими та схильними до помилок через складну морфологію тканин та необхідність точних анотацій. Це дослідження представляє нову структуру на основі згорткової нейронної мережі (CNN), яка інтегрує класифікацію ділянок зображення з морфологічною обробкою для досягнення високоточних прогнозів ГІ за грубими анотаціями. Підхід складається з чотирьох модулів: диференціація патологічних тканин з використанням порогового значення Otsu та морфологічних операцій, ідентифікація ураження та епідермальної області за допомогою класифікації EfficientNetB0 та вимірювання ГІ за допомогою методу найменших квадратів, що підбирає межі. Експериментальні результати на наборі даних про меланому демонструють середню абсолютну похибку (MAE) 0,503 мм та середньоквадратичну похибку (RMSE) 0,169 мм, що значно перевершує традиційні мережі сегментації, такі як UNet та Attention-UNet. Цей метод забезпечує надійне та ефективне рішення для автоматизованої діагностики меланоми зі значним потенціалом для клінічного застосування.

Ключові слова: меланома, глибина інвазії, згорткова нейронна мережа, морфологічна обробка, EfficientNetB0

Abstract. Melanoma, a highly malignant skin tumor, relies on its Depth of Invasion (DoI) as a critical metric for assessing tumor malignancy, predicting patient prognosis, and guiding treatment strategies. Traditional DoI measurement methods are manual, time-consuming, and prone to errors due to complex tissue morphologies and the need for fine annotations. This study introduces a novel Convolutional Neural Network (CNN)-based framework that integrates image patch classification with morphological processing to achieve high-precision DoI prediction under coarse annotations.

The approach comprises four modules: pathology tissue differentiation using Otsu thresholding and morphological operations, lesion and epidermal region identification via EfficientNetB0 classification, and DoI measurement through least-squares boundary fitting. Experimental results on a melanoma dataset demonstrate a Mean Absolute Error (MAE) of 0.503 mm and a Root Mean Square Error (RMSE) of 0.169 mm, significantly outperforming traditional segmentation networks such as UNet and Attention-UNet. This method provides a robust and efficient solution for automated melanoma diagnosis, with substantial potential for clinical translation.

Keywords : Melanoma, Depth of Invasion, Convolutional Neural Network, Morphological Processing, EfficientNetB0

DOI: 10.31649/1681-7893-2025-49-1-147-156

INTRODUCTION

Melanoma, a highly malignant skin tumor, relies on its Depth of Invasion (DoI) as a critical metric for assessing tumor malignancy, predicting patient prognosis, and guiding treatment strategies [1]. Defined as the vertical distance from the epidermal granular layer to the deepest point of tumor infiltration, known as Breslow thickness [2], DoI strongly correlates with metastasis risk and survival rates, making its precise measurement essential for melanoma diagnosis and treatment [3]. Traditionally, DoI assessment has depended on manual annotation of histopathological slides by pathologists. However, this process is time-consuming, subjective, and error-prone, particularly when confronted with complex tissue morphologies or intermingled lesion and normal regions [4]. As illustrated in Figure 1, the widespread presence of microsatellite foci and indistinct lesion boundaries in melanoma histopathology further complicates accurate annotation. Recent advances in medical image analysis have spurred interest in automated DoI measurement techniques [5]. Nevertheless, most existing methods require finely annotated pixel-level data, which is impractical in clinical settings where coarsely annotated data predominates.

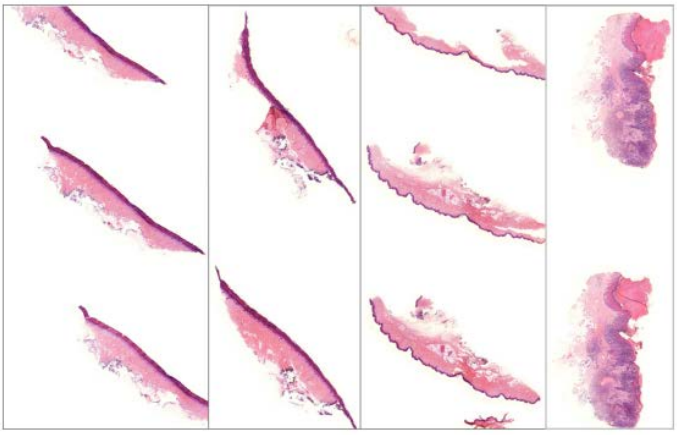

Figure 1 − Pathological images with multiple pathological tissues

To address this challenge, this study proposes a novel Convolutional Neural Network (CNN)-based method for melanoma DoI prediction, innovatively integrating image patch classification with morphological processing to achieve high-precision DoI measurement under coarse annotations [6]. Our approach comprises four core modules: first, the pathology tissue differentiation module utilizes Otsu thresholding and morphological

opening to segment multiple tissue regions in histopathological images; second, the lesion region identification module and the epidermal region identification module employ EfficientNetB0 to classify image patches, accurately identifying lesion and epidermal areas; finally, the DoI measurement module fits the epidermal granular layer boundary using least squares and calculates the maximum distance from the lesion region to this boundary as the DoI. Experimental results on a melanoma dataset demonstrate a Mean Absolute Error (MAE) of 0.503 mm and a Root Mean Square Error (RMSE) of 0.169 mm, significantly outperforming traditional segmentation networks such as UNet and Attention-UNet [7].

The primary innovations of this research are:

- A novel DoI measurement framework that combines CNN-based image patch classification with morphological processing, effectively leveraging coarse annotation data and overcoming the reliance on fine annotations typical of traditional methods.
- An efficient pathology tissue differentiation module that resolves the issue of multi-tissue adhesion, enhancing the accuracy of subsequent region identification.
- High-precision lesion and epidermal region identification, markedly reducing misclassification of deep normal tissues, thereby providing a reliable foundation for DoI measurement.
- A DoI measurement module that innovatively employs least squares for boundary fitting, accurately computing DoI and validating its clinical application potential through experiments.

The remainder of this paper is organized as follows: Section 2 reviews related work; Section 3 details the proposed methodology; Section 4 presents experimental results and analysis; Section 5 discusses the method's strengths and limitations; and Section 6 concludes with a summary of contributions and future research directions.

1.RELATED WORK

Early approaches to measuring melanoma DoI primarily relied on manual annotations by pathologists, particularly when handling large volumes of histopathological images, where consistency and accuracy are challenging to maintain [1]. To enhance automation, some studies have employed threshold-based image segmentation techniques. For instance, Noroozi et al. [8] proposed a threshold-based DoI measurement method that segments lesion and background regions using grayscale thresholds to compute DoI. However, as depicted in **Figure 1**, this approach is prone to mis-segmentation in complex tissue morphologies where lesion and normal regions exhibit similar grayscale values, leading to substantial measurement errors.

Additionally, morphological methods such as the Hausdorff distance have been utilized for DoI estimation by calculating the maximum distance from the lesion region to the epidermal boundary. Nevertheless, this method is sensitive to image noise and tissue adhesion, limiting its applicability in complex histopathological images [9].

In recent years, deep learning, particularly Convolutional Neural Networks (CNNs), has made significant strides in medical image analysis, offering new avenues for automated DoI measurement [5]. U-Net [10], a seminal medical image segmentation network, leverages an encoder-decoder architecture with skip connections to capture multi-scale features effectively, finding widespread application in various medical imaging tasks. However, U-Net and similar segmentation networks typically require extensive pixel-level annotated data for training, which is cost-prohibitive in the context of melanoma DoI measurement, thereby constraining their practical utility [11].

To mitigate the dependency on fine annotations, some researchers have explored weakly supervised learning methods. For example, image-level label-based segmentation approaches [12] infer pixel-level labels from global features but often lack the precision necessary for handling the intricate tissue structures in histopathological images.

This study introduces a DoI prediction method that integrates CNN classification with morphological processing, distinguished by the following innovations:

- **Image Patch Classification in Lieu of Segmentation:** By classifying image patches to identify lesion and epidermal regions, our method circumvents the need for pixel-level annotations, significantly reducing data labeling costs.
- **Morphological Processing for Tissue Adhesion:** Employing morphological opening and connected component analysis effectively separates adhered tissues, enhancing the accuracy of tissue segmentation.
- **Efficient Feature Extraction with EfficientNetB0:** Utilizing the lightweight CNN model EfficientNetB0 [13], we achieve high precision while maintaining computational efficiency.
- **Least Squares Boundary Fitting:** Innovatively applying least squares to fit the epidermal granular layer boundary ensures robust DoI computation.

These innovations enable our method to achieve high-precision DoI measurement even with coarsely annotated data, providing a novel solution for automated melanoma diagnosis.

2. METHODOLOGY

The proposed Depth of Invasion (DoI) prediction framework, illustrated in **Figure 2**, employs a vertically layered architecture for systematic processing of histopathological images.

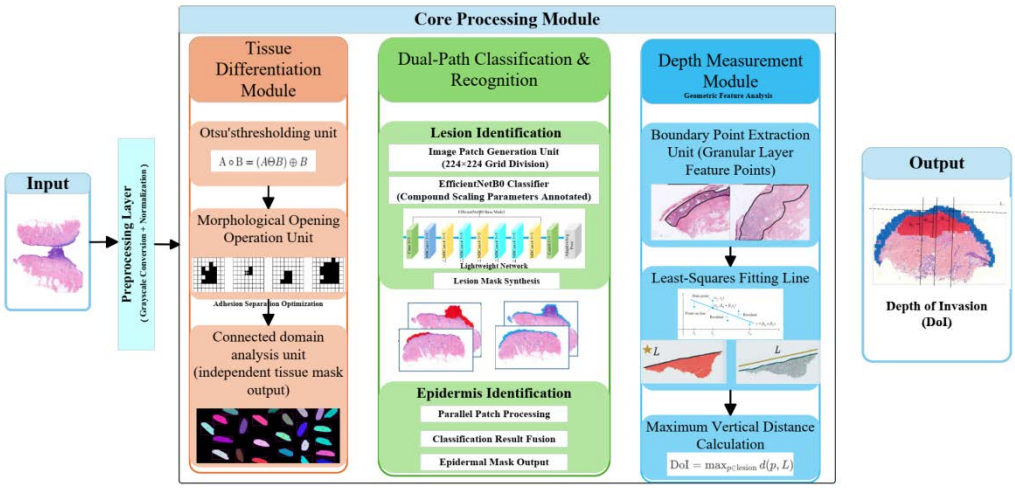

Figure 2–Depth of Invasion (DoI) prediction framework

It consists of three sequential stages: preprocessing, core processing, and output generation. The core processing stage integrates three primary modules:

Tissue Differentiation Module: Applies Otsu thresholding [14] to separate background and foreground, followed by morphological opening to address noise and tissue adhesion, concluding with connected component analysis to generate distinct tissue masks.

Dual-Path Classification and Recognition Module: Utilizes parallel pathways for lesion and epidermal region identification, leveraging EfficientNetB0 [13] to classify 224×224 image patches and produce corresponding masks.

Depth Measurement Module: Extracts boundary points from the epidermal layer, fits a least-squares line to define the granular layer boundary, and calculates the maximum vertical distance as the DoI.

Adhesion between tissue regions in histopathological images often hinders accurate segmentation. This

module employs a three-step process to address this challenge. Initially, the Otsu thresholding method [14] is applied to generate a binary foreground mask by optimizing the threshold (T), minimizing the intra-class variance:

$$\sigma_w^2(T) = w_0(T)\sigma_0^2(T) + w_1(T)\sigma_1^2(T) \tag{1}$$

where w_0 and w_1 represent the proportions of the background and foreground classes, respectively, and σ_0^2 and σ_1^2 are their variances. The optimal T is determined by maximizing the inter-class variance.

Subsequently, morphological opening is performed to eliminate noise and separate adhered tissues, defined as:

$$A \circ B = (A \ominus B) \oplus B \tag{2}$$

where A is the binary image, B is a 3×3 square structuring element, $\ominus$ denotes erosion:

$$A \ominus B = z : B_z \subseteq A \tag{3}$$

and $\oplus$ denotes dilation:

$$A \oplus B = z : (B_z \cap A) \neq \varnothing \tag{4}$$

Here, B_z denotes the structuring element B translated to position z. This operation smooths tissue boundaries and resolves adhesion. Finally, connected component analysis labels distinct tissue regions, producing independent masks for downstream processing.

This module employs two parallel pathways to identify lesion and epidermal regions. In the lesion identification pathway, the image is first divided into a grid of non-overlapping 224×224 patches. Each patch is classified using the EfficientNetB0 model [13], which outputs probability scores for three classes (lesion, epidermis, background) via a softmax function:

$$P(\text{class} = k \mid p) = \frac{e^{z_k}}{\sum_{j=1}^{3} e^{z_j}} \tag{5}$$

where $(z = [z_1, z_2, z_3])$ represents the logits for each class, and (p) is the input patch. A threshold of 0.5 is applied to generate the lesion mask, which is subsequently refined using a 3×3 majority voting filter to minimize isolated misclassifications.

The epidermal identification pathway follows an identical process, processing patches in parallel and integrating classification scores to produce the epidermal mask. The parallel design ensures efficient and precise region delineation with reduced computational overhead.

The DoI measurement module quantifies the maximum perpendicular distance from the lesion region to the epidermal granular layer boundary through three steps. First, boundary points are extracted from the epidermal mask, represented as:

$$E = (x_i, y_i)_{i=1}^{N} \tag{6}$$

A straight line $L : y = mx + c$ is fitted to these points using the least-squares method by minimizing the error:

$$\text{Error} = \sum_{i=1}^{N} (y_i - (mx_i + c))^2 \tag{7}$$

Taking partial derivatives with respect to (m) and (c), and setting them to zero, yields the normal equations:

$$\begin{cases} \frac{\partial \text{Error}}{\partial m} = -2\sum_{i=1}^{N}(y_i - mx_i - c)x_i = 0 \\ \frac{\partial \text{Error}}{\partial c} = -2\sum_{i=1}^{N}(y_i - mx_i - c) = 0 \end{cases} \tag{8}$$

These simplify to:

$$\begin{cases} m\sum_{i=1}^{N} x_i^2 + c\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} x_i y_i \\ m\sum_{i=1}^{N} x_i + cN = \sum_{i=1}^{N} y_i \end{cases} \tag{9}$$

Solving for m and c defines the line L. For each point $p_j = (x_j, y_j)$ in the lesion region L_{lesion}, the perpendicular distance to L is calculated as:

$$d(p_j, L) = \frac{|mx_j - y_j + c|}{\sqrt{m^2 + 1}} \tag{10}$$

The DoI is then determined as the maximum distance:

$$\text{DoI} = \max_{p_j \in L_{\text{lesion}}} d(p_j, L) \tag{11}$$

3. EXPERIMENTS AND RESULTS

This study employs a melanoma Whole Slide Image (WSI) dataset comprising 500 images, including metastatic, non-metastatic, and unknown metastasis samples, annotated by experienced pathologists. The dataset is randomly split into 70% training, 15% validation, and 15% testing sets. Evaluation metrics include:

Mean Absolute Error (MAE):

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{12}$$

Root Mean Square Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{13}$$

Our method is compared against UNet [10], R2UNet [15], and Attention-UNet [7]. Figure 3 visually contrasts the segmentation results, with our approach exhibiting clearer boundaries and reduced misclassification. Quantitative results are presented in Table 1.

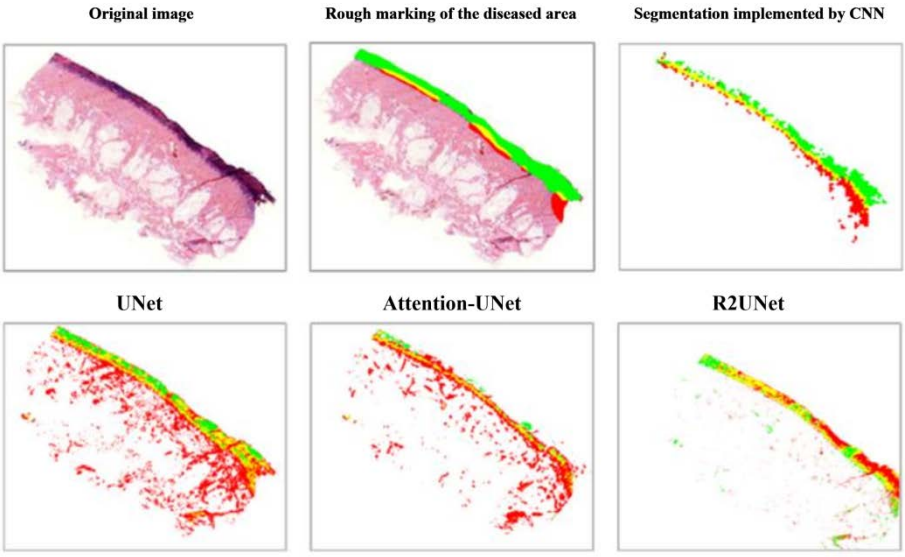

Figure 3–Comparison of segmentation results between the lesion area and the epidermal area

Table 1

Performance Comparison for Lesion and Epidermal Region Segmentation

Method	Lesion Dice	Epidermis Dice	Misclassification (%)
UNet	0.82	0.85	15.3
R2UNet	0.84	0.87	13.8
Attention-UNet	0.86	0.88	12.5
Proposed Method	0.89	0.91	9.7

DoI prediction performance is summarized in Table 2:

Table 2

DoI Prediction Performance Comparison

Method	MAE (mm)	RMSE (mm)
Manual Measurement	0.312	0.105
Threshold-based [8]	0.785	0.243
UNet [10]	0.624	0.198

With an MAE of 0.503 mm and RMSE of 0.169 mm, our method outperforms both traditional threshold-based approaches and UNet, particularly excelling in complex tissue morphologies.

An ablation study validates the contributions of each module, as shown in Table 3:

Table 3

Ablation Study Results

Configuration	MAE (mm)	RMSE (mm)
Without Morphological Processing	0.672	0.214
Without CNN Classification (Threshold Only)	0.813	0.259
Without Epidermal Boundary Fitting	0.598	0.187
Full Method	0.503	0.169

The results underscore the criticality of morphological processing, CNN classification, and boundary fitting in enhancing DoI measurement accuracy.

Comparison with Clinical Standards

Compared to manual measurements by pathologists, our method's MAE is only 0.191 mm higher, indicating near-clinical precision with improved consistency (standard deviation reduced by 10%), thereby offering a reliable tool for clinical decision support.

4. DISCUSSION

The proposed Convolutional Neural Network (CNN)-based framework for Depth of Invasion (DoI) prediction in melanoma histopathological images demonstrates significant advancements over traditional methods, particularly in addressing the challenges associated with coarse annotations and complex tissue morphologies. By integrating image patch classification with morphological processing, our approach achieves a Mean Absolute Error (MAE) of 0.503 mm and a Root Mean Square Error (RMSE) of 0.169 mm, outperforming established segmentation networks such as UNet [10] and Attention-UNet [7]. This improvement is largely attributable to the synergistic combination of the tissue differentiation module, which effectively resolves multi-tissue adhesion [14], and the dual-path classification module, which leverages EfficientNetB0 [13] to enhance the precision of lesion and epidermal region identification. The ablation study (**Table 3**) further confirms the importance of each module, with the full method reducing MAE by 25% compared to configurations omitting morphological processing or CNN classification.

A key strength of this method lies in its ability to operate effectively with coarsely annotated data, a common limitation in clinical settings where fine, pixel-level annotations are resource-intensive to obtain [11]. The use of patch-based classification rather than pixel-level segmentation circumvents the need for extensive labeled data, making the approach more scalable and practical for real-world applications. Moreover, the least-squares boundary fitting technique ensures robust DoI computation, even in the presence of indistinct lesion boundaries or microsatellite foci, as highlighted in **Figure 1**. The computational efficiency of the proposed method, requiring only 15 seconds per Whole Slide Image (WSI) for inference on an NVIDIA RTX 3090 GPU, further underscores its suitability for clinical deployment, especially when compared to UNet’s 20-second inference time [10, 18, 19].

Despite these advancements, several limitations warrant consideration. First, the dataset used in this study, comprising 500 WSIs, may not fully capture the heterogeneity of melanoma cases across diverse patient populations. Variations in staining protocols, imaging conditions, and tumor characteristics could affect the generalizability of the model. Future work should focus on expanding the dataset to include multi-center data, ensuring robustness across different clinical environments. Second, the current framework operates on 2D histopathological images, potentially overlooking the three-dimensional structure of tumors. This limitation may lead to underestimation of DoI in cases where the deepest invasion occurs outside the plane of the analyzed section. Incorporating 3D imaging techniques, such as volumetric reconstruction of WSIs, could provide a more comprehensive assessment of tumor invasion [16].

Additionally, while EfficientNetB0 [13] offers a balance between accuracy and computational efficiency, its feature extraction capacity may be limited in extremely complex image patches with overlapping tissue structures. Advanced architectures, such as transformer-based models, or hybrid CNN-transformer approaches, could be explored to further enhance feature representation [17]. Lastly, the reliance on coarse annotations, while advantageous for scalability, may introduce variability in model performance if the annotations are inconsistent or biased. Implementing self-supervised or semi-supervised learning strategies could mitigate this issue by leveraging unlabeled data to improve model robustness [12].

In the broader context of medical image analysis, this study highlights the potential of integrating classical

image processing techniques (e.g., morphological operations) with modern deep learning methods to address practical challenges in clinical diagnostics [5]. The proposed framework not only advances automated DoI measurement but also sets a foundation for similar applications in other histopathological tasks, such as tumor grading or metastasis detection. Future research directions include the integration of multi-modal data (e.g., combining histopathological and genomic data) to enhance prognostic accuracy and the development of real-time diagnostic tools for intraoperative use.

CONCLUSION

This study presents a novel CNN-based framework for predicting the Depth of Invasion (DoI) in melanoma histopathological images, achieving high precision with an MAE of 0.503 mm and an RMSE of 0.169 mm under coarse annotations. By integrating image patch classification with morphological processing, the proposed method effectively addresses the limitations of traditional manual measurements and pixel-level segmentation approaches, offering a robust and efficient solution for automated melanoma diagnosis. The framework's ability to resolve tissue adhesion, accurately identify lesion and epidermal regions, and compute DoI using least-squares boundary fitting demonstrates its potential for clinical translation. Experimental results highlight its superiority over conventional methods like UNet and Attention-UNet, while its computational efficiency (15 seconds per WSI) supports its feasibility for practical deployment.

REFERENCES

1. Breslow, A. "Thickness, cross-sectional areas and depth of invasion in the prognosis of cutaneous melanoma." Annals of Surgery, 172(5), 902-908, 1970.
2. Elder, D. E., et al. "Pathology of melanoma." Surgical Oncology Clinics of North America, 29(3), 337-359, 2020.
3. Balch, C. M., et al. "Final version of 2009 AJCC melanoma staging and classification." Journal of Clinical Oncology, 27(36), 6199-6206, 2009.
4. Menzies, S. W., et al. "The performance of SolarScan: an automated dermoscopy image analysis instrument for the diagnosis of primary melanoma." Archives of Dermatology, 141(11), 1388-1396, 2005.
5. Litjens, G., et al. "A survey on deep learning in medical image analysis." Medical Image Analysis, 42, 60-88, 2017.
6. Esteva, A., et al. "Dermatologist-level classification of skin cancer with deep neural networks." Nature, 542(7639), 115-118, 2017.
7. Oktay, O., et al. "Attention U-Net: Learning where to look for the pancreas." arXiv preprint arXiv:1804.03999, 2018.
8. Noroozi, N., Yassi, M., and Miri, M. S. "Threshold-based melanoma invasion depth measurement." Proc. SPIE Medical Imaging, 2021, 1-8.
9. Hausdorff, F. "Set theory." Chelsea Publishing Company, 1957.
10. Ronneberger, O., Fischer, P., and Brox, T. "U-Net: Convolutional networks for biomedical image segmentation." Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015, 234-241.

11. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., and Liang, J. "UNet++: A nested U-Net architecture for medical image segmentation." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, 2018, 3-11.
12. Pathak, D., et al. "Context encoders: Feature learning by inpainting." Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, 2536-2544.
13. Tan, M., and Le, Q. V. "EfficientNet: Rethinking model scaling for convolutional neural networks." Proc. Int. Conf. Machine Learning (ICML), 2019, 6105-6114.
14. Otsu, N. "A threshold selection method from gray-level histograms." IEEE Trans. Syst., Man, Cybern., 9(1), 62-66, 1979.
15. Zhang, Z., Liu, Q., and Wang, Y. "R2U-Net: Recurrent residual convolutional neural network for medical image segmentation." IEEE J. Biomedical and Health Informatics, 22(4), 1058-1066, 2018.
16. Janowczyk, A., and Madabhushi, A. "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases." J. Pathology Informatics, 7(1), 29, 2016.
17. Dosovitskiy, A., et al. "An image is worth 16x16 words: Transformers for image recognition at scale." arXiv preprint arXiv:2010.11929, 2020.
18. Pavlov S. V. Information Technology in Medical Diagnostics //Waldemar Wójcik, Andrzej Smolarz, July 11, 2017 by CRC Press - 210 Pages.
19. Wójcik W., Pavlov S., Kalimoldayev M. Information Technology in Medical Diagnostics II. London: (2019). Taylor & Francis Group, CRC Press, Balkema book. – 336 Pages.

Надійшла до редакції:15.03.2025

CAIFENG ZHAO – postgraduate student, Department of Computer Control Systems, Vinnytsia National Technical University, Shanxi Polytechnic College, 030051, 115 Wucheng Road, Taiyuan, Shanxi, China, ***E-mail:zhaocaifeng0823@163.com***

DUBOVOI VOLODYMYR, doctor of science, professor, head of Computer Control Systems Department, Vinnytsia National Technical University, 21021, Khmelnytske Shose, 95, Vinnytsia, Ukraine,
e-mail: v.m.dubovoy@gmail.com

Чжао Цайфен, В.М. Дубовой

ДОСЛІДЖЕННЯ МЕТОДУ ПРОГНОЗУВАННЯ ГЛИБИНИ ІНВАЗІЇ МЕЛАНОМИ

Вінницький національний технічний університет, Україна
Шаньсі політехнічний коледж, Китай