



<http://dx.doi.org/10.35596/1729-7648-2025-23-6-103-109>

УДК 004.8: 616.441

МНОГОМАСШТАБНАЯ НЕЙРОННАЯ СЕТЬ ДЛЯ КЛАССИФИКАЦИИ ФРАГМЕНТОВ ГИСТОЛОГИЧЕСКИХ ИЗОБРАЖЕНИЙ

А. А. КОСАРЕВА

*Белорусский государственный университет информатики и радиоэлектроники
(Минск, Республика Беларусь)*

*Объединенный институт проблем информатики Национальной академии наук Беларуси
(Минск, Республика Беларусь)*

Аннотация. В статье представлена архитектура многомасштабной нейронной сети с несколькими входами, предназначенная для одновременной обработки групп фрагментов гистологических изображений, полученных при различных увеличениях. Предлагаемая модель интегрирует признаки на уровне скрытых слоев, что позволяет эффективно объединять информацию о тканевой структуре на разных масштабах детализации. Экспериментальные результаты показали, что применение трех уровней увеличения по одной стороне фрагмента (98,58; 197,16 и 394,32 мкм) обеспечивает оптимальный баланс между информативностью входных данных и устойчивостью модели. Использование архитектуры позволило повысить среднее значение F1-меры на 5 % по сравнению с одношкальным подходом, достигнув величины $0,8962 \pm 0,0508$, а в отдельных запусках – 0,9697. Наблюдаемое стандартное отклонение обусловлено не нестабильностью модели, а естественной вариабельностью медицинских данных в ходе формирования обучающих выборок. Полученные результаты подтверждают перспективность многомасштабного анализа для задач цифровой патологии и демонстрируют потенциал предложенного решения в качестве средства автоматизированного выделения подозрительных участков на полнослайдовых изображениях.

Ключевые слова: многомасштабная нейронная сеть, глубокое обучение, гистопатологические изображения, полнослайдовые изображения, искусственный интеллект в медицине.

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Для цитирования. Косарева, А. А. Многомасштабная нейронная сеть для классификации фрагментов гистологических изображений / А. А. Косарева // Доклады БГУИР. 2025. Т. 23, № 6. С. 103–109. <http://dx.doi.org/10.35596/1729-7648-2025-23-6-103-109>.

MULTI-SCALE NEURAL NETWORK FOR CLASSIFICATION OF HISTOLOGICAL IMAGE FRAGMENTS

ALEKSANDRA A. KOSAREVA

*Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)
The United Institute of Informatics Problems of the National Academy of Sciences of Belarus
(Minsk, Republic of Belarus)*

Abstract. This paper presents the architecture of a multi-scale, multi-input neural network for simultaneously processing groups of histological image fragments acquired at different magnifications. The proposed model integrates features at the hidden layer level, which allows for efficient fusion of tissue structure information at different detail scales. Experimental results have shown that applying three magnification levels on one side of a fragment (98.58; 197.16, and 394.32 μm) provides an optimal balance between the information content of the input data and the stability of the model. Using this architecture, the average F1-score increased by 5 % compared to a single-scale approach, reaching a value of 0.8962 ± 0.0508 , and in some runs – 0.9697. The observed standard deviation is due to the natural variability of medical data during the formation of training samples, rather than

to model instability. The obtained results confirm the promise of multiscale analysis for digital pathology tasks and demonstrate the potential of the proposed solution as a means of automated detection of suspicious areas in full-slide images.

Keywords: multi-scale neural network, deep learning, histopathological images, full-slide images, artificial intelligence in medicine.

Conflict of interests. The author declares that there is no conflict of interests.

For citation. Kosareva A. A. (2025) Multi-Scale Neural Network for Classification of Histological Image Fragments. *Doklady BGUIR*. 23 (6), 103–109. <http://dx.doi.org/10.35596/1729-7648-2025-23-6-103-109> (in Russian).

Введение

При анализе гистологических изображений диагностически значимые признаки проявляются на разных уровнях увеличения. Так, изображения с низким увеличением позволяют оценить общую архитектуру тканей и пространственное расположение патологических очагов, тогда как высокое увеличение дает возможность рассмотреть детали клеточного строения и морфологические особенности отдельных структур. Использование только одного масштаба может привести к утрате важной информации и ухудшению качества анализа.

Современные исследования в области цифровой патологии демонстрируют значительный прогресс в применении многомасштабных нейронных сетей для анализа гистопатологических изображений. Основная сложность заключается в высокой вариабельности морфологических признаков, перекрытии клеток и неоднородности окрашивания, что делает традиционные подходы недостаточно эффективными. Многомасштабные архитектуры позволяют преодолеть эти ограничения за счет интеграции информации из различных уровней детализации.

Одно из ключевых направлений – использование многомасштабного слияния признаков для повышения точности обнаружения малых объектов. В [1] показано, что комбинация алгоритма SSD с добавлением многомасштабного слияния признаков позволяет значительно улучшить детектирование мелких и средних объектов в условиях изменчивого освещения и угла обзора. Несмотря на уменьшение скорости обработки по сравнению с базовой версией SSD, достигается оптимальный баланс между точностью и производительностью, превосходя алгоритм Faster R-CNN по скорости при сохранении высокой чувствительности.

Альтернативный вариант использования многомасштабных признаков предложен в [2], где рассмотрена многомасштабная унифицированная сеть. Архитектура разделяет поверхностные слои на многомасштабные подсети для извлечения признаков из входных данных различного масштаба, затем объединяет низкоуровневые признаки в глубокие слои для формирования высокоуровневых семантических представлений. Введение масштабно-инвариантного ограничения обеспечивает согласованность признаков при разных масштабах, что позволяет повысить точность классификации до 44,53 % для входных изображений размером 32×32.

Для решения задачи сегментации папиллярного рака щитовидной железы (ПРЩЖ, Multi-Scale Adaptive Convolutional (MSAC)) в [3] предложены специализированные архитектуры MSAC-DBUNet и MSAC-DD с двойными декодерами. Эти сети используют адаптивную свертку на разных масштабах, что позволяет более точно выделять ядерные характеристики, являющиеся ключевыми для диагностики ПРЩЖ. Эксперименты показали, что MSAC-DD обеспечивает сопоставимую с MSAC-DBUNet точность при существенном сокращении времени вычислений и потребления GPU-ресурсов.

Отдельное внимание в статьях, посвященных использованию многомасштабных признаков, уделяется автоматическому обнаружению ядер клеток. Подход, предложенный в [4], основан на полностью сверточных сетях с многомасштабной регрессией плотностной карты. Архитектура без полносвязных слоев позволяет обрабатывать изображения произвольного размера, а использование морфологических операций и фильтров повышает надежность локализации центров клеток. На тестовых наборах достигнуты впечатляющие результаты: F1-мера составляла 0,9924 для флуоресцентных микроскопических изображений и 0,8440 – для гистопатологических срезов молочной железы.

Таким образом, современные многомасштабные архитектуры демонстрируют значительный потенциал в анализе гистопатологических изображений папиллярного рака. Их основное преиму-

щество заключается в способности интегрировать информацию из различных пространственных масштабов, в частности, при выявлении характерных ядерных аномалий при ПРИЦЖ.

Целью представленного в статье исследования являлась разработка архитектуры многомасштабной нейронной сети с несколькими входами для одновременной обработки групп гистологических изображений, полученных на различных уровнях увеличения, с последующим объединением признаков на уровне скрытых слоев.

Материалы и методы исследований

В качестве данных для обучения нейронных сетей использовался набор фрагментов гистопатологических изображений папиллярного рака, описанный в [5]. Набор данных был разделен на два класса: класс нормальной ткани и класс фрагментов изображений с наличием признаков папиллярного рака. Во второй класс вошли фрагменты, соответствующие разметке папиллярной, фолликулярной, солидной архитектоники рака. При обучении нейронных сетей учитывался индивидуальный идентификатор пациента, позволяющий исключить попадание фрагментов одного пациента одновременно в обучающую, валидационную или тестовую выборку. Размер фрагментов на входе сети составлял 224×224 пиксела. Были сформированы четыре набора, каждый из которых содержал фрагменты определенного увеличения по одной стороне фрагмента (49,29; 98,58; 197,16 и 394,32 мкм) (рис. 1).

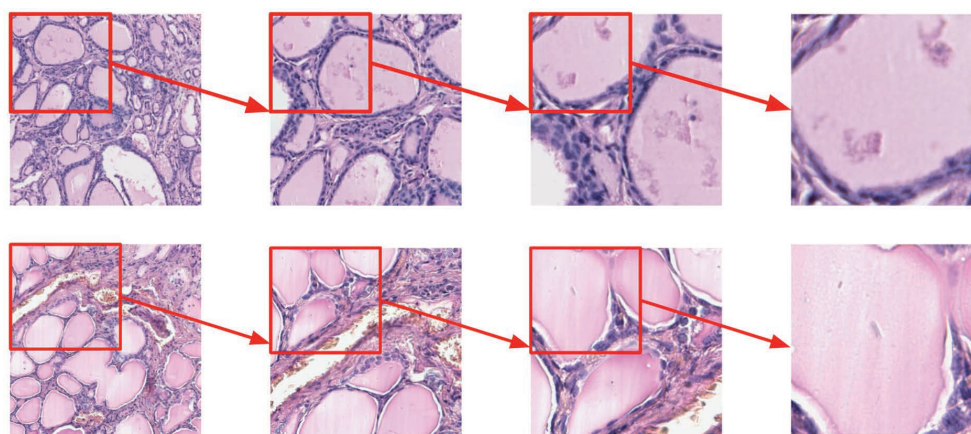


Рис. 1. Примеры сформированных групп фрагментов изображений разного увеличения
Fig. 1. Examples of formed groups of image fragments at different magnifications

Фрагменты изображений нарезались с привязкой к левому верхнему углу размеченного полигона. Такой подход был выбран с учетом особенностей реализованного скрипта обработки полнослайдовых изображений, а также для обеспечения строгого соответствия между фрагментом и его истинной морфологической принадлежностью. Учитывая значительную вариабельность размеров и формы областей разметки (от мелких очагов до крупных протяженных структур), выбор любой другой опорной точки (например, центра или другого угла) мог привести к тому, что один фрагмент частично перекрывал бы зоны с различными типами тканей. Это создало бы неоднозначность в разметке и исказило бы обучающую выборку, снижая качество обучения модели.

При обучении все фрагменты изображений аугментировались с использованием:

- геометрических трансформаций (отражения, повороты);
- трансформаций цвета (размытие, изменение контрастности и яркости, случайное удаление части изображения).

Всего набор данных включал 7244 группы фрагментов изображения нормальной ткани и 10 469 фрагментов изображений ткани, содержащей признаки наличия раковых опухолей. При обучении выборка формировалась сбалансированно по классам с ограничением размера каждого класса количеством фрагментов наименьшего класса, что обеспечивало равномерное представление всех категорий в процессе обучения.

В ходе исследований была разработана архитектура многомасштабной нейронной сети, реализован скрипт, позволяющий динамически выбирать количество входов нейронной сети

при обучении. Изображение с каждого входа проходило через сверточную нейронную сеть, после поступало на блок Cross-Scale Attention, основанный на слое multi-head attention [6]. Данный блок играл ключевую роль в интеграции информации между разными уровнями увеличения: он позволил модели адаптивно обмениваться признаками между низко- и высокомасштабными представлениями одного и того же изображения. Результаты, полученные для каждого из уровней увеличения, объединялись с исходными признаками через остаточный блок, что предотвращало потерю полезной информации и способствовало стабильности обучения. Далее применялась нормализация активации по каждому образцу, ускоряя сходимость и повышая устойчивость к колебаниям в данных. После данные конкатенировались и поступали на классификатор, в качестве которого выступала простая двухслойная полносвязная нейронная сеть.

В качестве сверточных блоков предлагаемой архитектуры использовалась модель EfficientNetB0 [7]. Ее выбор был обусловлен как результатами предыдущих исследований, так и наличием уже проведенных экспериментов на рассмотренном наборе данных с использованием одношкального подхода. Это позволило обеспечить корректное сравнение и оценку вклада многомасштабной стратегии за счет использования ранее полученного опорного результата. На рис. 2 представлена архитектура разработанной многомасштабной нейронной сети.

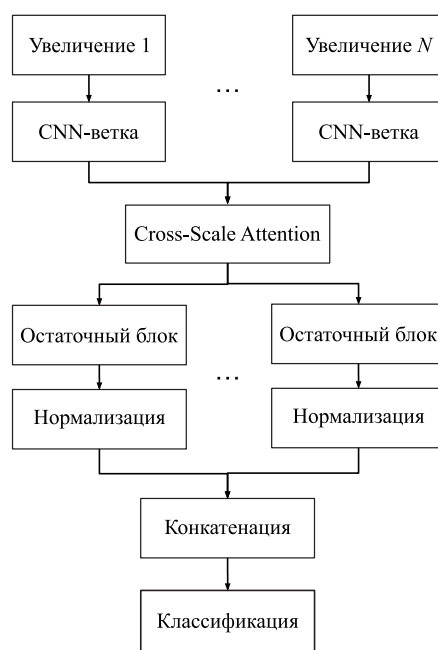


Рис. 2. Предлагаемая архитектура многомасштабной нейронной сети
Fig. 2. Proposed architecture of the multi-scale neural network

Нейронные сети с разным количеством входов обучались пять раз, и каждый раз класс наличия признаков папиллярного рака формировался случайным образом для обеспечения сбалансированности классов. Фрагменты изображений случайно перемешивались между выборками, исключая попадания в разные выборки фрагментов изображений одного пациента.

При обучении модели использовались следующие ключевые гиперпараметры: размер батча был равен 25, начальная скорость обучения составляла 10^{-5} с циклическим планировщиком (OneCycleLR) и максимальным значением 3×10^{-4} . Оптимизация проводилась с помощью оптимизатора Adam с применением обрезки градиентов по норме. Ранняя остановка осуществлялась по макроусредненной F1-мере на валидационной выборке при отсутствии улучшений в течение 15 эпох. Реализованный скрипт с подробным описанием можно найти на GitHub в репозитории по ссылке [8].

Результаты исследований и их обсуждение

В табл. 1 приведены усредненные результаты обучения по пяти экспериментам для нейронных сетей с разным количеством входов увеличения. При одном входе использовалась оригинальная

нейронная сеть EfficientNetB0, в качестве метрики для сравнения – средняя мера F1, рассчитанная по результатам бинарной классификации изображений тестового набора данных (mean \pm std).

Таблица 1. Усредненные результаты обучения многомасштабной нейронной сети с разным количеством входов

Table 1. Average results of training a multiscale neural network with different numbers of inputs

Количество входов нейронной сети	Размер изображения по одной стороне фрагмента, мкм	Средняя мера F1
Два входа	197,16 и 394,32	0,8850 \pm 0,0526
Три входа	98,58; 197,16 и 394,32	0,8962 \pm 0,0508
Четыре входа	49,29; 98,58; 197,16 и 394,32	0,8695 \pm 0,0504
Один вход	394,32	0,8481 \pm 0,0404

Наблюдаемое стандартное отклонение значений F1-меры отражает естественную вариабельность гистопатологических изображений, присущую разным подвыборкам данных. Примеры разнообразия фрагментов изображений, влияющих на колебание F1-меры в ходе формирования обучающих выборок, приведены на рис. 3. Поскольку разбиение на обучающую, валидационную и тестовую выборки проводилось случайным образом, состав каждой обучающей и валидационной подгруппы мог существенно различаться по доминирующим морфологическим паттернам, степени неоднородности окрашивания (рис. 3, *a*) и по внутриклассовому разнообразию (рис. 3, *b*). Это приводило к колебаниям метрик качества между запусками даже при одинаковой архитектуре и настройках обучения. Таким образом, высокое стандартное отклонение обусловлено не нестабильностью модели, а гетерогенностью данных – характерной особенностью медицинских изображений. Схожий характер изменения F1-меры во всех экспериментах подтверждает устойчивость поведения модели, несмотря на разнообразие входных данных.

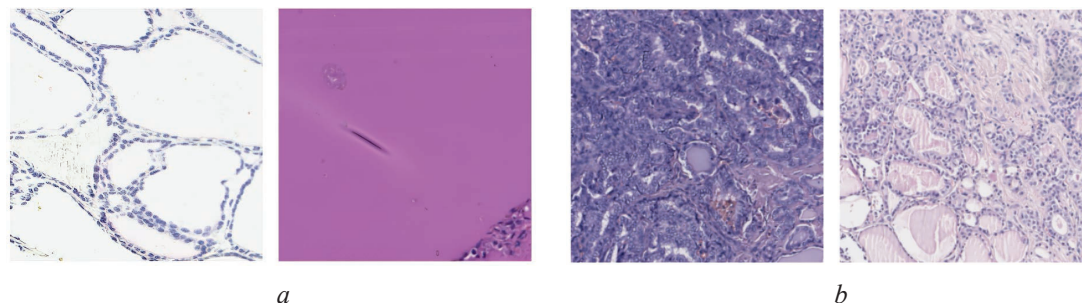


Рис. 3. Примеры разнообразия фрагментов изображений:

a – нормальной ткани одного уровня увеличения разной степени окрашивания;
b – разной архитектуры рака, относящегося к одному классу

Fig. 3. Examples of the diversity of image fragments: *a* – normal tissue at the same magnification level with varying degrees of staining; *b* – different architecture of cancer belonging to the same class

Проведенное исследование подтвердило высокую воспроизводимость результатов: предыдущий эксперимент с обучением модели EfficientNetB0 [6] был успешно воспроизведен в альтернативном программном окружении и с использованием иных технологий подготовки данных, что свидетельствует о надежности полученных ранее результатов. Применение многомасштабной архитектуры позволило достичь значимого улучшения качества классификации. Среднее значение F1-меры при использовании многомасштабной архитектуры возросло для всех рассмотренных конфигураций количества входов по сравнению с одношкальным подходом. Наилучшие показатели были достигнуты для модели с тремя входами разного увеличения, что обеспечило наиболее полное представление морфологических признаков на различных уровнях детализации (увеличение средней меры F1 на 5 % по сравнению с одномасштабным подходом). Данная конфигурация продемонстрировала среднюю метрику F1-меры на уровне $0,8962 \pm 0,0508$, а в отдельных запусках достигла значения 0,9697.

Несмотря на то что в предыдущих исследованиях увеличение масштаба фрагментов при обучении бинарных классификаторов приводило к снижению качества классификации моделью EfficientNetB0 (F1-мера = 0,66 для фрагмента 49,29 мкм по одной стороне, 0,68 – для 98,58 мкм,

0,70 – для 197 мкм и 0,83 – для 394,32 мкм) [6], в данной статье продемонстрирован противоположный эффект. При применении многомасштабной архитектуры три уровня увеличения 98,58; 197,16 и 394,32 мкм вносили положительный вклад в финальное качество классификации, обеспечивая более полное и устойчивое представление морфологических признаков за счет интеграции информации на разных уровнях детализации.

Одним из недостатков предлагаемого подхода являлся способ подготовки набора данных. Выбор левого угла при формировании групп увеличений был обусловлен спецификой реализованного скрипта и разметки. При выборе опорной точки в правом нижнем углу размеченной области фрагмент с более высоким увеличением был бы извлечен из соответствующей зоны в правой части изображения. Таким образом удавалось избегать попадания в фрагмент увеличенной области, содержащей ткани другого класса. В рамках тестирования многомасштабной архитектуры такого подхода было достаточно.

Более перспективным подходом могло бы стать формирование последовательности фрагментов, соответствующих требуемым уровням увеличения. Однако в рамках данного исследования был сознательно выбран упрощенный подход, основанный на стандартной архитектуре EfficientNetB0 без модификации ее внутренней структуры, имеющей строгое ограничение по размеру входа (224×224×3). Такой выбор обусловлен необходимостью обеспечить четкое и недвусмысленное понимание влияния многомасштабного представления данных на качество классификации. Использование немодифицированной, широко распространенной базовой архитектуры позволило изолировать эффект интеграции информации с разных масштабов от возможных искажений, вносимых изменениями в топологию сети, что способствует более объективной оценке вклада многомасштабного анализа.

Заключение

1. Интеграция информации из трех масштабов оказалась оптимальной стратегией, балансирующей между информативностью входных данных и устойчивостью модели, и может служить основой для дальнейшего развития систем автоматизированного анализа гистопатологических изображений. Применение разработанной архитектуры, интегрирующей данные с трех различных уровней увеличения, позволило повысить среднюю F1-меру на 5 % по сравнению с одношкальным подходом, что подтверждает важность комплексного анализа тканевой структуры на разных масштабах детализации. Наиболее высокие показатели качества были достигнуты при использовании фрагментов размерами 98,58; 197,16 и 394,32 мкм по стороне, где модель продемонстрировала устойчивое поведение с метрикой F1-меры $0,8962 \pm 0,0508$, а в отдельных запусках – рекордное значение 0,9697. При этом наблюдаемое стандартное отклонение объясняется не нестабильностью модели, а естественной вариабельностью медицинских данных, связанной с различиями в окраске, морфологии и стратегии разметки.

2. Полученные результаты открывают перспективы для использования данной технологии в системах компьютерной помощи диагностике в качестве средства для автоматического выделения подозрительных участков на полнослайдовых изображениях. Однако предложенный подход имеет ряд ограничений: зависимость от качества ручной разметки, недостаточная интерпретируемость решений, принимаемых моделью, и эмпирический выбор конфигурации входных масштабов. Одно из направлений дальнейшего исследования заключается во внедрении методов поиска оптимальных архитектур для оптимизации комбинации увеличений, в использовании обучаемых механизмов слияния признаков, а также в применении подходов к визуализации карт активаций для повышения интерпретируемости принимаемых решений.

3. Работа выполнена при поддержке Белорусского республиканского фонда фундаментальных исследований в рамках проекта «Наука М 2025» № Ф25М-001. Отдельная благодарность выражается М. В. Фридману за предоставленную разметку гистопатологических изображений, а также В. А. Ковалеву за ценные научные консультации.

Список литературы

1. Multi-Scale Feature Fusion Convolutional Neural Network for Indoor Small Target Detection / L. Huang [et al.] // *Frontiers in Neurorobotics*. 2022. Vol. 16. DOI: 10.3389/fnbot.2022.881021.
2. Liu Wenzhuo. Multi-Scale Unified Network for Image Classification / Liu Wenzhuo, Zhu Fei, Liu Cheng-Lin // *arXiv preprint*. 2024. Mode of access: <https://arxiv.org/abs/2405.12345>. Date of access: 09.10.2025.

3. Papillary Thyroid Carcinoma Semantic Segmentation Using Multi-Scale Adaptive Convolutional Network with Dual Decoders / T. Parasporins [et al.] // IEEE Access. 2025. Vol. 13. P. 17340–17353. DOI: 10.1109/ACCESS.2025.3532505.
4. Cell Detection in Pathology and Microscopy Images with Multi-Scale Fully Convolutional Neural Networks / X. Pan [et al.] // World Wide Web. 2018. Vol. 21. P. 1721–1743. DOI: 10.1007/s11280-017-0520-7.
5. Методика формирования базы гистопатологических изображений папиллярного рака щитовидной железы для глубокого обучения / М. В. Фридман [и др.] // Информатика. 2023. Т. 20, №. 2. С. 28–38. DOI: 10.37661/1816-0301-2023-20-2-28-38.
6. Attention Is All You Need / A. Vaswani [et al.] // arXiv preprint. 2017. Mode of access: <https://arxiv.org/abs/1706.03762>. Date of access: 09.10.2025.
7. Tan, M. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks / M. Tan, Q. V. Le // Proceedings of the 36th International Conference on Machine Learning (ICML 2019), Long Beach, California, USA, 9–15 June 2019. Long Beach, 2019. P. 6105–6114.
8. Multi_input_network-Histologia [Electronic Resource] // GitHub. Mode of access: https://github.com/Rui4317/Multi_input_network-Histologia. Date of access: 14.10.2025.

Поступила 14.10.2025

Принята в печать 27.11.2025

References

1. Huang L., Zhang Y., Yang X., Wang Z., Li H. (2022) Multi-Scale Feature Fusion Convolutional Neural Network for Indoor Small Target Detection. *Frontiers in Neurorobotics*. 16. DOI: 10.3389/fnbot.2022.881021.
2. Liu Wenzhuo, Zhu Fei, Liu Cheng-Lin (2024) Multi-Scale Unified Network for Image Classification. *arXiv preprint*. Available: <https://arxiv.org/abs/2405.12345> (Accessed 9 October 2025).
3. Parasporins T., Payatsuporn P., Jitaree A., Boonkong C. (2025) Papillary Thyroid Carcinoma Semantic Segmentation Using Multi-Scale Adaptive Convolutional Network with Dual Decoders. *IEEE Access*. 13, 17340–17353. DOI: 10.1109/ACCESS.2025.3532505.
4. Pan X., Yang D., Li L., Xie W., Yang H., Chen S. (2018) Cell Detection in Pathology and Microscopy Images with Multi-Scale Fully Convolutional Neural Networks. *World Wide Web*. 21, 1721–1743. DOI: 10.1007/s11280-017-0520-7.
5. Fridman M. V., Kosareva A. A., Snezhko E. V., Kamlach P. V., Kovalev V. A. (2023) Papillary Thyroid Carcinoma Whole-Slide Images as a Basis for Deep Learning. *Informatics*. 20 (2), 28–38. DOI: 10.37661/1816-0301-2023-20-2-28-38 (in Russian).
6. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., et al. (2017) Attention Is All You Need. *arXiv preprint*. Available: <https://arxiv.org/abs/1706.03762> (Accessed 9 October 2025).
7. Tan M., Le Q. V. (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *Proceedings of the 36th International Conference on Machine Learning (ICML 2019), Long Beach, California, USA, 9–15 June*. Long Beach. 6105–6114.
8. Multi_input_network-Histologia. *GitHub*. Available: https://github.com/Rui4317/Multi_input_network-Histologia (Accessed 14 October 2025).

Received: 14 October 2025

Accepted: 27 November 2025

Сведения об авторе

Косарева А. А., ассист. каф. электронной техники и технологии, Белорусский государственный университет информатики и радиоэлектроники; мл. науч. сотр. Объединенного института проблем информатики Национальной академии наук Беларуси

Адрес для корреспонденции

220013, Республика Беларусь,
Минск, ул. П. Бровки, 6
Белорусский государственный университет
информатики и радиоэлектроники
Тел.: +375 17 293-88-60
E-mail: kosareva@bsuir.by
Косарева Александра Андреевна

Information about the author

Kosareva A. A., Assistant at the Electronic Engineering and Technology Department, Belarusian State University of Informatics and Radioelectronics; Junior Researcher at the The United Institute of Informatics Problems of the National Academy of Sciences of Belarus

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovki St., 6
Belarusian State University
of Informatics and Radioelectronics
Tel.: +375 17 293-88-60
E-mail: kosareva@bsuir.by
Kosareva Aleksandra Andreevna