

РАЗРАБОТКА ПРИНЦИПОВ АСАМБЛЕВОГО ОБУЧЕНИЯ ИНФОРМАЦИИ В ЗАДАЧАХ МЕДИЦИНСКОЙ ДИАГНОСТИКИ, В ЧАСТНОСТИ ВЫЯВЛЕНИЯ ПРЕДРАКОВЫХ СОСТОЯНИЙ

DEVELOPMENT OF PRINCIPLES OF ASSEMBLY-BASED INFORMATION LEARNING IN THE TASKS OF MEDICAL DIAGNOSTICS, IN PARTICULAR, THE DETECTION OF PRECANCEROUS CONDITIONS

Ju. Rubekina

Summary. Oral cancer is one of the most aggressive and dangerous types of cancer, claiming the lives of thousands of people every year. In recent decades, there has been an alarming trend in the spread of the disease: the number of cases of oral cancer is increasing, while young people are also suffering from the disease. Despite the fact that oral cancer is a visually detectable neoplasm, the rate of advanced cases of the disease continues to be critically high. Of particular importance is the development of flexible screening algorithms that consider specific risk factors. This approach will make it possible to move from simple compliance with formal standards to the creation of a truly effective and working system for the early detection of oral diseases [6].

However, effective diagnosis of cancer and precancerous diseases of the oral cavity requires in-depth knowledge of their clinical manifestation and modern methods of their detection. The subject of the development is an automated diagnostic technique based on hardware — multispectral imaging in the UV and IR ranges combined with a standardized positioning system, and software — an algorithm for data preprocessing and ensemble classification of components.

Keywords: data processing, automated diagnostics, ensemble classification, screening, early diagnosis, automated system.

Рубекина Юлия Сергеевна

Инженер-разработчик,

ООО «Медиана», г. Санкт-Петербург

Ru.julia@gmail.com

Аннотация. Рак ротовой полости — один из наиболее агрессивных и опасных видов онкологических заболеваний, ежегодно уносящий жизни тысяч людей. В последние десятилетия отмечается тревожная тенденция в динамике распространения болезни: число случаев рака полости рта растет, при этом от болезни страдают и молодые люди. Несмотря на то, что рак полости рта относится к визуально определяемым новообразованиям, уровень запущенных случаев заболевания продолжает оставаться критически высоким. Особое значение приобретает разработка гибких скрининговых алгоритмов с учетом специфических факторов риска. Такой подход позволит перейти от простого соблюдения формальных норм к созданию действительно эффективной и рабочей системы раннего выявления заболеваний ротовой полости [6].

Однако эффективная диагностика рака и предраковых заболеваний полости рта требует глубоких знаний об их клиническом проявлении и современных методах их выявления. Предметом разработки является методика автоматизированной диагностики на основе аппаратного — мультиспектральная визуализация в УФ и ИК диапазонах, совмещенная со стандартизированной системой позиционирования, и программного — алгоритм предобработки данных и ансамблевая классификация, компонентов.

Ключевые слова: обработка данных, автоматизированная диагностика, ансамблевая классификация, скрининг, ранняя диагностика, автоматизированная система.

За последние годы в развитии онкологии слизистой полости рта прослеживается тревожная динамика, указывающая на то, что с данным диагнозом могут столкнуться не только пожилые мужчины злоупотребляющими вредными привычками, но молодые люди, в том числе женщины без вредных привычек. С каждым годом диагностируется все больше случаев раковых форм новообразований ротовой полости, при раннем выявлении которых выживаемость пациентов может достигать 80–90 %, но на III–IV стадиях этот показатель не превышает 40 %. Предраковые состояния, такие как лейкоплакия и эритроплакия могут выглядеть даже для опытного стоматолога как несерьезная травма или стоматит. Неуклонный рост заболеваемости диктует необходимость трансформации существующей системы скрининга.

Сложившаяся ситуация требует принципиально новых организационных решений в системе здравоохранения, направленных на оптимизацию выявления предопухолевой и опухолевой патологии челюстно-лицевой области. Особое значение приобретает разработка гибких скрининговых алгоритмов с учетом региональных особенностей и специфических факторов риска. Такой подход позволит перейти от простого соблюдения формальных норм к созданию действительно эффективной и рабочей системы раннего выявления заболеваний ротовой полости.

Однако эффективная диагностика рака и предраковых заболеваний полости рта требует глубоких знаний об их клиническом проявлении и современных методах

их выявления. Предметом разработки является методика автоматизированной диагностики на основе аппаратного — мультиспектральная визуализация в УФ и ИК диапазонах, совмещенная со стандартизированной системой позиционирования, и программного — алгоритм предобработки данных и ансамблевая классификация, компонентов.

В основе предлагаемого подхода к классификации лежат передовые принципы ансамблевого обучения, доказавшие свою эффективность в задачах медицинской диагностики. Главное преимущество такой методологии заключается в возможности объединить различные алгоритмы для комплексного анализа разнородных данных. Как подтверждают современные исследования в области медицинского искусственного интеллекта, ансамблевые методы демонстрируют более высокую диагностическую точность по сравнению с отдельно взятыми моделями, особенно в сложных случаях, требующих комплексной оценки множества связанных факторов.

Эффективность ансамблевого подхода обусловлена несколькими ключевыми факторами. Во-первых, комбинация 3D CNN (3D Convolutional Networks, трехмерные сверточные сети) и традиционных методов машинного обучения позволяет одновременно учитывать как сложные пространственно-спектральные зависимости в исходных изображениях, так и количественные характеристики, извлеченные в ходе предварительной обработки изображений. Согласно исследованиям, подобные гибридные архитектуры обеспечивают на 15–20 % более высокую точность по сравнению с отдельными моделями обучения.

Во-вторых, ансамблевый подход демонстрирует повышенную устойчивость к характерным для медицинских изображений артефактам и шумам. Это имеет решающее значение при работе с ультрафиолетовыми и инфракрасными каналами, где часто наблюдаются проблемы с засветкой или неравномерным освещением. Разнообразие моделей в ансамбле позволяет компенсировать искажения в отдельных каналах за счет информации из других источников данных. Особого внимания заслуживает клиническая интерпретируемость результатов, обеспечиваемая механизмами *soft voting*, которые позволяют количественно оценивать степень уверенности системы на основе согласованности предсказаний отдельных классификаторов.

Ансамблевая методология демонстрирует эффективность в условиях ограниченных выборок, характерных для клинической практики, где количество подтвержденных случаев предраковых состояний часто недостаточно для обучения отдельных моделей. Современные исследования демонстрируют, что комбинация трансферного обучения на предобученных CNN с классиче-

скими алгоритмами машинного обучения позволяет достигать высокой диагностической точности даже при небольшом количестве подтвержденных случаев. Это актуально для диагностики предраковых состояний, где количество подтвержденных случаев часто невелико.

Решающим аспектом является и возможность параллельной обработки данных. Параллельная обработка данных, когда 3D сверточные сети работают с исходными изображениями, а методы типа XGBoost анализируют извлеченные признаки, обеспечивает не только высокую точность, но и оперативность диагностики, ключевые в клинической практике.

Из этих рассуждений становится ясно, что ансамблевый подход представляет собой оптимальное решение для задачи классификации патологий полости рта, сочетая высокую точность, устойчивость к артефактам и клиническую интерпретируемость результатов. Это подтверждается как теоретическими исследованиями, так и практическими результатами, полученными в ходе медицинских испытаний.

Входными данными для системы служат мультиспектральные изображения полости рта, содержащие четыре спектральных канала: видимый спектр (RGB), ультрафиолетовый диапазон (365 нм) и два инфракрасных канала (850 и 950 нм). Эти данные дополняются бинарными масками сегментации, выделяющими области патологических изменений, и при наличии — клиническими метаданными пациента.

Первичный анализ направлен на многоуровневое выделение гетерогенных характеристик из входных изображений. Для получения глубинных признаков используется модифицированная архитектура ResNet50, где последние слои адаптированы под специфику медицинских изображений. Архитектура ResNet50 основана на концепции остаточного обучения, которое решает ключевую проблему глубоких нейронных сетей — деградацию точности при увеличении глубины. Основной строительный блок сети — остаточный модуль, содержащий так называемые «skip-connections», позволяющие градиенту свободно распространяться в обратном направлении в процессе обучения. Эти соединения реализуют принцип тождественного отображения, когда при необходимости входные данные могут передаваться на более глубокие слои без изменений.

Сеть состоит из пяти стадий обработки данных. Первая стадия включает начальную свертку большого ядра (7×7) с последующей операцией *max-pooling* с целью выделения низкоуровневых признаков. Далее следуют четыре группы остаточных блоков, каждая из которых работает с определенным уровнем абстракции признаков. В ResNet50 используется модификация остаточного

блока с «бутылочным горлышком» (bottleneck), где сначала происходит уменьшение размерности с помощью свертки 1×1 , затем применяется свертка 3×3 , и наконец — восстановление размерности. Такой подход значительно снижает вычислительную сложность без потери качества.

Особенностью архитектуры является постепенное увеличение количества фильтров (от 64 до 2048) при одновременном уменьшении пространственных размерностей карт признаков (от 56×56 до 7×7). Завершающая часть сети включает глобальный средний пулинг, преобразующий пространственные карты признаков в компактный вектор размерностью 2048 элементов, который затем используется полносвязным слоем для классификации. Алгоритм данного этапа отображен на схеме рисунка 1. Благодаря остаточным соединениям и продуманной организации блоков, ResNet50 демонстрирует высокую эффективность при работе с различными типами изображений, включая медицинские данные.

Параллельно вычисляются спектрально-текстурные характеристики. Для каждого спектрального канала строится нормированная гистограмма распределения интенсивностей пикселей, которое вычисляется по формуле (1). Это дает инструмент для количественной оценки особенностей светопоглощения в различных патологических состояниях:

$$H(i) = \frac{N_i}{N_{total}}, i \in [0, 255], \quad (1)$$

где: N_i — количество пикселей со значением интенсивности i ;

N_{total} — общее число пикселей в области интереса.

Текстура анализируется с использованием матрицы GLCM (Gray-Level Co-occurrence Matrix), предоставляющей пять диагностически значимых параметров: контраст, энергию, однородность, корреляцию и дисперсию. Наибольшее диагностическое значение имеют ультрафиолетовый и дальний инфракрасный каналы, демонстрирующие наибольшую чувствительность к клеточным структурам. Морфологический анализ сегментированных областей включает расчет трех ключевых показателей:

— площади поражения по формуле (2):

$$A = \sum_{i=1}^N I(p_i \in R), \quad (2)$$

где A — площадь в пикселях;

I — индикаторная функция (1 если пиксель принадлежит области);

p_i — координаты пикселя;

R — сегментированная область.

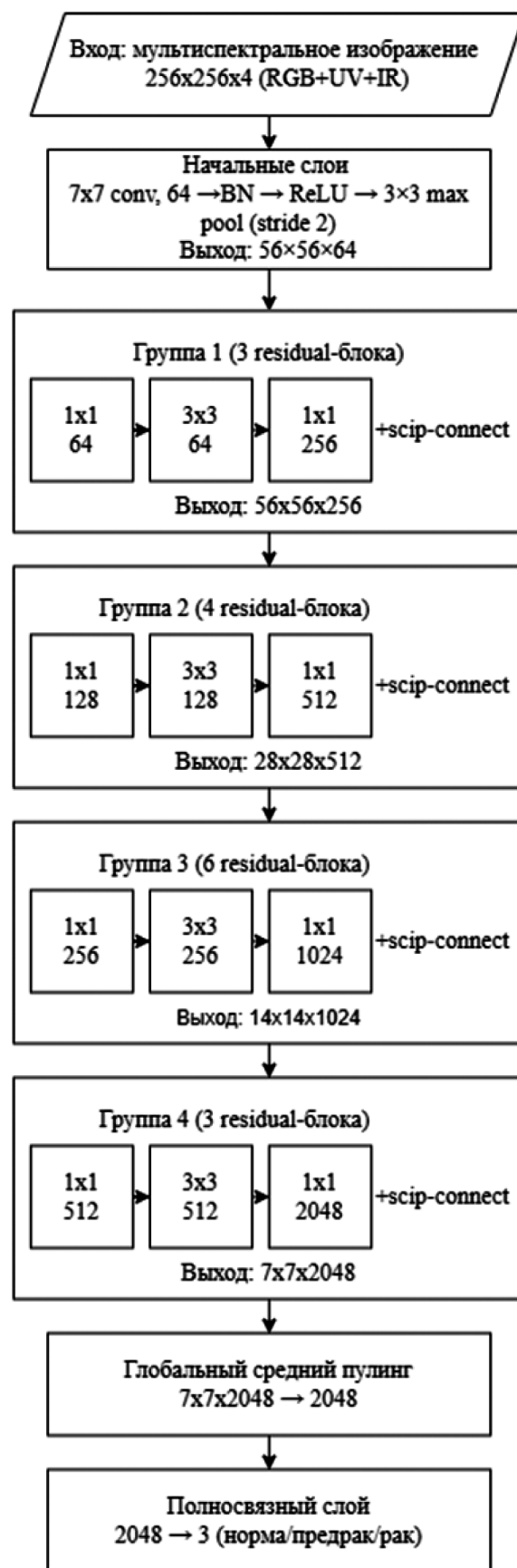


Рис. 1. Архитектура модели ResNet50 для задачи классификации

— коэффициента эксцентриситета (количественно характеризующего отклонение формы от окружности) по формуле (3):

$$e = \sqrt{1 - \frac{b^2}{a^2}}, \quad (3)$$

где a — длина большой полуоси эллипса;
 b — длина малой полуоси.

— параметра solidity, отражающего степень выпуклости контура интересующей для анализа области по формуле (4):

$$S = \frac{A}{A_{convex}} \quad (4)$$

где A — площадь реального контура;
 A_{convex} — площадь выпуклой оболочки.

Эти метрики обладают высокой диагностической ценностью, поскольку злокачественные новообразования обычно проявляют выраженную нерегулярность формы и фрагментированность границ.

На этапе обучения создается гибридная ансамблевая модель, объединяющая три компонента. Ядром системы служит трехмерная сверточная сеть (3D CNN), специально разработанная для обработки мультиспектральных данных. Архитектура сети включает последовательность трехмерных сверточных слоев с ядрами размером $3 \times 3 \times 2$ с целью анализа пространственно-спектральные корреляции. Вторым компонентом — градиентный бустинг (XGBoost), принимающий на вход объединенный вектор из глубинных, спектральных и морфологических признаков входных данных. Третий элемент — метод опорных векторов с ядром (SVM), добавляющий линейную разделяющую способность.

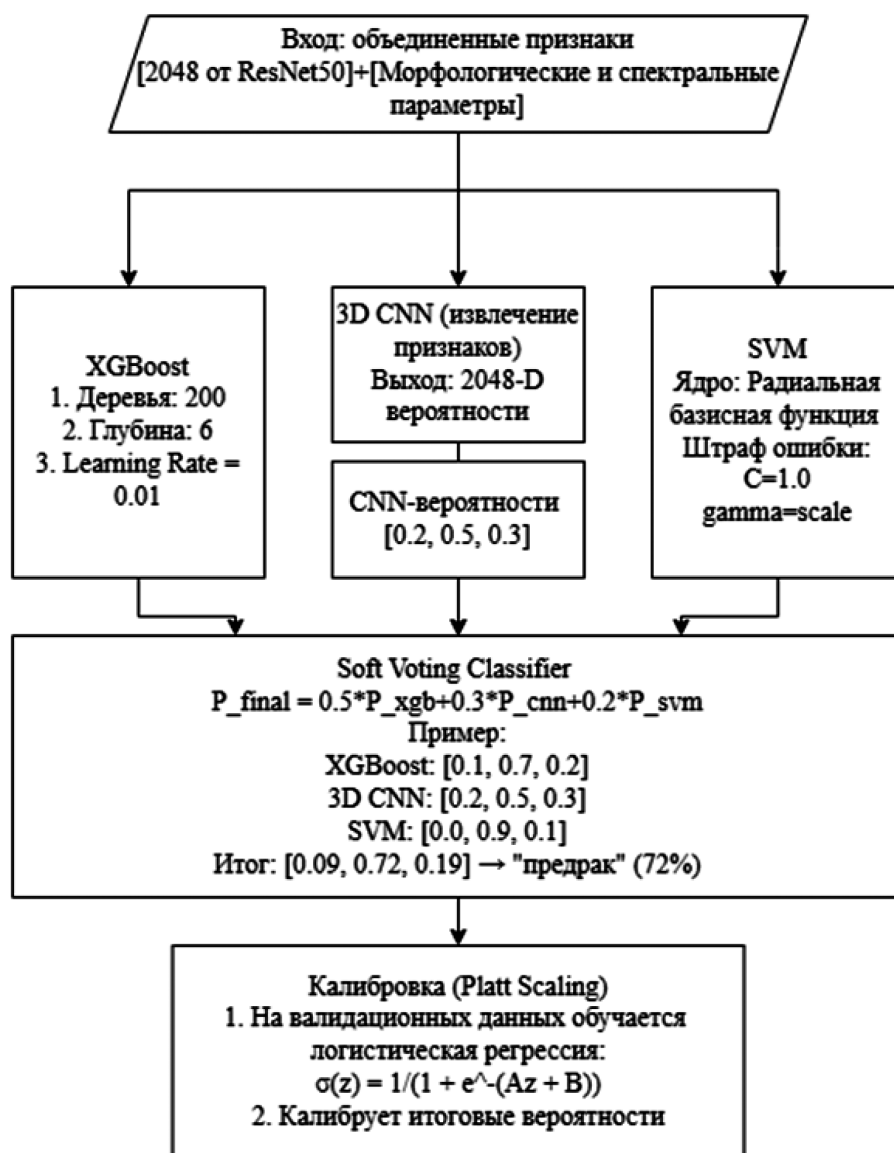


Рис. 2. Этап ансамблевой классификации

Финальное решение принимается путем мягкого голосования ансамбля, где каждый классификатор вносит взвешенный вклад в соответствии с его точностью на валидационной выборке. Для калибровки выходных вероятностей применяется метод Платта, который преобразует «сырые» оценки классификаторов в достоверные вероятностные оценки с помощью логистической регрессии. Детализированный этап ансамблевой классификации представлен на рисунке 2. Это играет ключевую роль для медицинских приложений, где критична точная интерпретация степени уверенности системы в полученных результатах обучения.

Оптимальный порог классификации определяется по индексу Юдена, максимизирующему разницу между истинно положительными и ложно положительными результатами. Для пространственного уточнения масок используется условное случайное поле (CRF), которое учитывает как вероятности пикселей, так и их пространственную согласованность, устраняя точечные артефакты сегментации.

Интерпретация результатов обеспечивается двумя методами: Grad-CAM визуализирует области изображения, наиболее повлиявшие на решение модели, выделяя диагностически значимые зоны тепловой картой. SHAP-анализ количественно оценивает вклад каждого признака, позволяя выявить ключевые спектральные биомаркеры. Например, анализ может показать, что гистограмма в УФ-диапазоне имеет наибольшую значимость для выявления дисплазии, а текстура в IR950 канале критична для дифференциации инвазивных форм рака. Полный алгоритм классификации отображен блок-схемой, представленной далее на рисунке 3.

Вычислительная эффективность системы обеспечивается оптимизированной архитектурой: 3D CNN обрабатывает изображение размером $256 \times 256 \times 4$ пикселя за 8 миллисекунд на графическом ускорителе NVIDIA T4. Ансамблевый подход (3D CNN + XGBoost + SVM) обеспечивает статистически значимое улучшение точности на 9 % по метрике F1-score по сравнению с отдельными моделями, сохраняя интерпретируемость решений — ключевое требование для медицинских приложений. Система поддерживает два режима работы: быстрый скрининг на основе 3D CNN и углубленный анализ сложных случаев с полным ансамблем классификаторов.

Разработанный алгоритм демонстрирует высокую диагностическую точность за счет комбинации современных методов глубокого обучения и машинного обучения. Ансамблевый подход позволяет анализировать пространственно-спектральные зависимости, текстуру и морфометрические параметры, обеспечивая дифференциацию между здоровыми тканями, предраковыми состояниями и злокачественными новообразованиями.

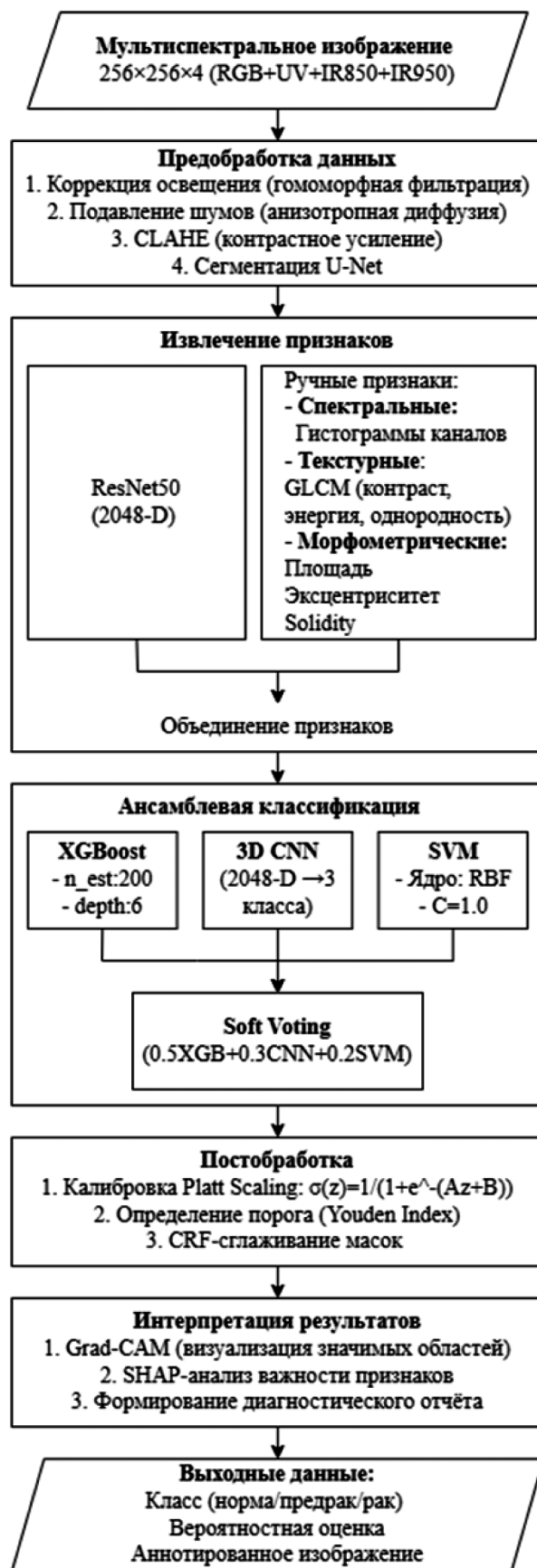


Рис. 3. Полный алгоритм модуля классификации

Для внедрения системы в клиническую практику следует не только обеспечить высокую точность алгоритмов, но и разработать удобный интерфейс для врачей.

Это требует создания специализированного программного обеспечения, интегрирующего предложенные методы в рабочий процесс специалиста диагностики.

ЛИТЕРАТУРА

1. Гажва С.И., Грехов А.В., Горячева Т.П., Сенина-Волжская И.В. Проблема ранней диагностики онкопатологии слизистой оболочки рта (социальные аспекты) // Современные проблемы науки и образования. — 2015. — № 3.
2. Булгакова Н.Н., Волков Е.А., Позднякова Т.И. Аутофлуоресцентная стоматоскопия как метод онкоскрининга заболеваний слизистой оболочки рта // Российский стоматологический журнал. — 2015. — № 1. — С. 34–39
3. Усанов М.С., Кульберг Н.С., Морозов С.П. Опыт применения адаптивных гомоморфных фильтров для обработки компьютерных томограмм // Информационные технологии и вычислительные системы. — 2017. — № 2. — С. 33–42
4. Рябых М.С., Сойникова Е.С., Батищев Д.С., Михелёв В.М. Высокопроизводительный метод морфологической обработки медицинских изображений // Интерактивная наука. — 2016. — № 5. — С. 34–41.
5. Ragab D.A., Sharkas M., Marshall S., Ren J. Breast cancer detection using deep convolutional neural networks and support vector machines // PeerJ. — 2019. — Vol. 7.
6. Abdulkadir A. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation // Medical Image Computing and Computer-Assisted Intervention — MICCAI 2016. — 2016. — P. 424–432.
7. Reza A.M. Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement // The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology. — 2004. — Vol. 38, № 1. — P. 35–44.
8. Baker M.J., et al. Using Fourier transform IR spectroscopy to analyze biological materials // Nature Protocols. — 2014. — Vol. 9, № 8. — P. 1771–1791.
9. Stockham T.G. Image processing in the context of a visual model // Proceedings of the IEEE. — 1972. — Vol. 60, № 7. — P. 828–842.
10. Gonzalez R.C., Woods R.E. Digital Image Processing. — 4th ed. — Pearson, 2018. — 1168 p.

© Рубекина Юлия Сергеевна (Ru.julia@gmail.com)

Журнал «Современная наука: актуальные проблемы теории и практики»