

DOI: <https://doi.org/10.17816/DD625866>

Использование методов машинного обучения при диагностике рака молочной железы

К.С. Дёмин, И.В. Гермашев

Волгоградский государственный университет, Волгоград, Россия

АННОТАЦИЯ

В последние несколько лет методы машинного обучения привлекают всё большее внимание в области диагностики, особенно при обнаружении рака молочной железы.

В статье проведён анализ актуальных исследований, посвящённых использованию методов машинного обучения при диагностике рака молочной железы, по трём направлениям: используемых для решения вспомогательных задач, возникающих в современной диагностике рака молочной железы, выступающих в качестве интеллектуальной оценки состояния пациента для предварительных диагностических решений и определяющих факторы риска рака молочной железы.

Анализ показал, что использование методов машинного обучения при диагностике рака молочной железы предоставляет большие возможности для улучшения точности и эффективности диагностики, а также позволяет решать дополнительные задачи.

В результате анализа литературы были выявлены признаки, которые используются в качестве входных данных в методах машинного обучения. В дальнейшем собранная информация будет использована при построении системы признаков для диагностики рака молочной железы методами машинного обучения.

Ключевые слова: рак молочной железы; скрининг рака; машинное обучение, искусственный интеллект.

Как цитировать:

Дёмин К.С., Гермашев И.В. Использование методов машинного обучения при диагностике рака молочной железы // Digital Diagnostics. 2024. Т. 5, № 3. С. 578–591. DOI: <https://doi.org/10.17816/DD625866>

DOI: <https://doi.org/10.17816/DD625866>

Machine learning techniques for breast cancer diagnosis

Kirill S. Dyomin, Ilya V. Germashev

Volgograd State University, Volgograd, Russia

ABSTRACT

In the last few years, machine learning techniques have been attracting even greater attention in the field of diagnostics, particularly when detecting breast cancer. Relevant studies dedicated to machine learning techniques in breast cancer diagnosis were analyzed in three areas: solving secondary problems that occur in modern-day breast cancer diagnostics, role in an intelligent assessment of the patient's condition for preliminary diagnostic decisions, and capability to detect breast cancer risk factors. The results revealed that machine learning techniques applied in breast cancer diagnosis have great potential for improving diagnostic accuracy and efficiency and solving secondary problems. The medical literature analysis has determined the parameters that are used as input data in machine learning techniques. Furthermore, the collected information will be applied to create a parameter system for breast cancer diagnosis using machine learning techniques.

Keywords: breast cancer; cancer screening; machine learning; artificial intelligence.

To cite this article:

Dyomin KS, Germashev IV. Machine learning techniques for breast cancer diagnosis. *Digital Diagnostics*. 2024;5(3):578–591.

DOI: <https://doi.org/10.17816/DD625866>

Submitted: 12.02.2024

Accepted: 30.05.2024

Published online: 13.09.2024

DOI: <https://doi.org/10.17816/DD625866>

利用机器学习方法诊断乳腺癌

Kirill S. Dyomin, Ilya V. Germashev

Volgograd State University, Volgograd, Russia

摘要

最近几年，机器学习方法在诊断领域，特别是对乳腺癌的检测受到越来越多的关注。

文章从三个方面分析了当前致力于在乳腺癌诊断中使用机器学习方法的研究：用于解决现代乳腺癌诊断中出现的辅助任务，作为初步诊断决策的患者病情智能评估，以及确定乳腺癌的风险因素。文章从三个方面分析了机器学习方法在乳腺癌诊断中的应用现状：解决现代乳腺癌诊断中出现的辅助课题，智能评估患者病情初步诊断决策，以及确定乳腺癌的风险因素。分析表明，利用机器学习方法诊断乳腺癌为提高诊断的准确性和效率提供巨大的可能性，并且还可以完成其他额外任务。

文献分析结果，确定了用作机器学习方法中输入数据的特征。今后，收集到的信息将用于构建一个利用机器学习方法诊断乳腺癌的特征系统。

关键词：乳腺癌；癌症筛查；机器学习；人工智能。

引用本文：

Dyomin KS, Germashev IV. 利用机器学习方法诊断乳腺癌. *Digital Diagnostics*. 2024;5(3):578–591. DOI: <https://doi.org/10.17816/DD625866>

收到: 12.02.2024

接受: 30.05.2024

发布日期: 13.09.2024

ВВЕДЕНИЕ

По данным 2022 г. рак молочной железы составляет 19% всех видов рака среди выявленных случаев и является лидирующим в этом списке [1]. Ранняя диагностика данной опухоли является решающим шагом, что сильно способствует успешному лечению и повышению выживаемости пациентов. В последние несколько лет методы машинного обучения привлекают всё большее внимание в области диагностики, в том числе при обнаружении рака молочной железы. Так, в исследовании S. Zou и соавт. [2] проводился сравнительный анализ методов машинного обучения, в которых в качестве входной информации были использованы демографические данные пациентов, истории болезни, данные визуализации: маммография, ультразвуковое исследование (УЗИ), магнитно-резонансная томография (МРТ) и генетические профили. Результаты показали, что алгоритмы машинного обучения, такие как машины опорных векторов (SVM), случайные леса, искусственные нейронные сети (ИНС) и модели глубокого обучения, имеют высокую точность для диагностики рака молочной железы.

В работе проведено выявление возможностей машинного обучения при диагностике рака молочной железы. В данном обзоре рассматриваются методы машинного обучения, используемые для решения вспомогательных задач, возникающих в современной диагностике рака молочной железы, выступающие в качестве интеллектуальной оценки состояния пациента для предварительных диагностических решений и применяемые в качестве предсказательной системы.

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РЕШЕНИЯ ВСПОМОГАТЕЛЬНЫХ ЗАДАЧ В МЕДИЦИНСКОЙ ДИАГНОСТИКЕ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ

Методы машинного обучения могут быть эффективно применены в диагностике рака молочной железы для решения вспомогательных задач. Одной из таких задач является сегментация изображений с помощью компьютерного зрения. Методы машинного обучения, такие как свёрточные нейронные сети (Convolutional Neural Networks, CNN), могут быть использованы для автоматической сегментации опухолей и других аномалий на маммограммах [3]. Ещё одним аспектом является классификация данных для определения наличия или отсутствия рака молочной железы. Здесь могут использоваться методы машинного обучения, включая метод опорных векторов [4] или случайный лес [5]. В рамках данного раздела было проанализировано порядка 120 работ, где использовались модели машинного обучения для автоматизации

рутинных или решения вспомогательных задач. Ниже приводятся некоторые примеры.

В дополнение к этому, методы машинного обучения могут использоваться для анализа генетической информации для прогнозирования рака молочной железы на основе геномных данных [6].

Далее будут рассмотрены опубликованные работы, где применились методы машинного обучения для решения этих задач в диагностике рака молочной железы при помощи классических методов диагностики: маммографии, УЗИ, МРТ и др.

Одним из ранних исследований диагностики рака молочной железы с применением методов машинного обучения была работа, в которой применялась ИНС [7]. В качестве входных данных были использованы поверхностная плотность желёз и коэффициент вариации люминального форм-фактора. ИНС правильно классифицировала 92% случаев в наборе данных, исходя из этого авторы предположили, что ИНС могут быть полезны для дифференциальной диагностики. В результате данного прорывного исследования был открыт путь к дальнейшему развитию методов машинного обучения в области диагностики рака молочной железы. С тех пор методы машинного обучения продолжали развиваться. Они помогают повысить точность и скорость постановки диагноза, а также уменьшить количество ошибочных результатов.

Методы машинного обучения в маммографии

Алгоритмы машинного обучения активно применяются в хорошо зарекомендовавших себя методах диагностики рака молочной железы. Так, например, в исследовании L. Tsochatzidis и соавт. [8] используется такой метод машинного обучения, как свёрточная ИНС, для уменьшения количества ложноположительных результатов и классификации опухолей на доброкачественные и злокачественные. Использование машинного обучения позволило достигнуть высокой точности диагностики и повысить эффективность исследования.

В другом исследовании применялась свёрточная ИНС для локализации опухоли, даже если она развилась в нескольких областях, а также для классификации опухоли на доброкачественную или злокачественную [3]. Пример работы данной модели представлен на рис. 1. Полученный классификатор показал точность в тестовых примерах около 85%, что может быть использовано в качестве вспомогательного инструмента для радиолога при выявлении рака молочной железы.

В похожем исследовании применялась нейронная сеть с радиально-базисными функциями для реализации автоматизированной классификации и сегментации опухолей молочной железы [9]. Данная работа показала высокую точность локализации области опухоли на исходном входном изображении с цветным результатом.

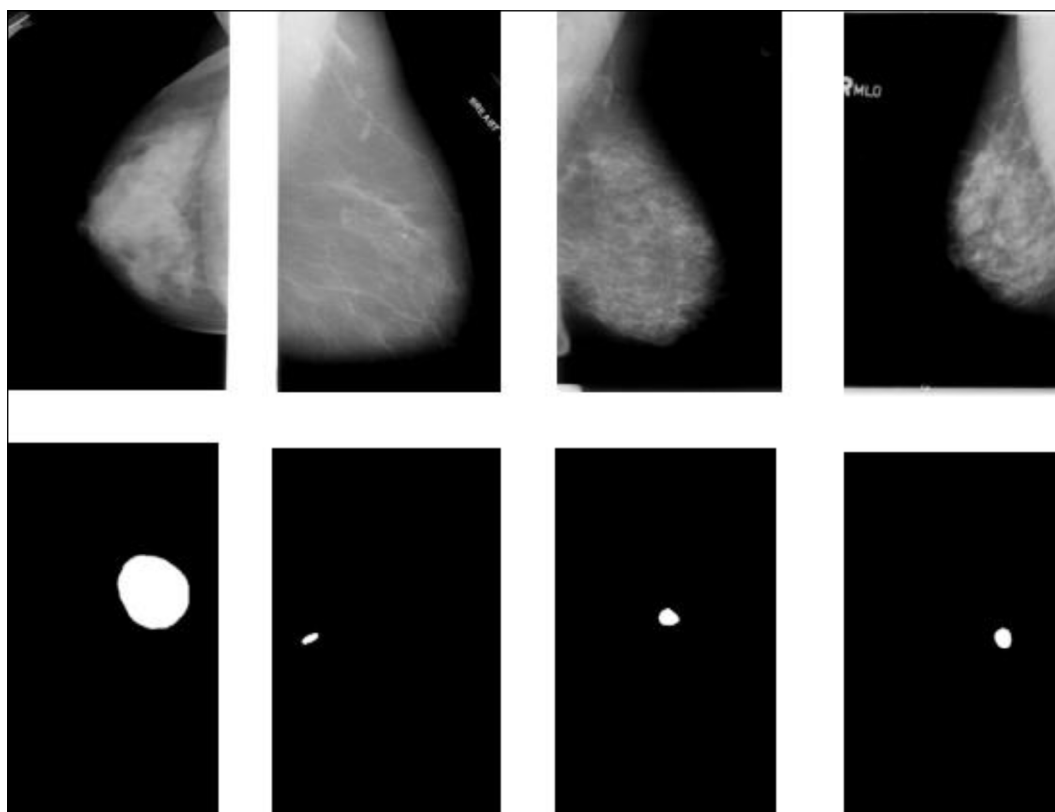


Рис. 1. Пример локализации опухоли [3].

В работе А.Н. Ручая и соавт. [10] используется метод классификации патологий молочной железы при помощи свёрточной ИНС в цифровой маммографии, позволяя с точностью 95,8% классифицировать патологии молочной железы.

В другом исследовании основной задачей было выполнение тестирования обученных моделей на наборе данных, значительно отличающихся от тех, которые использовались при их обучении [11]. При этом применялась предобработка данных, которая заключалась в нормализации значений каналов изображений в диапазоне от 0 до 1. По итогу работы была выявлена проблема с классификацией изображений с изменениями в балансе цвета, контрастности и яркости. Эти изменения привели к значительному ухудшению метрик: точности, метрики среднего отклика (average recall) и ROC-кривой.

Методы машинного обучения в ультразвуковом исследовании

Несмотря на высокую точность диагностики рака с использованием маммографа, существует серьёзный недостаток — ионизирующее излучение, — поэтому данный метод применяется только для женщин старше 40 лет и с дополнительными ограничениями¹. На сегодняшний

день широко используемым методом, помимо маммографии, является УЗИ. Для данного метода также ведутся исследования с использованием методов машинного обучения для повышения точности диагноза. Преимущества УЗИ заключаются в простоте в эксплуатации, безопасности, неинвазивности и недорогой стоимости.

В исследовании Computational Intelligence and Neuroscience [12] описывается процесс диагностики рака молочной железы с использованием ультразвукового диагностического устройства при помощи модели логистической регрессии для анализа формы, границ и классификации кровотока при дифференциальной диагностике доброкачественных и злокачественных опухолей. Результатом данного исследования стала модель, которая повышает точность диагностики до 92,4%.

В другом исследовании была сконструирована модель с использованием ИНС и применена к УЗИ-изображениям рака молочной железы для их сегментации [13]. Точность данного метода составила 97,3%, однако набор данных в этом исследовании невелик (90 пациенток), что может сказаться непредсказуемым образом на результатах работы метода при валидации.

В исследовании X. Zheng и соавт. [14] проводились измерения при помощи УЗИ с использованием свёрточной

¹ Приказ Министерства здравоохранения Российской Федерации № 404н от 27 апреля 2021 г. (ред. от 1 февраля 2022 г.) «Об утверждении Порядка проведения профилактического медицинского осмотра и диспансеризации определенных групп взрослого населения». Режим доступа: https://rcmp-nso.ru/profila/m_mater/docs1/order_rf404n.pdf?ysclid=m0q99isd18984245105 Дата обращения: 09.01.2024.

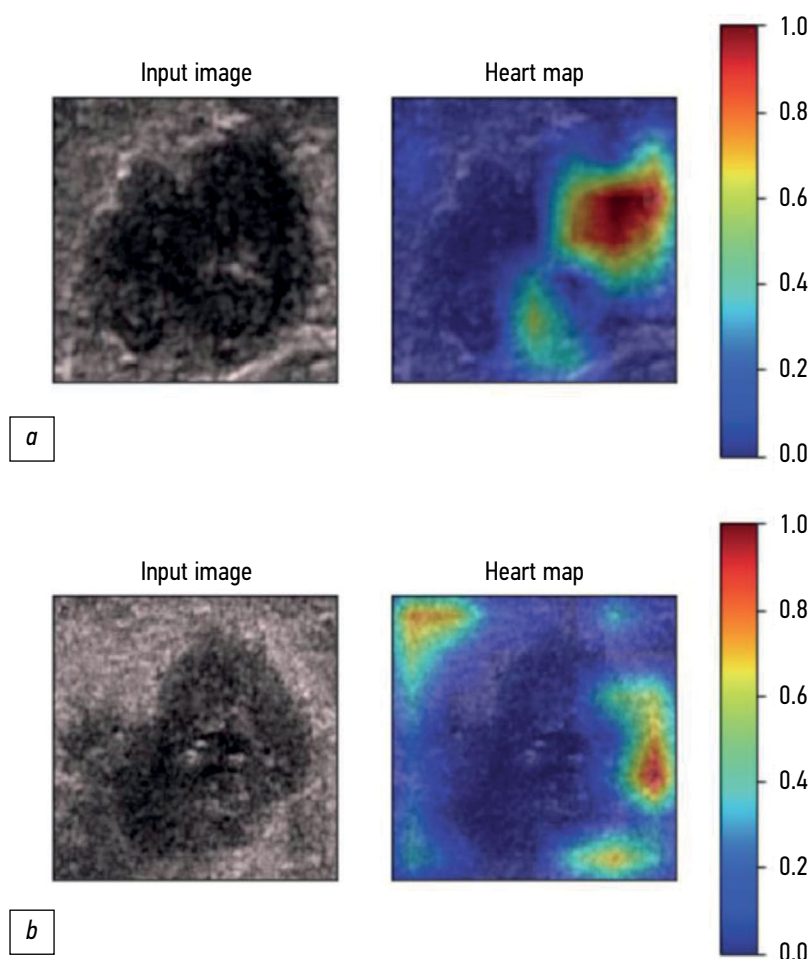


Рис. 2. Визуализация двух примеров пациентов [14]: *a* — низкочастотная область внутри опухоли ценна для предсказания статуса околососудистых лимфатических узлов; *b* — граница опухоли.

ИНС для идентификации ранней стадии рака молочной железы. На рис. 2 показана визуализация двух примеров пациентов. Каждый пример показывает изображение ультразвукового сканирования в градациях серого и соответствующую карту сердца, красная область представляет собой большой вес, который можно расшифровать с помощью цветовой шкалы справа. Рис. 2, *a* показывает, что низкочастотная область внутри опухоли ценна для предсказания статуса околососудистых лимфатических узлов, в то время как на рис. 2, *b* границей опухоли служит сама опухоль.

Но у УЗИ также есть недостатки, результаты могут зависеть от опыта и навыков врача, проводящего исследование, также УЗИ не всегда обнаруживает маленькие опухоли или ранние стадии рака молочной железы [15].

Методы машинного обучения в магнитно-резонансной томографии

На сегодняшний день высокоточным методом определения рака молочной железы является МРТ с применением внутривенного контрастирования. По данным M.S. Jochelson и соавт. [16], МРТ выявляет онкологию

в 96% случаев, что относительно выше по сравнению с маммологией и УЗИ, но при этом разработки методов с применением машинного обучения также ведутся в этом направлении для повышения точности диагностики.

В исследовании J. Witowski и соавт. [17] применялся метод глубокого обучения на основе МРТ для вычисления вероятности развития рака молочной железы. Исходя из результатов исследования, разработанная система показала результат ROC-AUC, равный 0,92.

В похожем исследовании использовали свёрточную ИНС для классификации изображений МРТ [18]. Точность сети составила 98,33%, а частота ошибок 0,0167.

В другом исследовании предлагается метод нормализации МРТ-сканов, поскольку изображения, генерируемые различными МРТ-сканерами, отличаются как по интенсивности, так и по распределению шума, что не позволяет алгоритмам, обученным на МРТ одного сканера, обобщать данные с других сканеров [19]. Была использована циклически согласованная генеративная состязательная сеть. Результатом метода стали нормализованные изображения, потенциально позволившие улучшить диагностику и выявление рака молочной железы.

Однако у данного метода диагностики есть недостатки: это достаточно дорогостоящая процедура, требующая высококвалифицированных специалистов и доступности томографов [20].

ПРЕДВАРИТЕЛЬНАЯ ОЦЕНКА ДИАГНОСТИКИ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

Помимо приведённых выше методов диагностики, существуют подходы, которые позволяют выявлять факторы риска и на этом основании предлагать более обстоятельное обследование [21–23]. Один из данных методов — микроволновая радиотермометрия с использованием компьютеризированного диагностического комплекса [21]. Этот комплекс разработан для измерения интенсивности естественного электромагнитного излучения, которое испускают внутренние ткани пациента в диапазоне сверхвысоких частот. Интенсивность излучения пропорциональна температуре тканей. Данный метод используется в онкологии, неврологии, урологии и гинекологии. Его главное преимущество в том, что он безопасен, не имеет противопоказаний и может быть использован при диагностике и контроле лечения различных заболеваний. На рис. 3 показан пример термограммы здорового пациента.

Однако анализ и интерпретация термометрических данных является довольно сложной задачей [24].

Другой метод, основанный на измерении теплового излучения с помощью инфракрасной камеры, — термография. Данный метод позволяет улавливать инфракрасное излучение и преобразовывать его в изображение — термограмму, которая регистрирует распределение тепла на поверхности тела [25].

Тепловая скintiография — это ещё один способ диагностики рака молочной железы с использованием радиоактивных веществ [26]. С помощью инфракрасной камеры можно увидеть, как в организме они скапливаются в раковых клетках, что позволяет обнаружить опухоль на снимке.

Помимо обработки тепловых данных, существуют модели машинного обучения, где в качестве входной информации используются генетические данные, связанные с раком молочной железы.

Подобные методы могут помочь в диагностике рака молочной железы на ранних стадиях, так как позволяют выявить проблему недорогими способами и при необходимости отправить пациента на традиционный метод диагностики. Однако стоит помнить, что они не дают полной картины и не позволяют произвести диагностику рака молочной железы, так как повышение температуры или мутация определённых генов не всегда связаны с раком молочной железы.

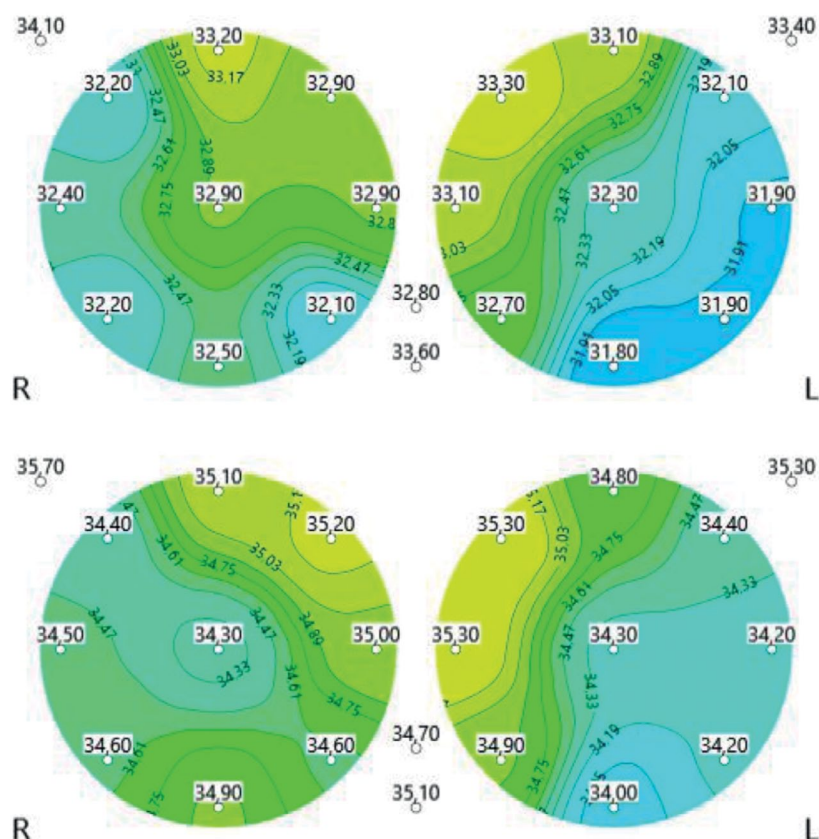


Рис. 3. Пример термограммы здорового пациента.

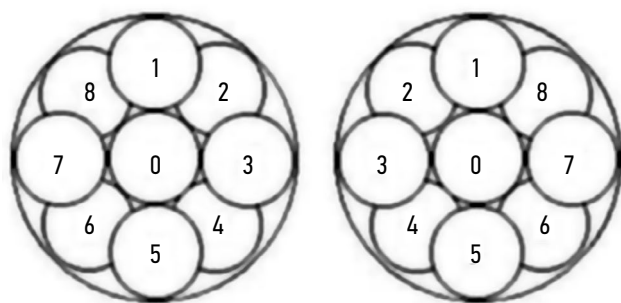


Рис. 4. Расположение точек для получения температур.

В данных методах машинное обучение используется как один из инструментов выявления факторов риска, так как получаемые данные затруднительны для понимания, а машинное обучение позволяет получить информацию, понятную врачу, дающую дополнительные доводы по необходимости проведения дополнительного обследования пациента. В рамках данного раздела было проанализировано около 90 работ, связанных с применением моделей машинного обучения для скрининга рака молочной железы. Ниже приводятся некоторые примеры.

Микроволновая радиотермометрия

В исследовании J. Li и соавт. [27] применяется нечувствительная к весам ИНС для определения рака молочной железы при помощи микроволновой радиотермометрии. Результаты показали, что данная модель обладает точностью 92%.

В другой работе применялся генетический алгоритм для локализации злокачественных новообразований в молочной железе на основе данных микроволновой радиотермометрии [28]. Данные были получены путём снятия температур с 18 точек на молочной железе, приведённых на рис. 4. Результаты работы алгоритма составили точность от 55% до 65% на тестовых выборках.

В исследовании Т.В. Замечник и соавт. [29] была разработана консультативная интеллектуальная система, которая помогает врачу быстро ориентироваться в изменениях температурных полей при постановке диагноза.

Приведённые исследования описывают потенциал микроволновой радиотермометрии, использующий диагностический инструмент на основе машинного обучения для выявления риска рака. Однако у данного метода есть свои недостатки. Не все показания данного инструмента могут быть полезными для диагностики рака, и необходимо предварительно подготовить список признаков, которые могут быть использованы [23]. Кроме того, микроволновая радиотермометрия выявляет только тепловые изменения, которые в первую очередь зависят от скорости роста опухоли, и при небольшой скорости оказывается неэффективной.

Термография

В исследовании S.T. Kakileti и соавт. [30] применялся метод термографии для определения рака молочной

железы на основе машинного обучения и был достигнут показатель ROC-AUC в 0,89.

В другом исследовании была разработана система с использованием термографии и метода опорных векторов для обнаружения рака молочной железы [4]. Достижимая при этом точность составила 96,57%. Кроме того, результатами исследования стала точность в различении доброкачественных и злокачественных случаев с точностью классификации 92,70% по сравнению с 82,05% при маммографии.

В похожем исследовании использовали свёрточную ИНС InceptionV3 для распознавания тепловых изображений молочной железы для выявления рака [31]. Полученная модель классифицирует изображение как «больное» с достоверностью 0,78, как «здоровое» с достоверностью 0,94.

В исследовании М.В. Макаровой и соавт. [32] использовалась инфракрасная термометрия в связке со средой SPSS, в которой присутствуют методы машинного обучения. Чувствительность термографии в данном исследовании у женщин в возрастной группе до 55 лет составила 89,9%, а специфичность в той же группе — 76,4%. У женщин старше 55 лет эти показатели составили: чувствительность — 90,3% для термометрии, специфичность — 78,9%.

Несмотря на высокую эффективность, термография представляет собой метод исследования организма человека, который не даёт возможности визуализировать анатомические структуры, а лишь регистрирует инфракрасное тепловое излучение, исходящее от поверхности кожи, что не всегда является фактором развития рака молочной железы.

Генетический анализ

В исследовании использовались алгоритмы машинного обучения, такие как обобщённая линейная модель, случайный лес, градиентный бустинг и модель глубокого обучения для выявления сложных закономерностей в зародышевой линии ДНК, которые коррелируют с заболеваемостью раком молочной железы [5]. Проводился анализ закономерностей при измерении изменений длины зародышевой линии ДНК в хромосомном масштабе. Эти измерения представляют собой сумму многих вставок, делений и изменений количества копий в хромосоме. В результате был разработан программный продукт с применением алгоритма градиентного бустинга, ROC-AUC которого составила 0,83 для бинарной классификации.

В другом исследовании, помимо информации генетических профилей, использовались такие данные, как образ жизни и репродуктивные факторы в модели логистической регрессии для выявления риска рака молочной железы [2]. Результатом работы модели стали ROC-AUC=0,73 и выводы о том, что данная модель может служить отправной точкой для скрининга рака молочной железы.

В работе N.A.K. Hussein и соавт. [33] была разработана система, помогающая врачу классифицировать рак

молочной железы по экспрессии генов. Применяли сразу несколько методов машинного обучения: рекуррентную ИНС, метод k-ближайших соседей, вероятностную ИНС, свёрточную ИНС и генетические алгоритмы. Лучшие результаты для процесса классификации выбирались путём объединения каждого метода с использованием генетического алгоритма и сравнения их по точности. Точность предложенного метода составила 97%.

В исследовании В. Yagin и соавт. [6] представлена модель, сочетающая алгоритмы градиентного бустинга и объяснимый искусственный интеллект для прогнозирования метастазов рака молочной железы на основе геномных данных. Результатом стала модель, показавшую точность 96% и ROC-AUC 99,3%.

У рассмотренного метода есть некоторые преимущества по сравнению с другими методами диагностики, такие как высокая точность — генетический анализ может выявить наличие конкретных генетических мутаций, которые могут быть связаны с развитием рака молочной железы; раннее выявление — метод помогает выявить генетические мутации, которые могут predispose к развитию болезни, ещё до появления первых симптомов.

Однако имеются и недостатки: высокая стоимость и ограниченность данных. На данный момент не все генетические мутации, связанные с раком молочной железы, известны. Это ограничивает возможности генетического анализа в полной и точной идентификации всех рисков и признаков развития болезни, но может служить отправной точкой для более высокоточных методов диагностики.

Лингвистическое описание состояния пациента

Помимо использования методов машинного обучения для анализа данных, полученных при обследовании пациентов, существуют подходы, в которых в качестве входных данных используются анамнез, заключение врача по УЗИ и другие лингвистические показатели. Так, в работе А.А.Р. Magna и соавт. [34] применяются методы глубокого обучения для системы рекомендаций по диагностике рака молочной железы с использованием историй болезни пациентов. В качестве входной информации выступали такие данные, как клиническая анатомия, тип заболевания, данные медицинских консультаций, пройденных процедур и принимаемых лекарств. Результаты 20 экспериментов с 5-кратной перекрёстной проверкой дали среднее значение точности измерений (Assigasy) и отзыва (Recall), равной 98% для классификации «рак» по сравнению с «не рак» и 98,6% для классификации «рак молочной железы» по сравнению с «другой рак».

В другом исследовании в качестве входных данных использовались семейный анамнез и данные из анкет, заполненных пациентами [35]. Были собраны 603 анкеты, включающие данные 309 пациентов с раком молочной

железы и 294 здоровых пациентов. Методы, использованные для сбора и анализа данных, подробно описаны в исследовании [36]. В качестве модели прогнозирования рака молочной железы были исследованы три алгоритма классификатора:

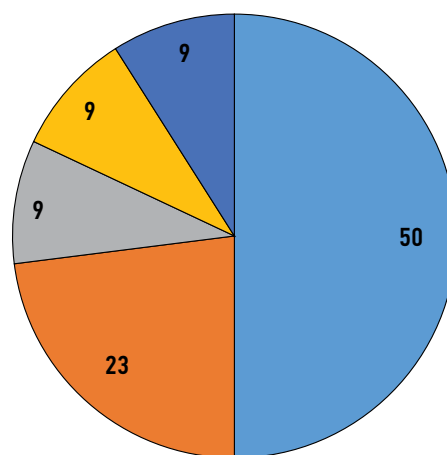
- 1) метод опорных векторов (SVM);
- 2) случайный лес (RF);
- 3) многослойный перцептрон (MLP).

Исследованные модели RF, SVM и MLP показали результаты для правильной классификации случаев рака молочной железы и здоровых случаев: средняя чувствительность >97,2%, средняя специфичность >96,4% и средняя точность >97,1%.

В исследовании М.Р. Afrash и соавт. [37] были проанализированы данные 3168 здоровых людей и 1742 истории болезни пациентов в больнице Ирана. Для прогнозирования рака молочной железы были выбраны такие показатели, как возраст, потребление молочных продуктов, семейный анамнез, биопсия молочной железы, рентген грудной клетки, гормональная терапия, употребление алкоголя, избыточный вес, наличие детей и статус образования. Результаты эксперимента показали, что дерево решений обеспечивает более высокую производительность, чем другие модели машинного обучения, со значениями точности, специфичности и чувствительности 99,3%, 99,5%, 98,26% соответственно.

ЗАКЛЮЧЕНИЕ

В рамках данной работы было проанализировано более 200 научных публикаций из открытых источников. Использовались работы из различных стран: России, США, Китая, Ирака и др. На рис. 5 приведена диаграмма, в которой показывается, какие модели машинного



- свёрточная нейронная сеть
- методы глубокого обучения
- логистическая регрессия
- метод опорных векторов
- генетические алгоритмы

Рис. 5. Частота использования моделей машинного обучения, %.

Таблица 1. Методы машинного обучения и возможные цели их применения в диагностике рака молочной железы

Решаемые задачи	Методы машинного обучения
Классификация маммограмм [7, 8, 10]; сегментация опухолей из данных маммографии [3]; классификации изображений МРТ-сканов и данных УЗИ [13, 14, 18]	Свёрточная нейронная сеть
Прогнозирование развития рака [35], выявление сложных закономерностей в ДНК [6, 33]	Методы глубокого обучения
Классификации кровотока с использованием УЗИ и генетических профилей [2, 12]	Логистическая регрессия
Классификация на основе тепловых данных [4]	Метод опорных векторов
Классификация по экспрессии генов, локализация злокачественных новообразований [28, 33]	Генетические алгоритмы

Примечание. МРТ — магнитно-резонансная томография; УЗИ — ультразвуковое исследование.

обучения чаще используются при диагностике рака молочной железы.

Можно выделить две основные задачи, решаемые методами машинного обучения: решение вспомогательных задач в медицинской диагностике рака молочной железы и предварительная оценка диагностики рака молочной железы.

В рамках первой задачи можно выделить исследования, посвящённые классификации образцов тканей на злокачественные и доброкачественные. Кроме классификации, методы машинного обучения также могут использоваться для идентификации факторов риска развития рака молочной железы.

Другим примером является анализ молекулярных данных, например, геномных. Такие задачи могут быть решены с помощью различных методов машинного обучения, которые представлены в табл. 1.

Использование методов машинного обучения при диагностике рака молочной железы предоставляет большие возможности для улучшения точности и эффективности диагностики, а также позволяет решать дополнительные задачи.

Так, например, свёрточные ИНС обладают способностью выявлять сложные иерархические признаки в изображениях. Это значит, что ИНС может автоматически изучать такие характеристики, как асимметрия тепловых полей, неоднородность, микрокальцинаты и т.д., которые являются ключевыми для диагностики рака молочной железы. Это позволяет сети выявлять патологические изменения, которые могут быть неочевидными для человеческого глаза.

Преимуществом применения вероятностной ИНС в маммографии является её способность обрабатывать неопределённость и учитывать вероятности различных классов. Это позволяет получить более адекватную и надёжную диагностику, особенно при работе с неоднозначными и сложными случаями.

Одним из основных преимуществ глубокого обучения является его способность автоматически извлекать сложные иерархические признаки из больших объёмов данных. Для прогнозирования развития рака молочной железы это может быть особенно полезно, так как медицинские

данные, используемые для диагностики, содержат множество характеристик, которые могут быть сложными для восприятия человеком.

Применение логистической регрессии для классификации кровотока с использованием УЗИ позволяет создать модель, которая может помочь в диагностике рака молочной железы. Основным преимуществом данной модели является её устойчивость к наличию выбросов и шума в данных. Это означает, что она может обрабатывать данные УЗИ, которые могут содержать некоторую степень шума или несовершенство, без значительного влияния на точность классификации.

Метод опорных векторов применяется для классификации рака молочной железы на основе тепловых данных. Данный метод хорошо справляется с классификацией в случае ограниченного количества обучающих данных. Это особенно полезно, когда доступно ограниченное количество случаев рака молочной железы для обучения модели.

Одним из важных аспектов диагностики рака молочной железы является выделение наиболее значимых генетических маркёров или признаков, которые могут быть связаны с предрасположенностью к данному виду рака. Генетические алгоритмы могут эффективно работать с большими объёмами данных, что особенно важно для классификации рака молочной железы, где доступно множество клинических и генетических данных.

Однако стоит учитывать ограничения и недостатки этих методов: необходимость большого объёма данных для обучения модели и проверки и валидации полученных результатов.

В результате анализа имеющихся источников можно сделать вывод, что ведётся активное применение компьютерных технологий в медицинской диагностике, в частности, методы машинного обучения, такие как ИНС, регрессия, случайный лес и др. В указанных исследованиях наглядно показано, что машинное обучение позволяет повысить точность и адекватность диагностики, в автоматическом режиме классифицировать новообразования и подготавливать данные для последующей обработки, проводить скрининговые обследования пациентов.

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Источник финансирования. Авторы заявляют об отсутствии внешнего финансирования при проведении поисково-аналитической работы.

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Вклад авторов. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, проведение поисково-аналитической работы и подготовку статьи, прочли и одобрили финальную версию перед публикацией). Наибольший вклад распределён следующим образом: И.В. Гермашев — вклад в разработку концепции исследования, методологию и общее руководство проектом; К.С. Дёмин — сбор и анализ данных, написание основных разделов статьи.

ADDITIONAL INFORMATION

Funding source. This article was not supported by any external sources of funding.

Competing interests. The authors declare that they have no competing interests.

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work. I.V. Germashev — contribution to the development of the research concept, methodology, and overall project management; K.S. Dyomin — data collection and analysis, writing the main sections of the article.

СПИСОК ЛИТЕРАТУРЫ

1. Состояние онкологической помощи населению России в 2022 году / под ред. Каприна А.Д., Старинского В.В., Шахзадовой А.О. Москва : МНИОИ им. П.А. Герцена – филиал ФГБУ «НМИЦ радиологии» Минздрава России, 2023.
2. Zou S., Lin Y., Yu X., et al. Genetic and lifestyle factors for breast cancer risk assessment in Southeast China // *Cancer Medicine*. 2023. Vol. 12, N 14. P. 15504–15514. doi: 10.1002/cam4.6198
3. Raza S.K., Sarwar S.S., Syed S.M., Khan N.A. Classification and Segmentation of Breast Tumor Using Mask R-CNN on Mammograms // *Research Square*. 2021. doi: 10.21203/rs.3.rs-523546/v1
4. Khan A.A., Arora A.S. Thermography as an Economical Alternative Modality to Mammography for Early Detection of Breast Cancer // *Journal of Healthcare Engineering*. 2021. P. 5543101. doi: 10.1155/2021/5543101
5. Ko C., Toh C., Brody J.P. Genetic risk scores for breast cancer based on machine learning analysis of chromosomal-scale length variation // *Clinical Cancer Research*. 2021. Vol. 27, Suppl. 5. P. PR-09. doi: 10.1158/1557-3265.ADI21-PR-09
6. Yagin B., Yagin F.H., Colak C., et al. Cancer Metastasis Prediction and Genomic Biomarker Identification through Machine Learning and eXplainable Artificial Intelligence in Breast Cancer Research // *Diagnostics (Basel)*. 2023. Vol. 13, N 21. P. 3314. doi: 10.3390/diagnostics13213314
7. O'Leary T.J., Mikel U.V., Becker R.L. Computer-assisted image interpretation: use of a neural network to differentiate tubular carcinoma from sclerosing adenosis // *Modern Pathology*. 1992. Vol. 5, N 4. P. 402–405.
8. Tsochatzidis L., Costaridou L., Pratikakis I. Deep Learning for Breast Cancer Diagnosis from Mammograms – A Comparative Study // *Journal of Imaging*. 2019. Vol. 5, N 3. P. 37. doi: 10.3390/jimaging5030037
9. Hamad Y.A., Simonov K., Naeem M.B. Breast Cancer Detection and Classification Using Artificial Neural Networks. In: 1st Annual International Conference on Information and Sciences (AiCIS); Nov 20–21, 2018; Fallujah. P. 51–57. doi: 10.1109/aicis.2018.00022
10. Ручай А.Н., Кобер В.И., Дорофеев К.А., и др. Классификация патологий молочной железы с использованием глубокой сверточной нейронной сети и трансферного обучения // *Информационные процессы*. 2020. Т. 20, № 4. С. 357–365.
11. Сасов Д.А., Зубков А.В., Орлова Ю.А., Турицына А.В. Классификация рака молочной железы с помощью сверточных нейронных сетей // *Инженерный вестник Дона*. 2023. № 6. С. 730–741.
12. Computational Intelligence and Neuroscience. Retracted: Value of Artificial Neural Network Ultrasound in Improving Breast Cancer Diagnosis // *Computational Intelligence and Neuroscience*. 2023. P. 9872174. doi: 10.1155/2023/9872174
13. Zhang L., Jia Z., Leng X., Ma F. Artificial Intelligence Algorithm-Based Ultrasound Image Segmentation Technology in the Diagnosis of Breast Cancer Axillary Lymph Node Metastasis // *Journal of Healthcare Engineering*. 2021. P. 8830260. doi: 10.1155/2021/8830260
14. Zheng X., Yao Z., Huang Y., et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer // *Nature Communications*. 2020. Vol. 11, N 1. P. 1236. doi: 10.1038/s41467-020-15027-z
15. Мустафин Ч.К. Современная диагностика заболеваний молочных желез // *Главный врач Юга России*. 2014. № 2. С. 20–23.
16. Jochelson M.S., Dershaw D.D., Sung J.S., et al. Bilateral contrast-enhanced dual-energy digital mammography: feasibility and comparison with conventional digital mammography and MR imaging in women with known breast carcinoma // *Radiology*. 2013. Vol. 266, N 3. P. 743–751. doi: 10.1148/radiol.12121084
17. Witowski J., Heacock L., Reig B., et al. Improving breast cancer diagnostics with deep learning for MRI // *Science Translational Medicine*. 2022. Vol. 14, N 664. P. eabo4802. doi: 10.1126/scitranslmed.abo4802
18. Yurttakal A.H., Erbay H., İkizceli T., Karacavus S. Detection of breast cancer via deep convolution neural networks using MRI images // *Multimedia Tools and Applications*. 2020. Vol. 79. P. 15555–15573. doi: 10.1007/s11042-019-7479-6

19. Modanwal G., Vellal A., Mazurowski M.A. Normalization of breast MRIs using cycle-consistent generative adversarial networks // *Computer Methods and Programs in Biomedicine*. 2021. Vol. 208. doi: 10.1016/j.cmpb.2021.106225
20. Цены на МРТ. В: Лайк Доктор [Internet]. Режим доступа: <https://like.doctor/ceny/diagnostika/mrt> Дата обращения: 09.01.2024.
21. Диагностический микроволновый радиотермометр РТМ-01-РЭС. В: Микроволновая радиотермометрия в медицине [Internet]. Режим доступа: <http://www.radiometry.ru/radiometry/mammology/> Дата обращения: 09.01.2024.
22. Polyakov M.V., Popov I.E., Losev A.G., Khoperskov A.V. Application of computer simulation results and machine learning in analysis of microwave radiothermometry data // *Mathematical Physics and Computer Simulation*. 2021. Vol. 24, N 2. P. 27–37. doi: 10.15688/mpcm.jvolsu.2021.2.3
23. Гермашев И.В., Дубовская В.И., Лосев А.Г., Попов И.Е. Факторный анализ влияния признаков на точность диагностики рака молочной железы по данным микроволновой радиотермометрии // *Прикаспийский журнал: управление и высокие технологии*. 2022. № 1. С. 139–148.
24. Лосев А.Г., Медведев Д.А., Применение нейронных сетей в диагностике рака молочной железы по данным микроволновой радиотермометрии // *Современная наука и инновации*. 2019. № 4. С. 22–28.
25. Шушарин А.Г., Морозов В.В., Половинка М.П. Медицинское тепловидение – современные возможности метода // *Современные проблемы науки и образования*. 2011. № 4.
26. Тицкая А.А., Чернов В.И., Синилкин И.Г., и др. Стандартизованные методики радионуклидной диагностики. Маммосцинтиграфия. Москва : ООО «НТЦ Амплитуда», 2014.
27. Li J., Galazis C., Popov L., et al. Dynamic Weight Agnostic Neural Networks and Medical Microwave Radiometry (MWR) for Breast Cancer Diagnostics // *Diagnostics (Basel)*. 2022. Vol. 12, N 9. P. 2037. doi: 10.3390/diagnostics12092037
28. Глазунов В.А. Тестирование алгоритма локализации опухоли при раке молочной железы на результатах моделирования температурных полей // XXV Региональная конференция молодых исследователей Волгоградской области : Тезисы докладов, Волгоград, 20 ноября – 13 2020 года / Редколлегия: А.Э. Калинина (отв. ред.) [и др.]. Волгоград: Волгоградский государственный университет, 2021. С. 343–347. EDN: DJQDLK
29. Замечник Т.В., Лосев А.Г., Левшинский В.В. Результаты оптимизации диагностических признаков рака молочной железы, выявляемых методом микроволновой радиотермометрии // *Медицинский вестник Северного Кавказа*. 2019. Т. 14, № 1.1. С. 48–52. doi: 10.14300/mnnc.2019.14047
30. Kakileti S.T., Madhu H.J., Manjunath G., et al. Personalized risk prediction for breast cancer pre-screening using artificial intelligence and thermal radiomics // *Artificial Intelligence in Medicine*. 2020. Vol. 105. P. 101854. doi: 10.1016/j.artmed.2020.101854
31. Mambou S.J., Maresova P., Krejcar O., et al. Breast Cancer Detection Using Infrared Thermal Imaging and a Deep Learning Model // *Sensors (Basel)*. 2018. Vol. 18, N 9. P. 2799. doi: 10.3390/s18092799
32. Макарова М.В., Юницына А.В. Тепловизионное исследование молочных желез в оценке объемных образований // *Вестник Северного (Арктического) федерального университета. Серия «Медико-биологические науки»*. 2013. № 4. С. 44–50.
33. Hussein N.A.K., Al-Sarray B. Deep Learning and Machine Learning via a Genetic Algorithm to Classify Breast Cancer DNA Data // *Iraqi Journal of Science*. 2022. Vol. 63, N 7. P. 3153–3168. doi: 10.24996/ij.s.2022.63.7.36
34. Magna A.A.R., Allende-Cid H., Taramasco C., et al. Application of Machine Learning and Word Embeddings in the Classification of Cancer Diagnosis Using Patient Anamnesis // *IEEE Access*. 2020. Vol. 8. P. 106198–106213. doi: 10.1109/ACCESS.2020.3000075
35. Mortazavi S.A.R., Tahmasebi S., Par-Saei H., Taleie A. Machine Learning Models for Predicting Breast Cancer Risk in Women Exposed to Blue Light from Digital Screens // *Journal of Biomedical Physics and Engineering*. 2022. Vol. 12, N 6. P. 637–644. doi: 10.31661/jbpe.v0i0.2105-1341
36. Mortazavi S.A.R. The Association of Screen Time and Female Breast Cancer – A Retrospective Case-Control Study: dissertation. Shiraz : Shiraz University of Medical Sciences, 2021.
37. Afrash M.R., Bayani A., Shanbehzadeh M., et al. Developing the breast cancer risk prediction system using hybrid machine learning algorithms // *Journal of Education and Health Promotion*. 2022. Vol. 11, N 1. P. 272. doi: 10.4103/jehp.jehp_42_22

REFERENCES

1. Kaprin AD, Starinsky VV, Shakhzadova AO, editors. *Sostoyaniye onkologicheskoy pomoshchi naseleniyu Rossii v 2022 godu*. Moscow: P.A. Herzen Institute of Medical Sciences – branch of the Federal State Budgetary Institution “NMIC of Radiology” of the Ministry of Health of the Russian Federation; 2023. (In Russ.)
2. Zou S, Lin Y, Yu X, et al. Genetic and lifestyle factors for breast cancer risk assessment in Southeast China. *Cancer Medicine*. 2023;12(14):15504–15514. doi: 10.1002/cam4.6198
3. Raza SK, Sarwar SS, Syed SM, Khan NA. Classification and Segmentation of Breast Tumor Using Mask R-CNN on Mammograms. *Research Square*. 2021. doi: 10.21203/rs.3.rs-523546/v1
4. Khan AA, Arora AS. Thermography as an Economical Alternative Modality to Mammography for Early Detection of Breast Cancer. *Journal of Healthcare Engineering*. 2021:5543101. doi: 10.1155/2021/5543101
5. Ko C, Toh C, Brody JP. Genetic risk scores for breast cancer based on machine learning analysis of chromosomal-scale length variation. *Clinical Cancer Research*. 2021;27 Suppl. 5:PR-09. doi: 10.1158/1557-3265.ADI21-PR-09
6. Yagin B, Yagin FH, Colak C, et al. Cancer Metastasis Prediction and Genomic Biomarker Identification through Machine Learning and eXplainable Artificial Intelligence in Breast Cancer Research. *Diagnostics (Basel)*. 2023;13(21):3314. doi: 10.3390/diagnostics13213314
7. O’Leary TJ, Mikel UV, Becker RL. Computer-assisted image interpretation: use of a neural network to differentiate tubular carcinoma from sclerosing adenosis. *Modern Pathology*. 1992;5(4):402–405.
8. Tsochatzidis L, Costaridou L, Pratikakis I. Deep Learning for Breast Cancer Diagnosis from Mammograms – A Comparative Study. *Journal of Imaging*. 2019;5(3):37. doi: 10.3390/jimaging5030037

9. Hamad YA, Simonov K, Naeem MB. Breast Cancer Detection and Classification Using Artificial Neural Networks. In: *1st Annual International Conference on Information and Sciences (AiCIS); Nov 20–21, 2018; Fallujah*. P. 51–57. doi: 10.1109/aicis.2018.00022
10. Ruchai AN, Kober VI, Dorofeev KA, et al. Classification of breast pathologies using a deep convolutional neural network and transfer learning. *Information processes*. 2020;20(4):357–365.
11. Sasov DA, Zubkov AV, Orlova YuA, Tupitsyna AV. Classification of breast cancer using convolutional neural networks. *Inženernyj vestnik Dona*. 2023;6:730–741. (In Russ.)
12. Computational Intelligence and Neuroscience. Retracted: Value of Artificial Neural Network Ultrasound in Improving Breast Cancer Diagnosis. *Computational Intelligence and Neuroscience*. 2023;9872174. doi: 10.1155/2023/9872174
13. Zhang L, Jia Z, Leng X, Ma F. Artificial Intelligence Algorithm-Based Ultrasound Image Segmentation Technology in the Diagnosis of Breast Cancer Axillary Lymph Node Metastasis. *Journal of Healthcare Engineering*. 2021;8830260. doi: 10.1155/2021/8830260
14. Zheng X, Yao Z, Huang Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nature Communications*. 2020;11(1):1236. doi: 10.1038/s41467-020-15027-z
15. Mustafin CK. Sovremennaya diagnostika zabolevanij molochnyh zhelez. *Glavnyj vrač Ūga Rossii*. 2014;(2):20–23. (In Russ.)
16. Jochelson MS, Dershaw DD, Sung JS, et al. Bilateral contrast-enhanced dual-energy digital mammography: feasibility and comparison with conventional digital mammography and MR imaging in women with known breast carcinoma. *Radiology*. 2013;266(3):743–751. doi: 10.1148/radiol.12121084
17. Witowski J, Heacock L, Reig B, et al. Improving breast cancer diagnostics with deep learning for MRI. *Science Translational Medicine*. 2022;14(664):eabo4802. doi: 10.1126/scitranslmed.abo4802
18. Yurttakal AH, Erbay H, İkizceli T, Karacavus S. Detection of breast cancer via deep convolution neural networks using MRI images. *Multimedia Tools and Applications*. 2020;79:15555–15573. doi:10.1007/s11042-019-7479-6
19. Gourav Modanwal, Adithya Vellal, Maciej A. Mazurowski, Normalization of breast MRIs using cycle-consistent generative adversarial networks // *Computer Methods and Programs in Biomedicine*, Vol. 208, 2021. doi: 10.1016/j.cmpb.2021.106225
20. Ceny na MRT. In: Like Doctor [Internet] [cited 2024 Jan 9]. Available from: <https://like.doctor/ceny/diagnostika/mrt> (In Russ.)
21. Diagnosticheskij mikrovolnovyj radiotermometr RTM-01-RES. In: *Mikrovolnovaya radiotermometriya v medicine* [Internet] [cited 2024 Jan 9]. Available from: <http://www.radiometry.ru/radiometry/mammology/> (In Russ.)
22. Polyakov MV, Popov IE, Losev AG, Khoperskov AV. Application of computer simulation results and machine learning in analysis of microwave radiothermometry data. *Mathematical Physics and Computer Simulation*. 2021;24(2):27–37. doi: 10.15688/mpcm.jvolsu.2021.2.3
23. Germashev IV, Dubovskaya VI, Losev AG, Popov IE. Factor analysis of the effect of signs on the accuracy of breast cancer diagnosis according to microwave radiothermometry. *Caspian Journal: Management and High Technologies*. 2022;(1):139–148.
24. Losev AG, Medvedev DA. The use of neural networks in the diagnosis of breast cancer according to microwave radiothermometry. *Modern Science and Innovation*. 2019;(4):22–28.
25. Shusharin AG, Morozov VV, Polovinka MP. Medical thermal imaging – modern possibilities of the method. *Modern problems of science and education*. 2011;(4).
26. Titskaya AA, Chernov VI, Sinilkin IG, et al. *Standartizirovannye metodiki radionuklidnoj diagnostiki. Mammoscintigrafiya*. Moscow: NTC Amplitude; 2014. (In Russ.)
27. Li J, Galazis C, Popov L, et al. Dynamic Weight Agnostic Neural Networks and Medical Microwave Radiometry (MWR) for Breast Cancer Diagnostics. *Diagnostics (Basel)*. 2022;12(9):2037. doi: 10.3390/diagnostics12092037
28. Glazunov VA. Testing the algorithm of tumor localization in breast cancer based on the results of modeling temperature fields. // XXV Regional Conference of Young Researchers of the Volgograd region: Theses of reports, Volgograd, November 20 – 13, 2020 / Editorial board: A.E. Kalinina (ed.) [et al.]. Volgograd: Volgograd State University, 2021. P.343–347. (In Russ.) EDN: DJQDLK
29. Zamechnik TV, Losev AG, Levshinsky VV. Results of optimization of diagnostic signs of breast cancer detected by microwave radiothermometry. *Medical News of North Caucasus*. 2019;14(1.1):48–52. doi: 10.14300/mnnc.2019.14047
30. Kakileti ST, Madhu HJ, Manjunath G, et al, Personalized risk prediction for breast cancer pre-screening using artificial intelligence and thermal radiomics. *Artificial Intelligence in Medicine*. 2020;105:101854. doi: 10.1016/j.artmed.2020.101854
31. Mambou SJ, Maresova P, Krejcar O, et al. Breast Cancer Detection Using Infrared Thermal Imaging and a Deep Learning Model. *Sensors (Basel)*. 2018;18(9):2799. doi: 10.3390/s18092799
32. Makarova MV, Unitsina AV. Thermal imaging of mammary glands in the assessment of volumetric formations. *Vestnik of Northern (Arctic) Federal University. Series "Humanitarian and Social Sciences"*. 2013;(4):44–50.
33. Hussein NAK, Al-Sarray B. Deep Learning and Machine Learning via a Genetic Algorithm to Classify Breast Cancer DNA Data. *Iraqi Journal of Science*. 2022;63(7):3153–3168. doi: 10.24996/ijis.2022.63.7.36
34. Magna AAR, Allende-Cid H, Taramasco C, et al. Application of Machine Learning and Word Embeddings in the Classification of Cancer Diagnosis Using Patient Anamnesis. *IEEE Access*. 2020;8:106198–106213. doi: 10.1109/ACCESS.2020.3000075
35. Mortazavi SAR, Tahmasebi S, Par-Saei H, Taleie A. Machine Learning Models for Predicting Breast Cancer Risk in Women Exposed to Blue Light from Digital Screens. *Journal of Biomedical Physics and Engineering*. 2022;12(6):637–644. doi: 10.31661/jbpe.v0i0.2105-1341
36. Mortazavi SAR. The Association of Screen Time and Female Breast Cancer – A Retrospective Case-Control Study [dissertation]. Shiraz: Shiraz University of Medical Sciences; 2021.
37. Afrash MR, Bayani A, Shanbehzadeh M, et al. Developing the breast cancer risk prediction system using hybrid machine learning algorithms. *Journal of Education and Health Promotion*. 2022;11(1):272. doi: 10.4103/jehp.jehp_42_22

ОБ АВТОРАХ

*** Дёмин Кирилл Сергеевич;**

адрес: Россия, 400062, Волгоград,

Университетский пр-т, д.100;

ORCID: 0009-0002-4571-3437;

e-mail: diominkirill@yandex.ru

Гермашев Илья Васильевич, д-р тех. наук, профессор;

ORCID: 0000-0001-5507-8508;

eLibrary SPIN: 2489-2628;

e-mail: germashev@volsu.ru

AUTHORS' INFO

*** Kirill S. Dyomin;**

address: 100 Universitetskiy avenue, 400062

Volgograd, Russia;

ORCID: 0009-0002-4571-3437;

e-mail: diominkirill@yandex.ru

Ilya V. Germashev, Dr. Sci. (Engineering);

ORCID: 0000-0001-5507-8508;

eLibrary SPIN: 2489-2628;

e-mail: germashev@volsu.ru

* Автор, ответственный за переписку / Corresponding author