На правах рукописи

Махов Денис Сергеевич

БИОТЕХНИЧЕСКАЯ СИСТЕМА АВТОМАТИЗИРОВАННОГО ОПРЕДЕЛЕНИЯ АМПЛИФИКАЦИИ ГЕНА HER2 ПО FISH-ИЗОБРАЖЕНИЯМ ПРИ ДИФФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКЕ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ

Специальность 2.2.12. Приборы, системы и изделия медицинского назначения

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата технических наук



Москва – 2026

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)».

Научный руководитель:	**Самородов Андрей Владимирович**, кандидат технических наук, доцент.
Официальные оппоненты:	**Обухов Юрий Владимирович,** доктор физико-математических наук, старший научный сотрудник, Федеральное государственное бюджетное учреждение науки Институт радиотехники и электроники им. В.А. Котельникова Российской академии наук, главный научный сотрудник; **Семенова Евгения Анатольевна**, кандидат технических наук, доцент, федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)», доцент кафедры биотехнических систем.
Ведущая организация:	Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет «Московский институт электронной техники», г. Москва, г. Зеленоград.

Защита диссертации состоится «21» октября 2026 г. в 12:00 часов на заседании Диссертационного совета 24.2.331.09 при Московском государственном техническом университете имени Н.Э. Баумана в зале Ученого совета по адресу: 105005, Москва, ул. 2-я Бауманская, д. 5, стр. 1.

С диссертацией можно ознакомиться в библиотеке МГТУ им. Н.Э. Баумана и на сайте www.bmstu.ru.

Автореферат разослан «___» _____ 2026 г.

Ваш отзыв в двух экземплярах, заверенный гербовой печатью, просьба направлять по адресу: 105005, г. Москва, 2-я Бауманская ул., д. 5, стр. 1, МГТУ им. Н.Э. Баумана, ученому секретарю диссертационного совета 24.2.331.09.

Ученый секретарь
диссертационного совета,
к.т.н., доцент — Самородов Андрей Владимирович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Каждый год число новых случаев онкологических заболеваний неуклонно растёт (Bray F. и др., 2024). У женщин рак молочной железы (РМЖ) является наиболее распространённым типом онкологии и занимает лидирующие позиции в структуре заболеваемости и смертности в России и мире (Каприн А.Д. и др., 2024, Bray F. и др., 2024). Для снижения смертности от РМЖ необходимо повышать эффективность методов ранней диагностики и осуществлять регулярный скрининг населения (Henderson J.T. и др., 2024). При установлении наличия РМЖ важно определить его гистологический и молекулярно-генетический подтип, то есть провести дифференциальную диагностику. Результаты дифференциальной диагностики позволяют подобрать наиболее эффективную тактику лечения (DeVitra V.T. и др., 2023).

Гистологическая характеристика РМЖ включает определение типа опухоли на основе ее морфологии, установление наличия инвазии и оценку степени дифференцировки (Cserni G., 2020). Гетерогенность РМЖ на молекулярно-генетическом уровне устанавливается с помощью суррогатной иммуногистохимической классификации на основе экспрессии рецепторов эстрогенов и прогестерона (ER и PR, соответственно), индекса пролиферативной активности Ki-67 и HER2-статуса (Cardoso F. и др., 2024).

HER2-положительные опухоли составляют примерно 15 % всех случаев РМЖ (Harbeck N. и др., 2019). Исторически они ассоциировались с агрессивным течением: как правило, такие опухоли низкодифференцированы, часто лишены ER и PR экспрессии и – в отсутствие таргетной терапии – характеризуются низкой общей и безрецидивной выживаемостью. Введение анти-HER2 терапии, в частности трастузумаба, кардинально изменило прогноз для этого подтипа, продемонстрировав высокую эффективность и значительно улучшив выживаемость пациенток (Slamon D. и др., 2011).

Для определения HER2-статуса опухоли первично выполняют иммуногистохимическое (ИГХ) исследование с полуколичественной оценкой мембранного окрашивания (Wolff A.C. и др., 2023). При неопределённом результате ИГХ-исследования (оценка 2+) назначается исследование методом флуоресцентной *in situ* гибридизации (FISH) для определения амплификации гена HER2 (Wolff A.C. и др., 2023). В рамках FISH-исследования проводят регистрацию цифровых изображений с использованием флуоресцентного микроскопа и последующий визуальный анализ с подсчетом флуоресцентных сигналов от меченных флуорофорами участков хромосом. На регистрируемых псевдоцветных изображениях области ядер визуализируются синим цветом (флуорофор DAPI), а внутри ядер присутствуют красные сигналы от участков гена HER2 (флуорофор TexasRed) и зеленые сигналы от участков CEP17 (флуорофор FITC). Согласно международным клиническим рекомендациям ASCO/CAP 2023 года наличие амплификации устанавливается при соотношении HER2/CEP17 больше или равном 2,0 и среднем на ядро числе копий гена HER2 больше или равном 4; при одновременном несоблюдении

данных условий устанавливается отсутствие амплификации; остальные случаи требуют проведения повторного исследования (Wolff A.C. и др., 2023).

При подсчёте сигналов и последующем определении амплификации, необходимо придерживаться инструкций производителя реагентов, локальных или международных рекомендаций (Wolff A.C. и др., 2023). Рекомендации включают в себя указания по выбору анализируемых ядер, учету кластеризации сигналов и гетерогенности их распределения по препарату, однако они не содержат формализованных критериев выбора репрезентативных областей препарата, что оставляет пространство для субъективной интерпретации и делает достоверность результатов зависимой от квалификации патолога (Rehman Z.U. и др., 2024).

Визуальный анализ FISH-изображений осложняется наличием аппаратных и биологических артефактов, часть из которых может быть частично или полностью устранена на этапе постобработки. К аппаратным артефактам относят перекрестные помехи (cross-talk), геометрические и хроматические аберрации, обусловленные характеристиками оптических элементов микроскопа, а также аналоговый и цифровой шумы в регистрирующем тракте; к биологическим – аутофлуоресценцию цитоплазмы, неспецифическое связывание зондов, недоотмытые зонды, а также некроз ядер и некорректную гибридизацию (Rehman Z.U. и др., 2024). Наличие таких артефактов делает подсчёт сигналов трудоёмким, а итоговое определение амплификации гена HER2 – зависящим как от квалификации патолога, так и от практики приготовления препаратов и используемого оборудования в конкретной лаборатории.

Для повышения объективности и воспроизводимости результатов FISH-исследования требуется его автоматизация (Xue T. и др., 2023, Кушнарев В.А., 2021). В ряде работ предложены алгоритмы автоматизации FISH-исследований как для конечного определения амплификации, так и отдельных его этапов, однако зачастую используются закрытые наборы данных, а сами модели недоступны для внешней валидации. Качество работы автоматизированных систем определения амплификации гена HER2 при FISH-исследовании РМЖ обычно оценивают по степени согласованности их результатов с заключениями патолога. Однако согласованность заключений даже между патологами не идеальна и зависит от качества изображений: для препаратов без артефактов уровень согласованности достигает каппы Коэна 0,655, тогда как при наличии артефактов степень согласованности падает до 0,450 (Zakrzewski F. и др., 2019). Вопрос системного устранения артефактов в современной литературе рассматривается лишь косвенно и остаётся актуальным направлением исследований. Соответственно, разработка и верификация метода автоматизированного анализа FISH-изображений для определения амплификации гена HER2 с учётом аппаратных и биологических артефактов представляет собой важную научно-практическую задачу.

Цель диссертационной работы: Разработка метода автоматизированного анализа цифровых изображений для определения

амплификации гена HER2 при FISH-исследовании для дифференциальной диагностики РМЖ с учетом аппаратных и биологических артефактов.

Для достижения поставленной цели в работе были сформулированы и решены следующие **задачи**:

1. Анализ особенностей формирования флуоресцентного изображения при FISH-исследовании с учетом аппаратных и биологических артефактов;
2. Обзор существующих решений в области автоматизации FISH-анализа с целью выбора базовых алгоритмов для сегментации ядер и детектирования сигналов;
3. Разработка методов прямого или косвенного устранения аппаратных и биологических артефактов для повышения качества работы базовых алгоритмов;
4. Разработка автоматизированного алгоритма поддержки принятия врачебных решений по определению амплификации гена HER2 методом FISH по выделенным бинарным маскам ядер и сигналов;
5. Разработка и апробация экспериментального образца биотехнической системы для автоматизированного определения амплификации гена HER2 по FISH-изображениям для дифференциальной диагностики РМЖ.

Научная новизна работы:

1. Показано, что при коррекции фоновой флуоресценции на FISH-изображениях с использованием аппроксимации полиномиальной поверхностью наибольший контраст Вебера для ядер достигается при использовании первой степени полинома, в отличие от широко применяемой для устранения данного артефакта аппроксимации с использованием полинома третьей степени.
2. Показано, что автоматизированное определение амплификации гена HER2 возможно на основе отношения площадей флуоресцирующих областей HER2/CEP17 вместо отношения числа соответствующих сигналов, подсчитываемого в соответствии с клиническими рекомендациями ASCO/CAP и Dako при ручном анализе. При этом необходима корректировка порогового значения, используемого для формирования автоматизированного заключения о наличии амплификации.
3. Показано, что использование стекинг-подхода к автоматизированному определению амплификации гена HER2, при котором первичное автоматизированное заключение, формируемое пороговым методом, и признаки распределения числа, площадей и интенсивности выделенных областей ядер и сигналов объединяются на входе мета-классификатора, позволяет повысить качество по сравнению с использованием только первичного автоматизированного заключения.

Теоретическая и практическая значимость работы:

1. Установлена необходимость коррекции фоновой флуоресценции для повышения качества работы автоматизированных алгоритмов сегментации ядер и детектирования сигналов на FISH-изображениях РМЖ при определении амплификации гена HER2.

2. Предложен способ адаптации клинических рекомендаций по определению амплификации гена HER2 методом FISH, предназначенных для ручного анализа. Показано, что автоматизированное определение амплификации гена HER2 возможно с использованием взвешенного среднего отношения площадей флуоресцирующих областей HER2/CEP17 с дальнейшей оптимизацией порогового значения.

3. Показана целесообразность совместного использования первичного автоматизированного заключения, формируемого пороговым методом, распределений геометрических характеристик выделенных областей ядер и сигналов и их интенсивностей в каналах регистрации для повышения качества автоматизированного определения амплификации гена HER2 по FISH-изображениям.

4. Собрана база данных верифицированных FISH-изображений, используемых для определения амплификации гена HER2 при дифференциальной диагностике РМЖ, содержащая ручную разметку областей ядер, сигналов HER2 и сигналов CEP17, а также заключение патолога об амплификации гена HER2. Общее число изображений составило 155 для 44 пациентов (21 с амплификацией и 23 – без).

5. На основании проведенных исследований создано программное обеспечение для автоматизированной поддержки принятия врачебных решений при определении амплификации гена HER2 по FISH-изображениям, использованное в научно-исследовательской практике отделения онкоцитологии МНИОИ им. П.А. Герцена – филиала ФГБУ «НМИЦ радиологии» Минздрава России.

Основные положения, выносимые на защиту:

1. Использование алгоритма коррекции фона на базе аппроксимации полиномиальной поверхностью первой степени приводит к максимизации контраста Вебера ядер относительно фона и, соответственно, к повышению коэффициента Дайса с 0,67 до 0,85 для автоматизированного алгоритма сегментации ядер. Дополнительно установлено, что применение данной коррекции также приводит к повышению коэффициента Дайса с 0,39 до 0,52 и с 0,33 до 0,40 для автоматизированной сегментации сигналов CEP17 и HER2, соответственно.

2. Использование взвешенного среднего отношения площадей флуоресцирующих областей HER2/CEP17 для формирования первичного автоматизированного заключения об амплификации гена HER2 позволяет достичь диагностических чувствительности, специфичности и эффективности (ДЧ, ДС и ДЭ, соответственно) и каппы Коэна 0,57, 0,87, 0,73 и 0,45, соответственно.

3. Было установлено, что при автоматизированном определении амплификации гена HER2 путем сравнения взвешенного среднего отношения площадей флуоресцирующих областей HER2/CEP17 с пороговым значением, последнее должно быть скорректировано. На исследованной выборке изменение порогового значения со значения 2,00, используемого в

клинических рекомендациях для ручного анализа, до 2,18 приводит к дополнительному повышению качества автоматизированного первичного заключения до уровня ДЧ, ДС, ДЭ и каппы Коэна 0,57, 0,91, 0,75 и 0,49, соответственно.

4. Применение мета-классификатора, на вход которому поступают первичное автоматизированное заключение, формируемое пороговым методом, и признаки распределения числа, площадей и интенсивностей выделенных областей ядер и сигналов, позволяет достичь на исследованной выборке ДЧ, ДС, ДЭ и каппы Коэна 0,87, 0,76, 0,79 и 0,58, соответственно.

Методы исследования. В работе использованы методы математической статистики, методы цифровой обработки сигналов, методы обработки и анализа данных, методы машинного обучения и распознавания образов, элементы теории биотехнических систем.

Достоверность результатов основывается на использовании в работе основных положений теории биотехнических систем, методов математической статистики и других известных методов исследования. Полученные результаты не противоречат общепризнанным принципам и результатам исследований, опубликованным в работах отечественных и зарубежных авторов.

Внедрение и использование. Результаты диссертационной работы использованы в научно-исследовательской деятельности отделения онкоцитологии МНИОИ им. П.А. Герцена – филиала ФГБУ «НМИЦ радиологии» Минздрава России для определения амплификации гена HER2 при дифференциальной диагностике РМЖ и в учебном процессе кафедры «Биомедицинские технические системы» МГТУ им. Н.Э. Баумана.

Апробация материалов диссертации. Основные положения и результаты диссертационной работы обсуждались на конференциях: «2018 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)» (Екатеринбург, 2018), «2019 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)» (Екатеринбург, 2019), «26th IEEE Conference of Open Innovations Association FRUCT» (Ярославль, 2020) и «Физика и радиоэлектроника в медицине и экологии (ФРЭМЭ-2022)» (Владимир, 2022), на научных семинарах факультета биомедицинской техники МГТУ им. Н.Э. Баумана.

Публикации. Основные результаты диссертационного исследования представлены в 7 научных работах, включая 3 статьи в рецензируемых журналах и изданиях из перечня ВАК РФ и 3 статьи, индексируемые в WoS/Scopus. Общий объем 3,4 п.л.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении показана актуальность диссертационного исследования, сформулированы цель и задачи работы, научная новизна и защищаемые положения, представлены общие сведения о работе.

В первой главе приведена медицинская статистика онкологических заболеваний в России и мире, из анализа которой следует, что РМЖ является

наиболее распространенным онкологическим заболеванием у женщин, а также одной из ведущих причин смертности населения. Определение подтипа РМЖ является ключевым при выборе тактики лечения и прогнозировании состояния пациента. Одним из маркеров при дифференциальной диагностике РМЖ является HER2-статус. Определение HER2-статуса проводится методом ИГХ и, в случае неопределенного ИГХ-результата, уточняется по результатам исследования методом FISH. Примеры FISH-изображений для случая отсутствия и наличия амплификации гена HER2 приведены на Рисунке 1 слева и справа, соответственно.

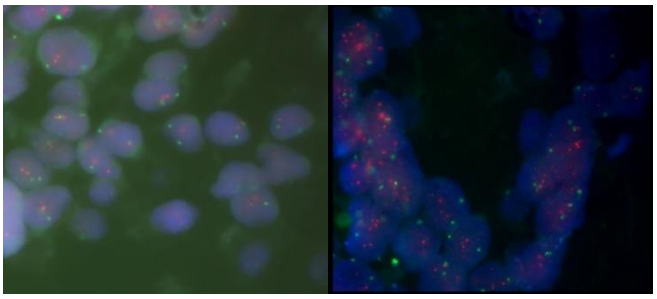


Рисунок 1. Примеры изображений FISH-исследования: слева – отсутствие, справа – наличие амплификации гена HER2

Анализ FISH-изображений проводится визуально и осложняется рядом аппаратных и биологических артефактов, на Рисунке 2 представлены некоторые из них. Наличие артефактов приводит к тому, что согласованность патологов между собой при определении амплификации гена HER2 методом FISH снижается с 0,655 каппы Коэна до 0,450 (Zakrzewski F. и др., 2019). Автоматизация данного анализа позволяет снизить нагрузку на патологов, а также повысить объективность и воспроизводимость получаемых результатов.

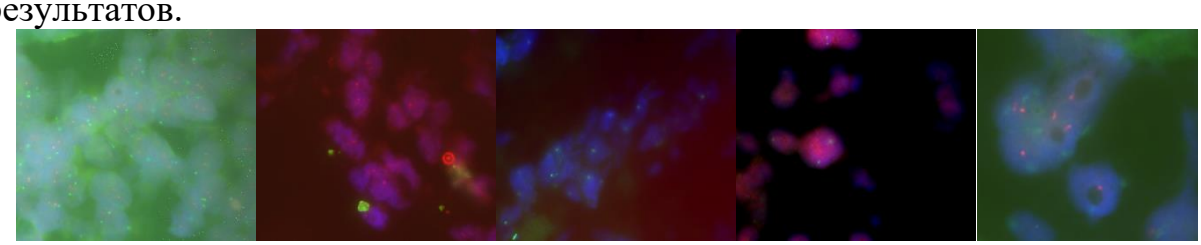


Рисунок 2. Примеры артефактов на FISH-изображениях. Слева направо: аутофлуоресценция цитоплазмы, проявляющаяся зеленоватым фоном, недоотмытые зонды, проявляющиеся красноватым фоном, неадекватная гибридизация, так как присутствуют сигналы только одного типа, неспецифическое связывание в области ядра, проявляющееся красными мелко дисперсными сигналами, некроз ядер – темные области внутри ядер из-за избыточной обработки пепсином

В работе представлен анализ современного состояния в области автоматизации определения амплификации гена HER2 методом FISH для дифференциальной диагностики РМЖ, из которого следует, что качество работы таких систем сопоставимо или даже превосходит качество ручного анализа в случае отсутствия артефактов на зарегистрированных

изображениях. Недостатками рассмотренных работ являются закрытые наборы обучающих и тестовых данных, неполное описание алгоритмов, а также отсутствие механизмов устранения аппаратных и/или биологических артефактов. Первые два недостатка приводят к невозможности сравнения имеющихся алгоритмов между собой, а последний – к невозможности обобщения результатов на FISH-изображения различного качества.

Таким образом, для разработки автоматизированного алгоритма определения амплификации гена HER2 по FISH-изображениям необходимо учитывать наличие и степень аппаратных и биологических артефактов.

Во второй главе приведен анализ особенностей формирования цифрового флуоресцентного изображения для установления причин возникновения аппаратных и биологических артефактов и разработки методов их устранения.

Для разработки автоматизированного алгоритма оценки амплификации гена HER2 по FISH-изображениям с учетом аппаратных артефактов была составлена схема формирования изображения в цифровом флуоресцентном широкопольном микроскопе (ЦФШМ). В соответствии с данной схемой была разработана математическая модель, описывающая процесс преобразования светового потока в ЦФШМ – от источника освещения до цифрового изображения, формируемого камерой:

$$I_{CCD}(i,j,k) = [\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [I^k_{фл}(x,y,z,\lambda)\cdot\delta(x-j)\cdot\delta(y-i)\cdot T_{obj}(\lambda)\cdot T^k_{em}(\lambda)\cdot T^k_{mir}(\lambda)\cdot \cdot T_{trino}(\lambda)\cdot T_{adaptor}(\lambda)\cdot R_{CCD}(\lambda)] \otimes PSF(x,y,z,\lambda)\cdot dz\cdot d\lambda]\cdot S\cdot t\cdot A_D, \quad (1)$$

где $I_{CCD}(i,j,k)$ – интенсивность пикселя цифрового изображения в позиции (i,j) в k-ом канале, обычно представляет 8-ми битную безразмерную величину;

$\delta(x)$ – дельта функция Дирака, отвечающая за пространственную регистрацию сигнала отдельным пикселем, ед;

$T_{obj}(\lambda)$ – спектральное распределение коэффициента пропускания микрообъектива, ед/нм;

$T^k_{em}(\lambda)$ – спектральное распределение коэффициента пропускания эмиссионного фильтра, ед/нм;

$T^k_{mir}(\lambda)$ – спектральное распределение коэффициента пропускания дихроического зеркала, ед/нм;

$T_{trino}(\lambda)$ – спектральное распределение коэффициента пропускания тринокуляра, ед/нм;

$T_{adaptor}(\lambda)$ – спектральное распределение коэффициента пропускания согласующего адаптера, ед/нм;

$R_{CCD}(\lambda)$ – спектральное распределение чувствительности ПЗС матрицы, ед/нм;

PSF – функция рассеяния точки оптической системы микроскопа, ед/нм;

S – площадь пикселя, м^2;

t – время экспозиции, с;

A_D – уровень выходного цифрового сигнала при заданной интенсивности, ед.

В предлагаемой модели были приняты следующие допущения:

– все оптические элементы микроскопа, за исключением флуоресцентных фильтров и дихроического зеркала, не изменяют спектрального состава излучения;

– функция рассеяния точки аппроксимируется двумерным гауссовым распределением и считается не зависящей от длины волны;

– толщина исследуемого препарата пренебрежимо мала, что позволяет использовать двумерную модель формирования изображения.

Разработанная модель позволила теоретически оценить вклад основного аппаратного артефакта – перекрестных помех (cross-talk), описываемого матрицей коэффициентов, близкой к единичной. Однако верификация на экспериментальных данных, полученных с использованием растворов отдельных флуорофоров показала, что флуоресценция одного флуорофора вносит вклад до 29 % в другой канал регистрации, что согласуется с ранее опубликованными результатами.

Биологические артефакты связаны с вариативностью свойств исследуемого биообъекта, диапазонами обработки на различных этапах приготовления препарата, а также стохастической природой межмолекулярного взаимодействия. К основным биологическим артефактам можно отнести явление фоновой флуоресценции (аутофлуоресценция цитоплазмы и недоотмытые зонды), некроз ядер, неспецифическое связывание и неадекватную гибридизацию. Современные протоколы регламентируют проведение всех этапов пробоподготовки, однако полностью избежать указанных артефактов не удается. На этапе постобработки зарегистрированных FISH-изображений часть артефактов, таких как фоновая флуоресценция, может быть частично устранена. К неустранимым биологическим артефактам можно отнести некроз ядер, некорректную гибридизацию и неспецифическое связывание, к появлению которых автоматизированная система должна быть устойчива.

Для устранения фоновой флуоресценции в литературных источниках зачастую используется аппроксимация фонового сигнала полиномиальной поверхностью третьей степени. Однако проведенный в диссертационной работе анализ зависимости контраста Вебера для различных групп изображений, разделенных по наличию артефактов, от степени полинома аппроксимирующей фоновый сигнал поверхности (Рисунок 3) показал, что первая степень дает наибольший контраст ядер для всех групп изображений. Соответственно, для дальнейших исследований по оценке влияния коррекции фона была выбрана первая степень аппроксимирующего полинома.

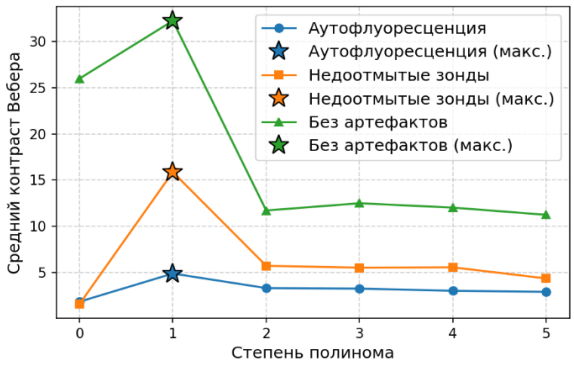


Рисунок 3. Зависимости контраста Вебера для ядер в зависимости от степени полиномиальной аппроксимации

В результате проведенного анализа установлено, что формирование цифрового флуоресцентного изображения при FISH-исследовании неотъемлемо сопровождается влиянием аппаратных и биологических артефактов. Разработка стандартизированных методик пробоподготовки и анализа позволяет снизить влияние ряда перечисленных выше артефактов, однако при разработке автоматизированного алгоритма определения амплификации HER2 по FISH-изображениям РМЖ необходимо в явном или косвенном виде устранять их влияние на финальное заключение.

Третья глава посвящена разработке биотехнической системы (БТС) автоматизированного определения амплификации HER2 при FISH-исследовании для дифференциальной диагностики РМЖ с учетом аппаратных и биологических артефактов.

Центральное место разрабатываемой БТС занимает программная подсистема, реализующая следующие этапы:

– устранение артефактов на цифровом изображении;
– сегментация ядер и детектирование сигналов;
– оценка наличия или отсутствия амплификации по полученным бинарным маскам.

Для оценки качества работы разработанной БТС были использованы диагностические чувствительность (ДЧ), специфичность (ДС) и эффективность (ДЭ), а также каппа Коэна для оценки степени согласованности автоматизированного заключения с патологом. Для обучения и верификации алгоритмов устранения артефактов, сегментации ядер и детекции сигналов на базе отделения онкоцитологии МНИОИ им. П.А. Герцена – филиала ФГБУ «НМИЦ радиологии» Минздрава России (руководитель – профессор, д.м.н. Волченко Н.Н., верификация базы данных – ведущий научный сотрудник, д.м.н. Славнова Е.Н.) была собрана база FISH-изображений (выборка №1) для 8 пациентов (4 – с амплификацией и 4 – без, всего 39 изображений), которые были размечены вручную с указанием контуров ядер и контуров сигналов от флуоресцентных зондов. Для оценки качества алгоритма определения амплификации была собрана верифицированная база из 155 изображений для 44 пациентов (21 с амплификацией, 23 – без амплификации) (выборка №2), для которых было установлено наличие или отсутствие амплификации гена HER2.

В результате оценки наличия артефактов на собранных изображениях в выборке №2 было установлено, что только 4, соответствующие одному пациенту, могут быть отнесены к изображениям с малым числом артефактов.

Задачи сегментации ядер и детекции сигналов были сформулированы как задачи бинарной сегментации, так как это кратно снижает трудозатраты на разметку данных, а также снижает так называемый шум разметки. Бинарная сегментация подразумевает предсказание маски, которая сопоставляет областям объектов логическую единицу и логический ноль – остальным областям. Для оценки качества работы алгоритмов сегментации ядер использовался коэффициент Дайса, вычисление которого проводилось по следующей формуле:

$$Dice = \frac{2 \cdot Pred \cap True}{True \cup Pred}, \quad (2)$$

где $Pred$ – бинарная маска на выходе алгоритма сегментации,
$True$ – бинарная маска, размеченная экспертом.

На выборке №1 было проведено сравнительное исследование базовых алгоритмов сегментации на основе методов компьютерного зрения и машинного обучения, однако использование нейросетевого подхода дало наилучшие результаты. Из проведенного анализа результатов работы базовых алгоритмов на примере алгоритма сегментации ядер было установлено, что классические алгоритмы демонстрируют нестабильное качество работы в зависимости от наличия и степени артефактов на изображении. Это вызвано неспособностью базовых алгоритмов учитывать априорные сведения об исследуемых объектах, в частности об их размерах и форме, а также локальные характеристики FISH-изображений. Использование сверточной нейронной сети (СНС) на базе U-net архитектуры позволило получить в задаче сегментации ядер коэффициент Дайса 0,67 против 0,62 для сбалансированного случайного леса – лучшего метода из методов машинного обучения в проведенном исследовании. Для детектирования сигналов использование СНС также позволило достичь наилучших метрик: коэффициент Дайса составил 0,39 и 0,33 для сигналов CEP17 и HER2, соответственно. Дальнейшие исследования проводились с использованием трех СНС: по одной на каждую задачу сегментации.

Нейросетевые алгоритмы позволяют учитывать крайне сложные пространственные локальные взаимосвязи, а также взаимодействия между каналами исходного изображения. Поэтому линейная зависимость каналов, вызванная перекрестными помехами, при использовании таких алгоритмов косвенным образом учтена в параметрах первого сверточного слоя. Шум на флуоресцентных изображениях для СНС при решении задачи бинарной сегментации не критичен, так как СНС может быть рассмотрена как последовательность нелинейных фильтров, позволяющих неявно устранять шум во внутренних представлениях сети. Одной из особенностей СНС, к которым относится U-net, является ограниченное рецептивное поле и

отсутствие механизмов адаптивного перестроения фильтров при смене пространственной локализации (веса свертки остаются фиксированными при свертке с входным тензором). Поскольку фоновая флуоресценция варьируется в масштабах, превышающих рецептивное поле сети, она не может быть эффективно скомпенсирована на уровне ее внутренних представлений. Как было установлено во второй главе для достижения максимального контраста Вебера ядер относительно фона необходимо использовать аппроксимацию фоновой флуоресценции полиномиальной поверхностью первой степени. В Таблице 1 приведены средние значения и среднеквадратичные отклонения (СКО) через «/» метрик качества работы нейросетевых алгоритмов для сегментации ядер и детектирования сигналов без и с коррекцией фона на основе полиномиальной поверхности первого порядка, обученных и протестированных на выборке №1.

Таблица 1.

Влияние коррекции фона на качество сегментации ядер и сигналов

Метрика	**Ядра**		**Сигналы CEP17**		**Сигналы HER2**	
	Без коррекции фона	С коррекцией фона	Без коррекции фона	С коррекцией фона	Без коррекции фона	С коррекцией фона
Коэффициент Дайса	0,67 / 0,33	**0,85 / 0,05**	0,39 / 0,22	**0,52 / 0,25**	0,33 / 0,30	**0,40 / 0,30**

В связи с малым в контексте глубокого обучения числом изображений было предложено использовать аугментации на базе геометрических преобразований: обрезка и отражение, а также гауссово размытие, которое позволяет имитировать расфокусированное изображение. Для учета вариативности свойств исследуемого биообъекта была использована аугментация на базе гамма-преобразования со случайным показателем степени для имитации различного контраста объектов на изображении. В Таблице 2 приведены выборочные средние коэффициента Дайса для алгоритмов сегментации ядер и детекции сигналов с указанием СКО через «/», демонстрирующие влияние различных видов аугментаций, полученные в результате обучения и тестирования на выборке №1 при использовании коррекции фона.

Таблица 2.

Коэффициент Дайса в задачах сегментации ядер и сигналов с применением различных комбинаций аугментаций при использовании коррекции фона

Аугментация на базе геом. преобр.	**Аугментация на базе гамма-преобр.**	**Сегментация ядер**	**Детектирование сигналов CEP17**	**Детектирование сигналов HER2**
–	–	**0,85 / 0,05**	0,52 / 0,25	0,40 / 0,31
–	+	0,76 / 0,07	**0,70 / 0,07**	0,43 / 0,28
+	–	0,67 / 0,36	0,51 / 0,30	0,49 / 0,29
+	+	0,72 / 0,05	0,69 / 0,10	**0,55 / 0,21**

Из полученных результатов следует, что при использовании коррекции фона аугментация данных приводит к улучшению сегментации сигналов.

Таким образом, в результате проведенных исследований были разработаны алгоритмы сегментации ядер и детектирования сигналов, способные в явном или неявном виде учитывать биологические и аппаратные артефакты. Примеры работы разработанных алгоритмов представлены на Рисунке 4.

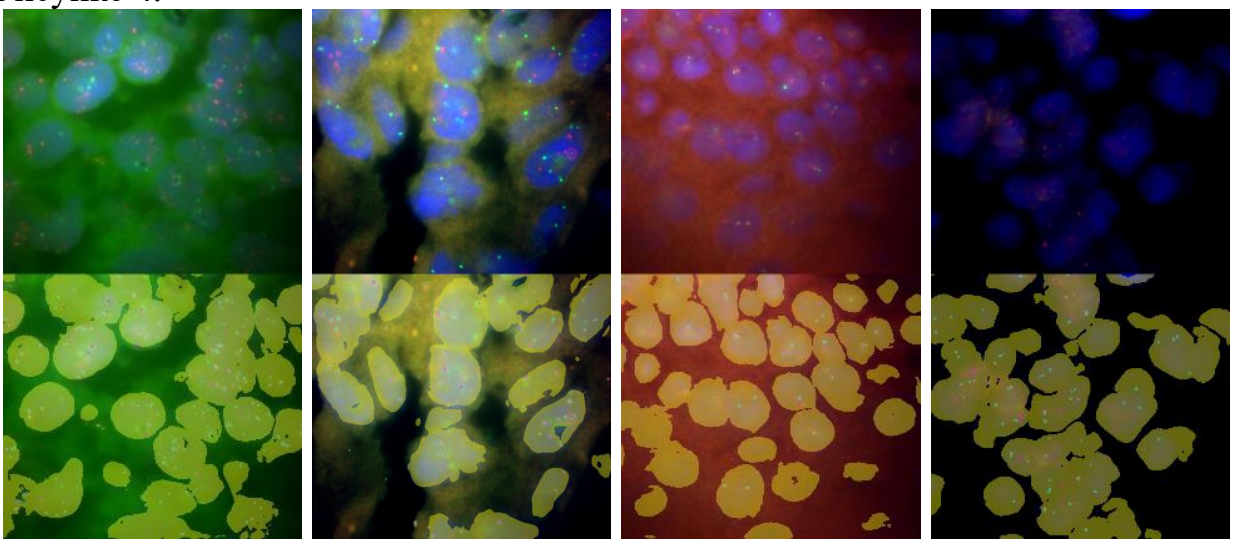


Рисунок 4. Примеры работы разработанных алгоритмов сегментации ядер, сигналов CEP17 и HER2

Четвертая глава посвящена разработке автоматизированного алгоритма поддержки принятия врачебных решений (ААППВР) с использованием полученных на предыдущих этапах бинарных масок ядер и сигналов на выборке №2. В качестве базовых ААППВР были использованы клинические рекомендации производителя реактивов Dako и международные клинические рекомендации ASCO/CAP 2023 для ручного анализа.

Для определения амплификации согласно рекомендациям Dako и ASCO/CAP необходимо определить среднее соотношение числа сигналов на ядро, а для рекомендаций ASCO/CAP – еще и среднее число сигналов HER2 на ядро. В связи с тем, что на вход алгоритму поступают бинарные маски ядер и сигналов, достоверно оценить число сигналов и его распределение по ядрам невозможно из-за наличия кластеров. В работе предложено определять среднее соотношение числа сигналов на ядро как среднее соотношение соответствующих площадей масок сигналов для каждой связанной области маски ядер. Число сигналов HER2 для использования рекомендаций ASCO/CAP по бинарным маскам определялось как число соответствующих связанных областей, что, как следует из результатов в Таблице 3, привело к отказу от выдачи заключения (неопределенная амплификация) в половине случаев. Таким образом, для дальнейших исследований за основу были выбраны клинические рекомендации производителя реагентов Dako.

Таблица 3.

Число случаев выдачи и отказа от выдачи заключения ААППВР на основе рекомендаций для ручного анализа

Рекомендации	**Число решений**	**Число отказов от решения**
Dako	44	0
ASCO/CAP 2023	22	22

Чтобы учесть в расчете среднего соотношения числа сигналов HER2/CEP17 различия площадей одиночных ядер и кластеров было предложено проводить взвешивание соотношений числа сигналов на площадь связанных областей, в пределах которых они были определены, при допущении о равном размере одиночных ядер по следующей формуле:

$$R = \frac{\sum_i R_i \cdot S_i}{\sum_i S_i}, \quad (3)$$

где R – итоговое среднее соотношение числа сигналов;

R_i – соотношение числа сигналов в i-ой связанной области на бинарной маске ядер;

S_i – площадь i-ой связанной области на бинарной маске ядер.

В Таблице 4 приведены метрики качества ААППВР, полученные на основе среднего и взвешенного среднего (по формуле 3) соотношений. Как следует из полученных результатов, взвешивание позволяет повысить ДС, ДЭ и каппу Коэна автоматизированного алгоритма.

Таблица 4.

Метрики качества ААППВР для определения амплификации гена HER2 с использованием рекомендаций Dako для ручного анализа. ИП, ЛО, ЛП, ИО – число истинно-положительных, ложно-отрицательных, ложно-положительных и истинно-отрицательных случаев, соответственно, и κ – каппа Коэна

Соотношение HER2 / CEP17	**ИП**	**ЛО**	**ЛП**	**ИО**	**ДЧ**	**ДС**	**ДЭ**	$\boldsymbol{\kappa}$
Среднее	12	9	5	18	0,57	0,78	0,68	0,36
Взвешенное среднее	12	9	3	20	0,57	0,87	0,73	0,45

Для принятия решения о наличии или отсутствии амплификации в рекомендациях Dako используется пороговое значение 2,00, которое предположительно является субоптимальным при использовании автоматизированных алгоритмов сегментации. На Рисунке 5 представлены графики зависимостей ДЭ и каппы Коэна от порогового значения соотношения числа сигналов HER2/CEP17 для порогового метода установления амплификации гена HER2. Использование порогового значения 2,18 вместо 2,00 позволило достичь на исследуемой выборке ДЧ, ДС, ДЭ и каппы Коэна 0,57, 0,91, 0,75 и 0,49, соответственно.

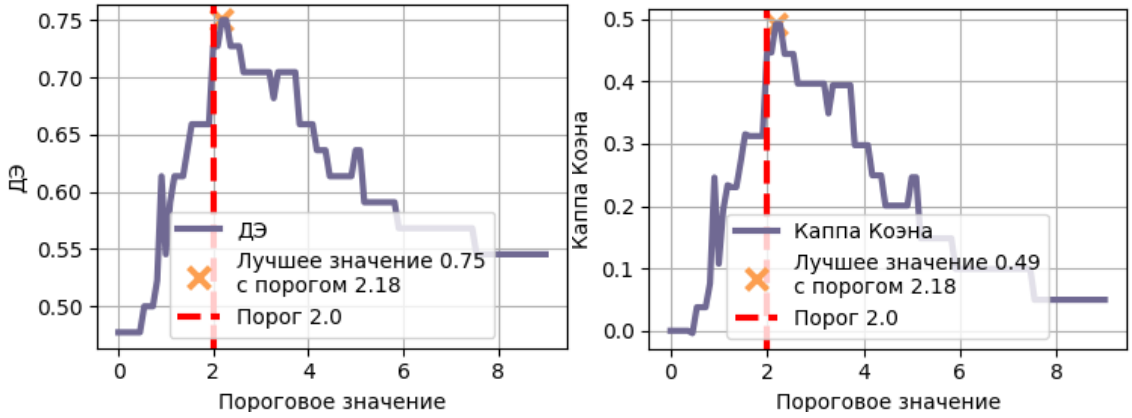


Рисунок 5. Зависимости ДЭ (слева) и каппы Коэна (справа) от порогового значения взвешенного среднего соотношения числа сигналов HER2/CEP17

Для повышения качества итогового автоматизированного заключения предложено использование стекинг-подхода, позволяющего выполнить уточнение первичного автоматизированного заключения, полученного пороговым методом. Для этого на вход мета-классификатора, в качестве которого использован алгоритм случайного леса, дополнительно к первичному автоматизированному заключению, являющемуся бинарным признаком (да/нет), подавались признаки, основанные на распределениях числа, площадей и интенсивностей выделенных объектов:

– числа связанных областей на масках ядер, сигналов CEP17 и сигналов HER2;

– квантили с 0-го до 100-го с шагом 5 площадей связанных областей ядер, сигналов CEP17 и сигналов HER2;

– квантили с 0-го до 100-го с шагом 5 интенсивностей областей ядер, сигналов CEP17 и сигналов HER2 в соответствующих каналах регистрации изображений.

В Таблице 5 приведены значения метрик качества ААППВР с мета-классификатором и без него при использовании различных комбинаций признаков. Оценка проводилась методом стратифицированной пятикратной кросс-валидации. Наилучшие значения ДЭ и каппы Коэна достигаются при использовании мета-классификатором всех указанных признаков.

Таблица 5.

Метрики качества ААППВР для определения амплификации гена HER2 с использованием мета-классификатора. 1 – с учетом признаков числа и площадей по бинарным маскам, 2 – с учетом признаков интенсивности, 3 – с учетом предварительного автоматизированного заключения

1	**2**	**3**	**ДЧ**	**ДС**	**ДЭ**	***к***
Предварительное автоматизированное заключение			0,57	**0,91**	0,75	0,49
Да	Нет	Нет	0,80	0,69	0,73	0,45
	Да	Нет	0,80	0,76	0,77	0,52
	Да	Да	**0,87**	0,76	**0,79**	**0,58**

На основе разработанных алгоритмов было создано программное обеспечение, которое было использовано в научно-исследовательской

деятельности отделения онкоцитологии МНИОИ им. П.А. Герцена – филиала ФГБУ «НМИЦ радиологии» Минздрава России при определении амплификации гена HER2 при дифференциальной диагностике РМЖ и в учебном процессе кафедры «Биомедицинские технические системы» МГТУ им. Н.Э. Баумана.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. Рак молочной железы является основным типом онкологии и одной из основных причин смертности у женского населения. Для выбора наиболее эффективной тактики лечения важно установить, в частности, HER2 статус, который, в случае неопределенной (2+) оценки результата ИГХ исследования, определяется по результатам FISH исследования. Анализ FISH-изображений проводится вручную и осложняется рядом аппаратных и биологических артефактов, что приводит к низкому уровню согласованности между патологами.
2. Современное состояние автоматизации определения амплификации гена HER2 по FISH-изображениям показало, что предлагаемые в научных публикациях алгоритмы не описаны полностью или не позволяют получить собственно оценку амплификации; обучение и верификация их работы проводились на закрытых базах данных. Вопрос устранения артефактов на FISH-изображениях, используемых для определения HER2 статуса не освещен в научных публикациях, однако показано, что в случае их наличия качество работы автоматизированных алгоритмов снижается.
3. Разработанные алгоритмы сегментации ядер и сигналов на FISH-изображениях на базе СНС с предварительным устранением фоновой флуоресценции, учитывают в явном или косвенном виде наличие аппаратных и биологических артефактов.
4. Было установлено, что при автоматизированном определении амплификации гена HER2 путем сравнения взвешенного среднего отношения площадей флуоресцирующих областей HER2/CEP17 с пороговым значением, последнее необходимо скорректировать. На исследованной выборке изменение порогового значения со значения 2,00, используемого в клинических рекомендациях для ручного анализа, до 2,18 приводит к дополнительному повышению качества автоматизированного первичного заключения до уровня ДЧ, ДС, ДЭ и каппы Коэна 0,57, 0,91, 0,75 и 0,49, соответственно.
5. Применение мета-классификатора, использующего автоматизированное предварительное заключение на базе клинических рекомендаций для ручного анализа, распределения числа, площадей и интенсивностей выделенных областей ядер и сигналов позволяет дополнительно повысить согласованность с патологом до уровня каппы Коэна 0,58.

6. Разработана БТС для автоматизированного определения амплификации гена HER2 по FISH-изображениям для дифференциальной диагностики РМЖ с учетом аппаратных и биологических артефактов. Формирование автоматизированного заключение в разработанной БТС выполнено путем последовательного решения задач коррекции фоновой флуоресценции, сегментации ядер и сигналов, классификации на базе стекинг-подхода. В результате апробации на выборке 155 изображений с различными артефактами для 44 пациентов были установлены показатели качества выполнения целевой функции на уровне ДЧ, ДС, ДЭ и каппы Коэна 0,87, 0,76, 0,79 и 0,58, соответственно.

7. Результаты работы были использованы в научно-исследовательской деятельности отделения онкоцитологии МНИОИ им. П.А. Герцена – филиала ФГБУ «НМИЦ радиологии» Минздрава России и в учебном процессе кафедры «Биомедицинские технические системы» МГТУ им. Н.Э. Баумана.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Махов Д.С., Самородов А.В. Автоматизированная поддержка принятия врачебных решений при определении амплификации гена HER2 по FISH изображениям при раке молочной железы // Биомедицинская радиоэлектроника. 2026. Т. 29. № 4. С. 18−22.
2. Разработка биотехнической системы автоматизированного определения HER2 статуса при раке молочной железы методом флуоресцентной *in situ* гибридизации (FISH) / Махов Д.С. [и др.] // Биомедицинская радиоэлектроника. 2022. Т. 25, № 5. С. 58-69.
3. Коррекция фона для автоматизации анализа FISH изображений / Махов Д.С. [и др.] // Биотехносфера. 2021. № 1(66). С. 23-29.
4. Convolutional neural network application for cells segmentation in immunocytochemical study / Parpulov D. [et al.] // 2018 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT). IEEE, 2018. С. 87-90.
5. Makhov D., Samorodov A., Slavnova E. Automatization of HER2 status assessment in breast cancer // 2019 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT). IEEE, 2019. С. 171-173.
6. Makhov D., Samorodov A., Slavnova E. Different approaches for automatic nucleus image segmentation in fluorescent *in situ* hybridization (FISH) analysis for HER2 status assesment // Conference of Open Innovations Association, FRUCT. 2020. No. 26. P. 270-277.
7. Аугментация на базе гамма преобразования для автоматизированного анализа флуоресцентных изображений / Махов Д.С. [и др.] // Физика и радиоэлектроника в медицине и экологии ФРЭМЭ'2022, Владимир-Суздаль. 2022. С. 149-152.