

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УДК 004.932.2: 004.2: 004.67

**ЛИСИЦА
ЕВГЕНИЯ ВЛАДИМИРОВНА**

**АНАЛИЗ ИЗОБРАЖЕНИЙ ПОПУЛЯЦИЙ РАКОВЫХ КЛЕТОК
В СИСТЕМАХ ОБРАБОТКИ ДАННЫХ ЛЮМИНЕСЦЕНТНОЙ
МИКРОСКОПИИ**

Автореферат диссертации
на соискание ученой степени кандидата технических наук
по специальности 05.13.01 – системный анализ, управление
и обработка информации

Минск, 2021

Научная работа выполнена в Белорусском государственном университете

Научный руководитель –

Апанасович Владимир Владимирович,
доктор физико-математических наук,
профессор, ведущий научный сотрудник
кафедры логистики ГУО «Института
бизнеса Белорусского государственного
университета».

Официальные оппоненты:

Краснопрошин Виктор Владимирович,
доктор технических наук, профессор,
заведующий кафедрой информационных
систем управления Белорусского
государственного университета;

Муха Владимир Степанович,
доктор технических наук, профессор,
профессор кафедры информационных
технологий автоматизированных систем
Учреждения образования «Белорусский
государственный университет
информатики и радиоэлектроники».

Оппонирующая организация:

**ГНУ «Объединенный институт проблем
информатики НАН Беларуси».**

Защита состоится «26» февраля 2021 года в 14⁰⁰ часов на заседании
совета по защите диссертаций Д 02.01.14 при Белорусском государственном
университете по адресу: 220030, г. Минск, ул. Ленинградская, 8 (корпус
юридического факультета), ауд. 407. Телефон ученого секретаря 209-57-09.

С диссертацией можно ознакомиться в библиотеке Белорусского
государственного университета.

Автореферат разослан « » января 2021 года.

Ученый секретарь совета по защите диссертаций
кандидат физико-математических наук доцент



Ю.И. Воротницкий

ВВЕДЕНИЕ

Методы цифровой обработки изображений находят все более широкое применение при проведении прикладных исследований (аэрокосмические системы, системы видеонаблюдения, биомедицина). Рассматриваемый в диссертационной работе новый класс решаемых задач связан с применением методов иммуногистохимии для анализа микроскопических объектов в медицине при проведении диагностики раковых заболеваний. В большинстве методов иммуногистохимии для исследования изображений срезов тканей и клеток используются системы люминесцентной микроскопии. Методы иммуногистохимии являются более точными и менее затратными по сравнению с рентгеноскопическими исследованиями, поэтому широко распространены в клинической практике. Они используют понятие биомаркера – белка, связанного с изучаемым биологическим процессом, – и позволяют оценить и спрогнозировать уровень экспрессии маркеров рака в зависимости от действия агентов анти-рака. В частности, в качестве биомаркера рака молочной железы обычно используется белок эстроген-рецептор. В ходе подготовки проб в структуру клетки, цитоплазму или ядро внедряется биомаркер, окрашенный красителем, люминесцирующим на определенной длине волны. В системах люминесцентной микроскопии происходит регистрация интенсивности биомаркера.

Общим недостатком существующих методов и программных средств обработки иммуногистохимических изображений является отсутствие адаптивной настройки параметров сегментации объектов на изображениях, что приводит к затруднениям при обработке больших объемов данных, обусловленным временными затратами и субъективностью при подборе параметров для сегментации.

Существующее программное обеспечение позволяет выполнять анализ изображений раковых клеток без учета локальной и внутриколониальной вариации экспрессии биомаркеров (что связано с неопределенностью и неоднородностью популяций раковых клеток). Интенсивность экспрессии биомаркера рассчитывается интегрально по всему изображению среза ткани. При этом не учитываются локальные особенности скоплений раковых клеток, а главное – не используется дифференциальная информация о них (размеры и форма каждой из клеток, интенсивность биомаркера в них). Это приводит к значительной потере информации в локальных областях, а также не позволяет достоверно дифференцировать степень распространения рака. В то же время, рассматривая цитоплазму, ядра, кластеры клеток как независимые объекты на изображении, на основе объектно-ориентированного подхода можно учесть их индивидуальные характеристики на пиксельном уровне.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Связь работы с научными программами (проектами), темами

Тема диссертационной работы утверждена на заседании совета факультета радиофизики и компьютерных технологий Белорусского государственного университета 29.11.2011 г. (протокол № 4).

Тема диссертации соответствует приоритетным направлениям фундаментальных и прикладных научных исследований Республики Беларусь на 2011–2015 гг., утвержденных Постановлением Совета Министров Республики Беларусь от 19.04.2010 № 585 (п. 5.1 «Методы математического и компьютерного моделирования, компьютерные технологии и интеллектуальные системы поддержки принятия решений»).

Исследование посвящено разработке моделей, алгоритмов и программных средств анализа люминесцентных изображений, что соответствует техническим наукам и специальности 05.13.01 – системный анализ, управление и обработка информации.

Диссертационная работа выполнялась в рамках следующих НИР:

– «Разработка методов цифровой обработки люминесцентных изображений биологических объектов» № Ф11МС-006 (номер госрегистрации 20114329, срок выполнения 2011–2013 гг.) в рамках государственной комплексной программы научных исследований «Фотоника»;

– «Разработка моделей, методов и программных средств анализа многоканальных люминесцентных изображений раковых клеток» (номер госрегистрации 20161301, срок выполнения 2016–2018 гг.) в рамках государственной программы научных исследований «Информатика, космос, безопасность»;

– «Разработка алгоритмов и программных средств анализа иммуногистохимических изображений раковых клеток» (номер госрегистрации 20190871, срок выполнения 2019–2020 гг.) в рамках государственной программы научных исследований «Информатика, космос, безопасность».

Цель и задачи исследования

Целью данной работы является разработка моделей, алгоритмов и программных средств для анализа люминесцентных изображений раковых клеток, полученных в системах люминесцентной микроскопии при диагностике онкологических заболеваний.

Данная цель предполагает решение следующих задач:

1. Разработка адаптивного алгоритма для автоматической сегментации и локализации ядер раковых клеток на изображениях, учитывающего индивидуальные характеристики объектов на пиксельном уровне изображений.

2. Разработка имитационной модели 3-канальных люминесцентных изображений, содержащих две популяции клеток, согласно протоколу эксперимента по регистрации изображений рака молочной железы при исследовании рецептора-эстрогена.

3. Разработка алгоритмов отбора наиболее информативных признаков интенсивности пикселей, расположения и формы ядер, а также методов анализа распределений этих характеристик для классификации и кластеризации объектов, сегментированных на люминесцентных изображениях раковых клеток.

4. Создание программного пакета для анализа экспериментальных и смоделированных данных люминесцентных изображений раковых клеток.

Объектом исследования являются системы регистрации и анализа люминесцентных изображений популяций клеток. Предмет исследования: модели и алгоритмы анализа люминесцентных изображений популяций клеток.

Научная новизна

1. Разработана имитационная модель люминесцентных изображений раковых клеток, которая позволяет воспроизводить изображения, содержащие несколько популяций клеток с учетом различной формы и расположения клеток в пространстве, распределений интенсивности их люминесценции, оптических искажений, возникающих в системе регистрации.

2. Разработан адаптивный алгоритм сегментации люминесцентных изображений раковых клеток, полученных согласно протоколу с использованием красителей DAPI, Cy3 и Cy5 и применением белка рецептора эстрогена. Разработанный алгоритм полностью автоматизирован, при этом его точность сопоставима с точностью полуавтоматических методов.

3. Разработан алгоритм отбора информативных признаков для описания сегментированных объектов на люминесцентных изображениях раковых клеток и с его помощью проведено ранжирование признаков по степени информативности, а также исследование методов классификации и кластеризации для анализа данного вида изображений.

4. Разработано программное обеспечение для сегментации и анализа люминесцентных изображений раковых клеток, которое позволяет выполнять анализ данных в интерактивном и автоматизированном режимах, используя методы классификации, кластеризации и снижения размерности данных.

Положения, выносимые на защиту

1. Имитационная модель изображений популяций раковых клеток, основанная на формализованном описании пространственного расположения клеток и распределения интенсивности регистрируемого излучения красителей с учетом искажений, вносимых системой регистрации оптического излучения, отличающаяся от известных тем, что позволяет воспроизводить люминесцентные изображения популяций клеток, регистрируемых при исследовании белка рецептор-эстроген для диагностики рака молочной железы.

2. Алгоритм сегментации ядер раковых клеток на люминесцентных изображениях, использующий метод главных компонент для учета корреляции между цветовыми каналами регистрации изображения и адаптивную пороговую сегментацию с учетом предварительной оценки геометрических характеристик исследуемых объектов, позволяющий по сравнению с известными алгоритмами снизить ошибку локализации ядер в два раза.

3. Алгоритм отбора и ранжирования наиболее информативных признаков ядер раковых клеток, основанный на использовании случайного леса с CART-деревьями для ранжирования с последующим объединением в группы по информативности, позволяющий сократить количество геометрических признаков ядер с 59 до 11 и признаков распределения интенсивности пикселей с 42 до 2 при сохранении точности классификации, уменьшая временные затраты при выполнении анализа в два раза и более.

4. Программный пакет CellDataMiner для анализа изображений популяций раковых клеток, основанный на использовании объектно-ориентированной методологии анализа люминесцентных изображений на пиксельном уровне, который реализует предложенный алгоритм сегментации объектов и позволяет в автоматическом режиме проводить классификацию, кластеризацию и визуализацию результатов анализа изображений раковых клеток. Пакет позволяет выделять группы раковых и здоровых клеток с точностью классификации превосходящей 95%.

Личный вклад соискателя

Соискателем выполнены все изложенные в диссертации разработки и исследования. Постановка задач и обсуждение результатов проводились совместно с научным руководителем и сотрудниками кафедры системного анализа и компьютерного моделирования Белорусского государственного университета. Соавторы опубликованных работ принимали участие в обсуждении промежуточных и конечных результатов. Обработка, интерпретация данных, а также выводы сделаны автором самостоятельно.

Апробация диссертации и информация об использовании её результатов

Основные положения диссертационной работы прошли апробацию на следующих научных конференциях: 68-ая научная конференция студентов и аспирантов БГУ (Минск, 2011); Pattern recognition and information processing (PRIP'2011) (Minsk, 2011); Международный конгресс по информатике: информационные системы и технологии (Минск, 2011); II Международная научно-практическая конференция молодых ученых «Научные стремления – 2011» (Минск, 2011); Modeling and simulation MS'2012 (Minsk, 2012); III Международная молодежная научно-практическая конференция «Научные стремления – 2012» (Минск, 2012); Информационные технологии и системы 2012 (ИТС 2012) (Минск, 2012); Медэлектроника – 2012. Средства медицинской электроники и новые медицинские технологии (Минск, 2012); Информационные технологии и системы 2013 (ИТС 2013) (Минск, 2013); Международный конгресс по информатике: информационные системы и технологии (Минск, 2013); Pattern Recognition and Information Processing (PRIP'2014) (Minsk, 2014); Информационные технологии и системы 2014 (ИТС 2014) (Minsk, 2014); Медэлектроника – 2014. Средства медицинской электроники и новые медицинские технологии (Минск, 2014); Информационные технологии и системы 2015 (ИТС 2015) (Минск, 2015); Генетика и биотехнология XXI века: проблемы, достижения, перспективы (к 50-летию ГНУ «Институт генетики и цитологии НАН Беларуси»). II Международная научная конференция (Минск, 2015); IX Съезд онкологов и радиологов стран СНГ и Евразии (Минск, 2016); International Congress on Computer Science: Information Systems and Technologies (Минск, 2016); Информационные технологии и системы 2017 (ИТС 2017) (Минск, 2017); Прикладные проблемы оптики, информатики, радиофизики и физики конденсированного состояния (Минск, 2019).

Промежуточные результаты диссертационного исследования были представлены для участия в конкурсах и отмечены следующими наградами: первое место на V Республиканском конкурсе инновационных проектов за проект «Программный пакет CellDataMiner для анализа изображений раковых клеток» (Минск, 2014); диплом на Санкт-Петербургской технической ярмарке (Санкт-Петербург, 2014); диплом на Молодежном инновационном форуме «Наука и бизнес 2015» (Минск, 2015).

Имеются акт о практическом использовании результатов исследования в учебном процессе для проведения комплекса лабораторных работ кафедрой системного анализа и компьютерного моделирования БГУ, справка о внедрении в учебный процесс для проведения лабораторных занятий кафедрой промышленной экологии Учреждения образования «Белорусский государственный техно-

логический университет», акт о практическом использовании результатов исследования в практической деятельности государственного учреждения «Республиканский научно-практический центр онкологии и медицинской радиологии им. Н.Н. Александрова».

Опубликованность результатов диссертации

Основные результаты диссертации опубликованы в 30 научных работах, в том числе в 5 статьях в научных изданиях в соответствии с п. 18 Положения о присуждении ученых степеней и присвоении ученых званий в Республике Беларусь (общий объем – 4,6 авт. л.), 2 статьях в других научных рецензируемых изданиях; 20 статьях в сборниках материалов научных конференций, 3 тезисах.

Структура и объем диссертации

Диссертационная работа состоит из перечня условных обозначений и сокращений, введения, общей характеристики работы, 4 глав, заключения, библиографического списка, 2 приложений. Полный объем диссертации составляет 156 страниц, включая 60 рисунков на 23 страницах, 23 таблицы на 6 страницах, 2 приложения на 19 страницах. Библиографический список содержит 183 наименования, включая список публикаций соискателя из 30 наименований, на 17 страницах.

ОСНОВНАЯ ЧАСТЬ

В первой главе приведено описание протокола получения люминесцентных изображений раковых клеток на основе использования красителей DAPI, Cy3, Cy5, рецептора эстрогена и системы регистрации люминесцентной микроскопии. Описана известная из литературных источников спектральная модель люминесцентных изображений раковых клеток. Данная модель, хотя и рассматривает влияние отдельных составляющих системы регистрации, не учитывает распределение регистрируемой интенсивности пикселей на изображении.

Выполнен обзор традиционных подходов к сегментации люминесцентных изображений раковых клеток. Общим недостатком приведенных методов обработки является отсутствие автоматической адаптивной настройки параметров сегментации, что приводит к большим временным потерям и затруднениям при обработке больших объемов данных.

Рассмотрены методы обработки данных, получаемых при сегментации изображений. Выполнен обзор существующих признаков объектов на изображениях и методов снижения размерности данных. Рассмотрены наиболее часто применяемые методы машинного обучения, выполнен обзор литературы по их применению при исследовании раковых заболеваний. Установлено, что на те-

кущий момент не существует единого информативного набора признаков для описания ядер клеток, сегментированных на люминесцентных изображениях раковых клеток.

Выполнен обзор программных средств анализа биомедицинских данных. Программное обеспечение можно разделить на следующие классы, в зависимости от решаемых задач: визуализация изображений, обработка изображений, классификация и поиск по содержимому. Однако не существует программного обеспечения, которое позволяло бы проводить адаптивный анализ люминесцентных изображений клеток на пиксельном уровне в автоматическом режиме.

Во второй главе предложен обобщенный подход к формализованно-логическому описанию экспериментов по регистрации трехканальных люминесцентных изображений раковых клеток.

Разработана имитационная модель люминесцентных изображений систем раковых клеток, позволяющая учитывать различные формы объектов раковых клеток и их расположение, экспериментальные эффекты, наблюдаемые на этапе регистрации люминесцентных изображений. Общая схема моделирования показана на рисунке 1. В основе моделирования как раковых, так и здоровых клеток лежит воспроизведение формы и местоположения клетки, а также воспроизведение распределения интенсивности пикселей, составляющих ядра и цитоплазму клеток.

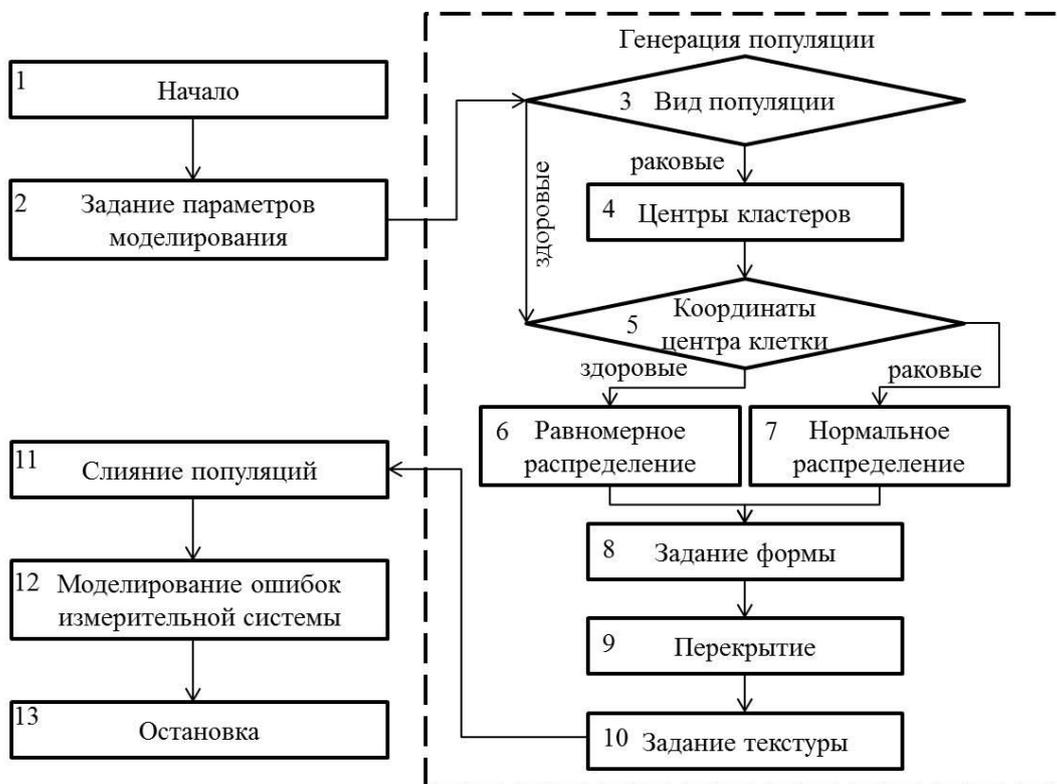


Рисунок 1. – Общая схема моделирования изображений

Блок 2. Задание параметров моделирования. Общие параметры для моделирования изображений популяции раковых клеток – размер изображения и количество клеток на нем N_{nuclei} . Популяции раковых и здоровых клеток отличаются между собой по нескольким параметрам. Между раковыми клетками существует перекрытие, у них регистрируется цитоплазма, они группируются в кластеры. Для раковых клеток необходимо задать параметры: размер ядра r_{cancer} , размер цитоплазмы $r_{cytoplasm}$, степень перекрытия L_{cancer} , количество кластеров N_{cancer} .

Для здоровых клеток отсутствует группирование в кластеры, нет люминесценции в области цитоплазмы, перекрытие ($L_{nuclei}=0$) отсутствует. Существенным параметром для моделирования этой группы клеток будет размер ядер здоровых клеток r_{nuclei} . Так же при моделировании используется параметр ξ , который задается для ядер здоровых клеток (ξ_{nuclei}), ядер (ξ_{cancer}) и цитоплазмы ($\xi_{cytoplasm}$) раковых клеток.

Генерация популяции. При моделировании раковых клеток на первом этапе (блок 4, рисунок 1) производится генерация центров кластеров. Затем происходит определение координат самих ядер (блок 7, рисунок 1), сгруппированных вокруг кластеров. При моделировании популяции здоровых (блок 6, рисунок 1) клеток их местоположение подчиняется равномерному закону распределения.

Блок 8. Задание формы. Для моделирования формы ядер и цитоплазм используется упрощенная формула:

$$x_{ij}^c = \xi \cdot r_i^c \cdot \cos(\theta_j) \text{ и } y_{ij}^c = \xi \cdot r_i^c \cdot \sin(\theta_j), \quad (1)$$

где j – номер пикселя на границе $j=1,2,\dots,K$; K – количество пикселей на границе цитоплазмы; θ_j – значение полярного угла для j -ого пикселя; i – номер клетки; ξ – реализация равномерно распределенной случайной величины на интервале $[0, \beta]$. Если степень перекрытия превышает заданное значение L , то происходит повторная генерация координат центра раковой клетки (блок 9 рисунка 1).

Блок 10. Задание текстуры. Законы распределения интенсивности пикселей в первом приближении можно рассматривать как нормальное распределение с заданными значениями математического ожидания и среднеквадратического отклонения.

Блок 11. Слияние популяций. В результате моделирования воспроизводятся две системы клеток. В процессе слияния полученных изображений формируется одно изображение, где на изображении здоровых клеток значения пикселей налагаются на соответствующие значения изображения раковых кле-

ток в областях построения опухоли. Вне опухоли значения пикселей на изображении остаются неизменными.

Блок 12. Моделирование ошибок измерительной системы. Наличие различных оптических узлов в измерительной системе приводит к возникновению шума от каждого из них на изображении. Данный шум моделируется по нормальному закону распределения $N(0, \sigma_d^2)$ с нулевым математическим ожиданием и дисперсией σ_d^2 . В качестве характеристики этого шума можно использовать понятие отношения сигнал/шум SNR .

Проведённая проверка адекватности разработанной модели показала, что модель позволяет воспроизводить изображения со схожими с экспериментально наблюдаемыми размерами и расположением ядер и распределением интенсивности пикселей в цветовых каналах.

Показана нелинейная зависимость между мощностью, поглощённой в области цитоплазмы под ядром раковых клеток, и мощностью, поглощённой в области самого ядра:

$$\tilde{P}_N^G = \tilde{P}_C^G \left(1 - \frac{\tilde{P}_N^B}{\bar{P}_N^B}\right), \quad (2)$$

где \tilde{P}_N^G – мощность, поглощённая в области цитоплазмы под ядром в зеленом канале; \tilde{P}_N^B – мощность, поглощённая в области самого ядра в синем канале; \bar{P}_N^B – мощность люминесценции, поглощенная в области ядер здоровых клеток в синем канале. Для исследования данной зависимости рассмотрено распределение поведения интенсивности пикселей в области ядер для зеленого и синего каналов изображения. Из исследованных различных наборов признаков для описания одного пикселя был отобран следующий, дающий наилучший результат при выполнении сегментации ядер:

$$X_1=B, X_2=G, X_3=B/G, X_4=G/B, X_5=1/B, X_6=1/G, \quad (3)$$

где B, G – интенсивности пикселей в синем и зеленом каналах люминесцентного изображения. Линейная взаимосвязь между признаками наблюдается для групп (X_3, X_2) , (X_1, X_4) , (X_3, X_5) , (X_4, X_6) . Выполненный анализ показал наличие статистической зависимости между признаками.

В третьей главе предложен адаптивный алгоритм сегментации ядер с использованием метода главных компонент (МГК) и проведен отбор наиболее информативных признаков выделенных объектов.

Применение МГК учитывает корреляцию между каналами регистрации изображения, что позволяет более точно восстановить распределение красителей в пространстве за счет учета информации, имеющейся в других каналах

изображения вследствие перекрытия спектров люминесценции. Алгоритм сегментации ядер состоит из двух этапов: вначале выполняется сегментация маски опухоли, а затем производится сегментация границ ядер раковых клеток.

Алгоритм сегментации маски опухоли состоит из пяти этапов, показанных на рисунке 2. На первом этапе для восстановления информации используется МГК (рисунок 2, блок 1), где в качестве признаков пикселей предложено использовать комбинацию между цветовыми компонентами исходного изображения, определяемую формулой 3. На втором этапе для отделения пикселей цитоплазмы от пикселей фона используется пороговый метод, где пороговое значение рассчитывается по методу Отцу для изображения, построенного по первой главной компоненте Z_1 .



Рисунок 2. – Структурная схема алгоритма сегментации опухоли

После бинаризации Z_1 необходимо дополнительно выполнить процедуру инвертирования. Затем (рисунок 2, блок 3) по результатам сегментации проводится анализ гистограмм эквивалентных диаметров пустот D_{eq} , на основе чего происходит численная оценка размеров ядер (рисунок 2, блок 4) по методу наименьших квадратов с оптимизацией по методу Левенберга-Маркварта.

Исходя из анализа гистограммы эквивалентных диаметров D_{eq} , функцию, аппроксимирующую распределение эквивалентных диаметров D_{eq} , предложено представить в виде суммы нормального и экспоненциального распределений:

$$F(D_{eq}; \lambda, m, \sigma) = p_1 \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(D_{eq}-m)^2}{2\sigma^2}} + p_2 \lambda e^{-\lambda D_{eq}}, \quad (4)$$

где p_1 и p_2 – вклады нормального и экспоненциального распределений соответственно; m и σ – математическое ожидание и среднеквадратическое отклонение для нормального распределения; $1/\lambda$ – математическое ожидание экспоненциального распределения.

В качестве максимального размера ядер предложено использовать двух-сигмовую границу, соответственно $D_{max} = m + 2\sigma$. Минимальный размер D_{min} ядер оценивается по формуле (блок 4 рисунка 2):

$$D_{min} = \arg(\min_{0 \leq D_{eq} \leq m} F(D_{eq}; \lambda, m, \sigma)). \quad (5)$$

Для улучшения качества найденных границ используется морфологическая обработка (блок 5 рисунка 2). Заполняются пустоты, обусловленные наличием шума от различных оптических узлов системы, когда эквивалентный диаметр не превосходит D_{min} . Удаляются пустоты, размеры которых больше D_{max} , и объекты, у которых эквивалентный диаметр D_{eq} объектов меньше D_{min} . Оператор морфологического замыкания диском с диаметром D_{min} используется для сглаживания границ найденной маски.

На рисунке 3 показаны итоговое отображение границ на исходном цветном изображении и маска опухоли отдельно.

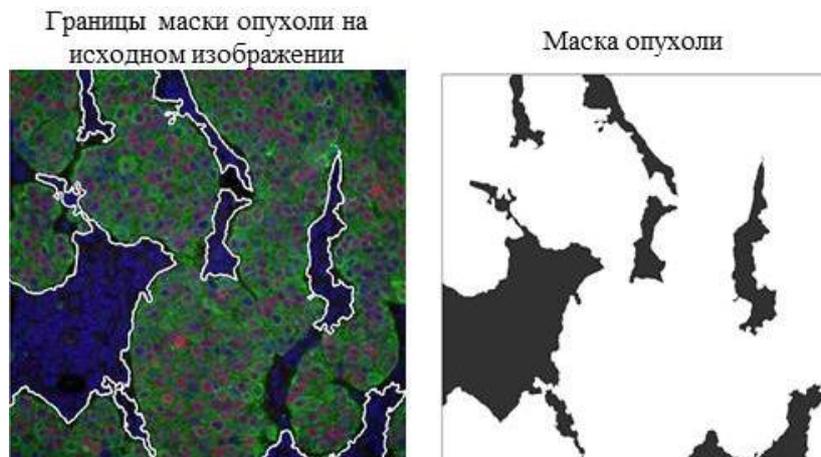


Рисунок 3. – Границы маски опухоли

Обобщенная схема сегментации ядер показана на рисунке 4. Для уменьшения воздействия шума применяется процедура пространственной фильтрации медианным фильтром с диаметром m . В работе предложено использовать адаптивную пороговую сегментацию по методу Отцу (блок 2 рисунка 4) с K различными вариантами деления изображения на блоки размером D_{max} . В этом случае каждый пиксель изображения классифицируется как фоновый или принадлежащий объекту (ядру) K раз (>9). Если пиксель классифицирован более $0,5 * K$ раз как пиксель ядра, то на итоговой маске он будет относиться к пикселям ядер BWN , иначе к фону. На конечном изображении BWN остаются только те объекты, которые также принадлежат маске опухоли BW .

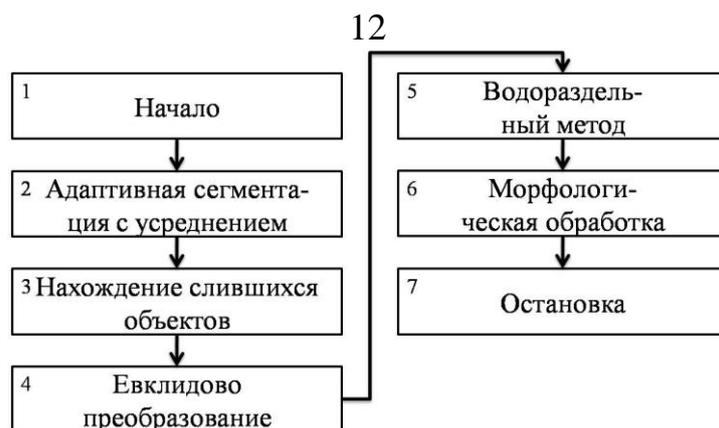


Рисунок 4. – Общая схема алгоритма сегментации ядер

Дальнейшая работа алгоритма связана с разделением слившихся ядер (блоки 3-5 рисунка 4). Для разделения слившихся ядер в работе используется водораздельный метод. С целью уменьшения чрезмерной сегментации используется фильтрация. Для усовершенствования работы метода на первом этапе строится маска $BWNC$, которая содержит ядра, площадь которых превосходит $Area_{max} = 2 * \pi * (0.5 * D_{max})^2$. Каждый пиксель изображения $DIST$ является элементарным дополнением к вычисленным расстояниям между текущим пикселем и ближайшим ненулевым пикселем инвертированного изображения $BWNC$. С целью устранения локальных неоднородностей, вызванных ошибкой пороговой сегментации, используется медианный фильтр диаметром D_{min} . В качестве границ раздела между слившимися ядрами на маске $BWNC$ используются линии водоразделов, найденные на изображении $DIST$ алгоритмом водораздела. В заключение выполняется морфологическая обработка (блок 6 рисунка 4). На рисунке 5 показан результат сегментации ядер раковых клеток на люминесцентных изображениях.

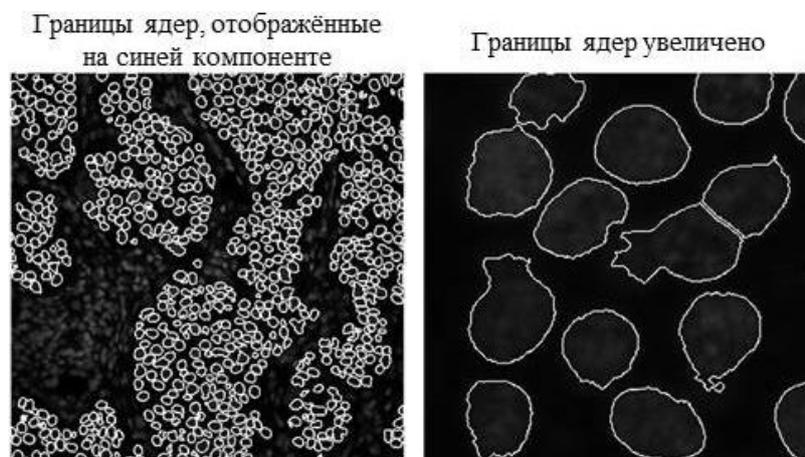


Рисунок 5. – Результат сегментации ядер

Разработанный адаптивный алгоритм на основе МГК ($АСМГК$) был сравнен с двумя алгоритмами сегментации CP (использует классический подход

для сегментации ядер) и *IJ* (алгоритм сегментации ядер, окрашенных с помощью гематоксилина и эозина), показавшими на этапе их сравнительного анализа лучшие результаты на экспериментальных и синтетических изображениях. В ходе вычислительных экспериментов проводилось исследование 9 экспериментальных люминесцентных изображений раковых клеток, полученных в ходе эксперимента на микрочипах срезов тканей опухолей молочной железы. В качестве эталонных масок изображений рассмотрены контуры ядер, выделенные экспертным путем. Количество раковых клеток на изображениях варьируется от 70% и более от общего количества клеток.

Количество выделенных экспертами объектов-ядер на 9 исследуемых изображениях следующее: 828, 708, 1154, 902, 675, 553, 1014, 704, 678. Результаты сравнительного анализа для 3 алгоритмов и пример фрагмента сегментированных границ ядер раковых клеток представлены на рисунке 6.

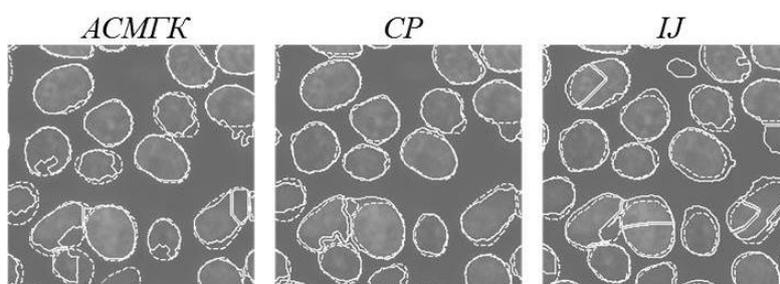


Рисунок 6. – Примеры сегментации экспериментальных изображений

Для всех рассматриваемых алгоритмов сегментации ошибка слившихся ядер ε_c значительно меньше других ошибок и не превышала 0,01. Ошибка пропущенных ядер ε_m для *АСМГК* составила $0,07 \pm 0,02$, что в 2-3 раза меньше чем для *СР* ($0,24 \pm 0,06$). Наименьшая ошибка пропущенных ядер ε_m была получена для алгоритма *IJ* ($0,02 \pm 0,01$). Количество ложно найденных ядер ε_f у разработанного алгоритма *АСМГК* составляет $0,10 \pm 0,04$, что немногим меньше чем для алгоритма *СР* ($0,15 \pm 0,03$) и меньше чем для алгоритма *IJ* ($0,6 \pm 0,1$).

Разработанный алгоритм *АСМГК* показал наименьшую ошибку ε_o для чрезмерно сегментированных ядер $0,1 \pm 0,03$, что меньше чем в алгоритме *IJ* ($0,20 \pm 0,01$) и *СР* ($0,32 \pm 0,06$). Интегральная ошибка локализации ε для разработанного алгоритма *АСМГК* составила $0,28 \pm 0,07$, что значительно меньше значений ошибок $0,71 \pm 0,12$ и $0,83 \pm 0,11$ для алгоритмов *СР* и *IJ*. Разработанный алгоритм *АСМГК* показал точность сегментации δ , сопоставимую с *IJ*, равную $0,74 \pm 0,03$, а наихудшая точность сегментации $0,69 \pm 0,04$ получена для *СР*.

Сравнительный анализ трех алгоритмов сегментации демонстрирует, что разработанный адаптивный алгоритм *АСМГК* обладает улучшенным качеством

локализации объектов, при этом его точность сегментации не меньше чем у алгоритмов *CP* и *IJ*.

Для исследования алгоритма сегментации ядер на основе разработанной модели были сгенерированы изображения со значениями среднеквадратического отклонения шума в интервале от 0 до 0,04. Размер смоделированного изображения 2048 пикселей на 2048, количество клеток 2200, из которых 1200 содержат цитоплазму (больные клетки), 1000 – нет (здоровые клетки). Между здоровыми клетками отсутствует перекрытие, перекрытие раковых клеток составляет $L = 0,1$. Радиус здоровых клеток $r^n = 20$, $\xi^n = 0$. Параметры размера для раковых клеток $r^n = 25$, $\xi^n = 0,1$ для ядер и $r^n = 40$, $\xi^n = 0,1$. Закон распределения интенсивности в ядрах здоровых клеток, цитоплазме, ядрах раковых клеток $N(0,5, 0,17)$.

В ходе вычислительных экспериментов на смоделированных изображениях установлено, что алгоритм *CP* успешно работает в диапазоне значений параметра $\sigma_d = 0-0,03$. Ошибка точности локализации объектов ε изменяется в пределах 0,3-0,6. Если шум от оптических узлов значительно возрастает, $\sigma_d > 0,03$, то метод не работает ($\varepsilon = \infty$). Точность сегментации уменьшается по мере увеличения шума, а при значении параметра шума $\sigma_d = 0,02$ происходит резкое ухудшение точности сегментации. Неработоспособность алгоритма в условиях высокого шума обусловлена неустойчивостью метода адаптивной пороговой сегментации. Рабочий диапазон разработанного алгоритма *АСМГК* по параметру σ_d находится в пределах от 0,002 до 0,25. Интегральная точность локализации объектов ε алгоритмом *АСМГК* варьируется в пределах 0,1-0,2 и почти в 2 раза меньше результатов *CP*. Ошибки определения пропущенных ε_m , чрезмерно сегментированных ε_o и слившихся ядер ε_c возрастают при увеличении шума. Для ошибки ложно-сегментированных объектов ε_f характерно поэтапное изменение: уменьшение ошибки при увеличении шума до 0,004, затем ошибка ε_f остается неизменной до уровня шума $\sigma_d = 0,125$, начиная с которого ошибка ε_f резко возрастает. Средняя точность сегментации $\langle \delta \rangle$ метода *АСМГК* не зависит от уровня шума в диапазоне исследуемых значений параметра σ_d и в среднем на 20% выше чем *CP*.

Разработан алгоритм отбора информативных признаков. На первом этапе оценивается информативность каждого из признаков методами на основе корреляции, случайного леса, логистической регрессии. Затем признаки отсортировываются по убыванию информативности, после чего оценивается оптимальное количество признаков для описания объектов. В качестве критерия оценки использована ошибка классификации, которая будет достигать минимального значения при оптимальном наборе признаков. В качестве метода классификации используется метод случайного леса с CART деревьями. В результате вы-

полненного анализа было показано, что 6 найденных наиболее информативных признаков интенсивности пикселей (из рассмотренных 42) являются полностью достаточными для описания всех объектов. При этом точность классификации сохраняется и составляет 0,025. При использовании всего 2 наиболее информативных признаков (верхний квантиль интенсивности в красном канале и математическое ожидание в зеленом канале) ошибка классификации составляет 0,03, что достаточно для уверенной классификации объектов на изображениях. Из 59 геометрических признаков формы были отобраны 11 наиболее информативных признаков: второе собственное значение матрицы тензора инерции, координата центра масс по оси абсцисс, координата центра масс по оси ординат, первое собственное значение матрицы тензора инерции, первый Ну-момент, нормированный центральный момент $\bar{\mu}_{21}$, коэффициент плотности, третий Ну-момент, пятый Ну-момент, нормированный центральный момент $\bar{\mu}_{20}$, площадь выпуклой оболочки для описания ядер на люминесцентных изображениях раковых клеток с сохранением точности классификации объектов. Ошибка классификации составляет 0,110. Временные затраты в результате уменьшения количества признаков для описания объектов были сокращены в 2,3 раза.

В четвертой главе изложено описание разработанного программного пакета CellDataMiner, интегрирующего объектно-ориентированную методологию анализа популяций клеток на люминесцентных изображениях на пиксельном уровне с последующим использованием реализованных методов кластеризации и классификации данных для выделения групп раковых клеток.

На рисунке 7 представлена последовательность обработки данных в разработанном программном обеспечении (ПО). Первая часть ПО отвечает за формирование набора данных (блоки 2-4 рисунка 7). Данные для анализа могут быть смоделированы (блок 3 рисунка 7) или загружены пользователем (блок 4 рисунка 7). Исходные данные, регистрируемые в эксперименте, сохранены в формате TIFF. У пользователя есть возможность загружать как единичное изображение, так и набор изображений.

Вторая часть ПО посвящена сегментации изображения (блоки 5-7 рисунка 7). Имеется возможность как автоматической адаптивной оценки параметров сегментации, так и ручного задания параметров метода сегментации. Имеется возможность использовать смоделированные изображения для настройки параметров сегментации. После настройки параметров происходит сегментация (блок 7 рисунка 7) маски опухоли и построение маски ядер, каждое из которых строится в отдельном окне.

В третьей части ПО (блоки 8-12 рисунка 7) реализованы возможности машинного обучения. Для решения задачи кластеризации (блок 9 рисунка 7) был реализован метод k-средних, который прост в реализации и показывает

устойчивые результаты сегментации. Среди методов классификации (блок 10 рисунка 7) были реализованы метод случайного леса (CART-деревья) и методы линейного и квадратичного дискриминантного анализа.

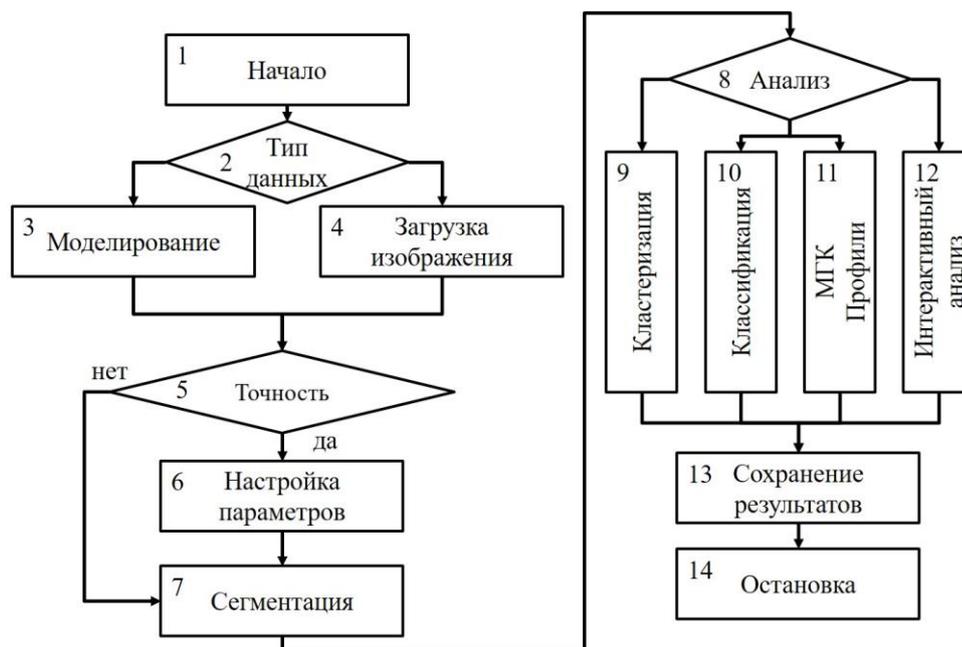


Рисунок 7. – Последовательность обработки данных в программном обеспечении CellDataMiner

В блоке МГК и блоке визуализации профилей признаков (блок 11 рисунка 7) используется МГК для представления данных и выделения значимой информации. Для представления в наглядной форме измеренных характеристик и соответствующих им групп объектов (ядер и клеток) используются первые две главные компоненты.

Программное обеспечение CellDataMiner было написано на языке М в среде MATLAB, интерфейс реализован с помощью конструктора графического интерфейса GUIDE. Представлены результаты анализа единичного изображения и группы изображений, полученных в ходе экспериментов по исследованию микрочипов тканей опухолей молочной железы.

Выполнено исследование методов классификации. Для большинства методов классификации точность увеличивается при использовании всех признаков, как формы, так и интенсивности пикселей. Все методы классификации показали значительное ухудшение точности при использовании только признаков формы. Метод опорных векторов и многослойный персептрон являются неустойчивыми к наличию выбросов в выборке, поэтому их использование на практике нецелесообразно. Наихудшие результаты были получены для метода опорных векторов, когда максимальная точность классификации составила 0,97 при использовании только признаков интенсивности. При этом результаты

классификации метода опорных векторов сильно зависят от выбранных параметров. Наиболее оптимальным методом для решения поставленной задачи является метод случайного леса, наилучшая точность классификации 0,97 не зависит значительно от параметров метода. Этот метод является устойчивым к наличию выбросов в обучающей выборке, т.к. в основе метода заложено коллективное решение по наборам деревьев. В группе байесовских методов при классификации по всем признакам и признакам интенсивности были получены результаты сопоставимые с методом случайного леса. Однако, при использовании только признаков формы точность классификации случайным лесом выше, чем у методов на основе теоремы Байеса.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1. Построена имитационная модель люминесцентных изображений раковых клеток молочной железы при исследовании рецептора-эстрогена, воспроизводящая несколько популяций клеток на одном изображении с учетом шума и размытия изображения. Разработанная модель позволяет учитывать различные красители, воспроизводить изображения со схожими признаками формы и интенсивности и была использована для исследования алгоритмов сегментации [1, 14, 15, 16, 22, 24, 26].

2. Разработан алгоритм сегментации цитоплазмы на изображении. Для восстановления информации о распределении красителя цитоплазмы использовался метод главных компонент. В отличие от других разработанный алгоритм оценивает размеры ядер на изображении, что позволяет выполнять автоматическую сегментацию наборов изображений [2, 6, 9, 10, 12, 13, 18].

3. Разработан алгоритм сегментации ядер на люминесцентных изображениях популяций клеток, основанный на использовании адаптивной пороговой сегментации и модифицированного водораздельного метода. Предложенный алгоритм снижает ошибку локализации в два раза и повышает точность сегментации ядер на изображении [2, 8, 11, 12, 19, 20, 22, 23, 25, 28].

4. Разработан алгоритм отбора информативных признаков, который основан на упорядочивании признаков по уровню информативности и построении наборов из наиболее информативных признаков. Количество признаков объектов было сокращено с 101 до 13. Использование только информативных признаков приводит к сокращению вычислительных затрат в 2 и более раза при сохранении сопоставимой точности анализа [4, 7, 29].

5. Программный пакет CellDataMiner для анализа изображений популяций раковых клеток, основанный на использовании объектно-ориентированной ме-

тодологии анализа люминесцентных изображений на пиксельном уровне, реализующий предложенный алгоритм сегментации объектов и позволяющий в автоматическом режиме проводить классификацию, кластеризацию и визуализацию результатов анализа изображений раковых клеток. Пакет позволяет выделять группы раковых и здоровых клеток с точностью классификации превосходящей 95% [3, 5, 14, 17, 21, 23, 27, 28, 29].

Рекомендации по практическому использованию результатов

1. Использование имитационной модели на практике позволяет объективно сравнивать различные алгоритмы анализа изображений по точности, исследовать алгоритмы на воздействие шума, перекрытие объектов. Такое сравнение невозможно при использовании только экспериментальных изображений, когда сравнение носит субъективный характер в силу ручной обработки изображений.

2. Применение разработанного алгоритма автоматической сегментации позволяет повысить скорость обработки данных, получаемых в медико-биологических исследованиях, оценить численно характеристики различных групп клеток и степень различия между ними, визуализировать полученные результаты. Увеличение точности сегментации и использование только информативных признаков позволяет повысить качество диагностики.

3. Практическая значимость полученных в диссертации результатов подтверждается актами внедрения в практическую деятельности государственного учреждения «Республиканский научно-практический центр онкологии и медицинской радиологии им. Н. Н. Александрова», в учебный процесс на факультете радиофизики и компьютерных технологий Белорусского государственного университета и справкой о внедрении в учебный процесс на факультете химической технологии и техники Белорусского государственного технического университета.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ УЧЕНОЙ СТЕПЕНИ

Статьи в научных изданиях в соответствии с п. 18 Положения о присуждении ученых степеней и присвоении ученых званий в Республике Беларусь и рецензируемых зарубежных изданиях

1. Имитационная модель трехканальных люминесцентных изображений популяций раковых клеток / Е. В. Лисица, Н. Н. Яцков, В. В. Апанасович, Т. В. Апанасович, М. М. Шитик // Журнал прикладной спектроскопии. – 2014. – Т. 81, № 6. – С. 907–913.

2. Алгоритм автоматической сегментации границ ядер раковых клеток на трехканальных люминесцентных изображениях / Е. В. Лисица, Н. Н. Яцков,

В. В. Апанасович, Т. В. Апанасович // Журнал прикладной спектроскопии. – 2015. – Т. 82, № 4. – С. 598–607.

3. Программный пакет CellDataMiner для анализа люминесцентных изображений раковых клеток / Е. В. Лисица, Н. Н. Яцков, В. В. Апанасович, Т. В. Апанасович // Информатика. – 2015. – № 4(48). – С. 73–84.

4. Отбор информативных геометрических признаков ядер клеток на люминесцентных изображениях раковых клеток / Е. В. Лисица, Н. Н. Яцков, В. В. Скакун, П. Д. Кривошеев, В. В. Апанасович // Информатика. – 2019. – Т. 16, № 2. – С. 7–17.

5. Исследование методов классификации для анализа сегментированных объектов на люминесцентных изображениях раковых клеток / Е. В. Лисица, Н. Н. Яцков, В. В. Скакун, В. В. Апанасович // Вестник Полоцкого государственного университета. Серия С. Фундаментальные науки. – 2020. – № 4. – С. 15–22.

Статьи в других научных рецензируемых изданиях

6. The correlation-based method for the movement compensation in the analysis of the results of FRAP Experiments / Y. Lisitsa, A. Halavatyi, E. Hadzic, E. Friederich, M. Yatskou, V. Apanasovich // Journal of Computational Optimization in Economics and Finance. – 2012. – Vol. 4, № 2/3. – P. 125–134.

7. Отбор характеристик распределения интенсивности в цветовых каналах на люминесцентных изображениях раковых клеток / Е. В. Лисица, В. В. Скакун, В. В. Апанасович // Журнал прикладной спектроскопии. – 2019. – Т. 86, № 3. – С. 394–400.

Статьи в сборниках материалов научных конференций

8. Лисица, Е. В. Объектно-ориентированный анализ люминесцентных изображений биологических объектов / Е. В. Лисица, Н. Н. Яцков // Сборник работ 68-ой науч. конф. студентов и аспирантов БГУ, Минск, 16–19 мая 2011 г. : в 3 ч. / Белорус. гос. ун-т. – Минск: БГУ, 2011. – Ч. 1. – С. 188–192.

9. Fully-automated segmentation of tumor nuclei in cancer tissue images / Y. U. Lisitsa, M. M. Yatskou, V. V. Apanasovich, H. Rui, T. V. Apanasovich // Pattern recognition and information processing (PRIP'2011) : proceedings of the 11th International Conference, Minsk, 18–20 May 2011 / Editorial board: R. Sadykhov [et al.]. – Minsk: BSUIR, 2011. – P. 116–120.

10. An algorithm for processing the frap images of actin filament treadmilling / Y. Lisitsa, A. Halavatyi, E. Hadzic, E. Friederich, M. Yatskou, V. Apanasovich // Международный конгресс по информатике: информационные системы и технологии : материалы междунар. науч. конгресса, Минск, 31 октября – 3 ноября

2011 г. : в 2 ч. / Редкол.: С. В. Абламейко [и др.]. – Минск: БГУ, 2011. – Ч. 2. – С. 238–242.

11. A fully-automated nuclear segmentation approach for analysis of cell fluorescent images / Y. Lisitsa, M. Yatskou, T. Apanasovich, V. Apanasovich // Международный конгресс по информатике: информационные системы и технологии : материалы междунар. науч. конгресса, Минск, 31 октября – 3 ноября 2011 г. : в 2 ч. / Редкол.: С. В. Абламейко [и др.]. – Минск: БГУ, 2011. – Ч. 2. – С. 242–247.

12. Алгоритмы сглаживания контуров биологических объектов на люминесцентных изображениях / Е. В. Лисица, Н. Н. Яцков, В. В. Апанасович, Т. В. Апанасович // Сборник материалов II Междунар. науч.-практ. конф. молодых ученых «Научные стремления – 2011», Минск, 14–18 ноября 2011 г. : в 2 т. / Совет молодых ученых НАН Беларуси ; редкол.: А. Н. Волченко [и др.]. – Минск: «Беларуская навука», 2011. – Т. 1. – С.740–744.

13. The Correlation-Based Method for the Movement Compensation in the Analysis of the Results of FRAP Experiments / Y. Lisitsa, A. Halavatyi, E. Hadzic, E. Friederich, M. Yatskou, V. Apanasovich // Modeling and simulation MS'2012 : proceedings of the International Conference, Minsk, 2–4 May, 2012 / Editors: V. Krasnoproshin [et al.]. – Minsk: Publ. Center of BSU, 2012. – P. 8–10.

14. Лисица, Е. В. Применение искусственных нейронных сетей для классификации раковых клеток / Е. В. Лисица, Н. Н. Яцков, Т. В. Апанасович // Информационные технологии и системы 2012 (ИТС 2012) : материалы междунар. науч. конф., Минск, 24 октября 2012 г. / Редкол.: Л. Ю. Шилин [и др.]. – Минск: БГУИР, 2012. – С. 254–255.

15. Воспроизведение люминесцентных изображений раковых клеток с использованием имитационных моделей / М. М. Шитик, Е. В. Лисица, Н. Н. Яцков, В. В. Апанасович // Информационные технологии и системы 2012 (ИТС 2012) : материалы междунар. науч. конф., Минск, 24 октября 2012 г. / Редкол.: Л. Ю. Шилин [и др.]. – Минск: БГУИР, 2012. – С. 256–257.

16. Воспроизведение люминесцентных изображений раковых клеток с использованием имитационных моделей / М. М. Шитик, Е. В. Лисица, Н. Н. Яцков, В. В. Апанасович // Сборник материалов III междунар. молодежной науч.-практ. конф. «Научные стремления – 2012», Минск, 6–9 ноября 2012 г. : в 2 т. / Совет молодых ученых НАН Беларуси ; редкол.: В. В. Казбанов [и др.]. – Минск: «Беларуская навука», 2012. – Т. 1. – С. 455–458.

17. Программный пакет CELLDATAMINER для интеллектуального анализа люминесцентных изображений раковых клеток / Е.В. Лисица, Н. Н. Яцков, Т. В. Апанасович, В. В. Апанасович // Медэлектроника – 2012. Средства медицинской электроники и новые медицинские технологии : сб. науч. ст. VII меж-

дунар. науч.-техн. конф., Минск, 13–14 декабря 2012 г. / МО Респ. Бел.; МЗ Респ. Бел.; ГКНТ; БГУИР; БелМАПО; БелИСА; редкол.: В. С. Улащик [и др.]. – Минск: БГУИР, 2012. – С. 67–70.

18. Антоненко, М. В. Сегментация конфокальных изображений клеток / М. В. Антоненко, Е. В. Лисица // Информационные технологии и системы 2013 (ИТС 2013) : материалы междунар. науч. конф., Минск, 23 октября 2013 г. / Редкол.: Л. Ю. Шилин [и др.]. – Минск: БГУИР, 2013. – С. 238–239.

19. Исследование алгоритмов разделения слившихся объектов / Е. В. Лисица, Н. Н. Яцков, В. В. Апанасович, Т. В. Апанасович // Информационные технологии и системы 2013 (ИТС 2013) : материалы междунар. науч. конф., Минск, 23 октября 2013 г. / Редкол.: Л. Ю. Шилин [и др.]. – Минск: БГУИР, 2013. – С. 284–285.

20. Шитик, М. М. Применение алгоритма отслеживания объектов для видеопоследовательностей эмбриональных клеток мыши / М. М. Шитик, Е. В. Лисица, В. В. Апанасович // Информационные технологии и системы 2013 (ИТС 2013) : материалы междунар. науч. конф., Минск, 23 октября 2013 г. / Редкол.: Л. Ю. Шилин [и др.]. – Минск: БГУИР, 2013. – С. 320–321.

21. Лисица, Е. В. Классификация объектов люминесцентных изображений / Е. В. Лисица, Н. Н. Яцков, В. В. Апанасович // Международный конгресс по информатике: информационные системы и технологии : материалы междунар. науч. конгресса, Минск, 4–7 ноября 2013 г. / Редкол.: С. В. Абламейко [и др.]. – Минск: БГУ, 2013. – С. 59–64.

22. Studying the advanced segmentation methods with the computer simulated images / Y. Lisitsa, M. Yatskou, V. Apanasovich, T. Apanasovich // Pattern Recognition and Information Processing (PRIP'2014) : proceedings of the 12th International Conference, Minsk, 28–30 May 2014 / Editors: Al. Tuzikov [et al.]. – Minsk: UIIP NASB, 2014. – P. 157–161.

23. Лисица, Е. В. Анализ люминесцентных изображений объектов микробиологии с использованием метода главных компонент / Е. В. Лисица, Н. Н. Яцков, В. В. Апанасович // Информационные технологии и системы 2014 (ИТС 2014) : материалы междунар. науч. конф., Минск, 29 октября 2014 г. / Редкол.: Л. Ю. Шилин [и др.]. – Минск: БГУИР, 2014. – С. 280–281.

24. Шитик, М. М. CellPainter – программный пакет для моделирования люминесцентных изображений микробиологических объектов / М. М. Шитик, Е. В. Лисица, Н. Н. Яцков // Медэлектроника – 2014. Средства медицинской электроники и новые медицинские технологии : сб. науч. ст. VIII междунар. науч.-техн. конф., Минск, 10–11 декабря 2014 г. / МО Респ. Бел.; МЗ Респ. Бел.; БГУИР; БелМАПО; БГМУ; Cisco Systems, Inc.; редкол.: В. С. Улащик [и др.]. – Минск: БГУИР, 2014. – С. 140–143.

25. Лисица, Е. В. Исследование методов распознавания на изображениях рентгенограмм / Е. В. Лисица, Е. Е. Рак // Информационные технологии и системы 2015 (ИТС 2015) : материалы междунар. науч. конф., Минск, 28 октября 2015 г. / Редкол.: Л. Ю. Шилин [и др.]. – Минск: БГУИР, 2015. – С. 288–289.

26. Формализованное описание многоканальных люминесцентных изображений раковых клеток [Электронный ресурс] / Е. В. Лисица, Н. Н. Яцков, В. В. Апанасович, Т. В. Апанасович, В. В. Скакун // International Congress on Computer Science: Information Systems and Technologies: материалы междунар. науч. конгресса, Республика Беларусь, Минск, 24–27 окт. 2016 г. / Редкол.: С. В. Абламейко [и др.]. – Минск: БГУ, 2016. – С. 623–628. – Режим доступа: <https://elib.bsu.by/handle/123456789/160466>. – Дата доступа: 24.08.2018.

27. Классификация раковых клеток на люминесцентных изображениях с использованием методов дискриминантного анализа / Е. В. Лисица, Н. Н. Яцков, В. В. Скакун, В. В. Апанасович // Прикладные проблемы оптики, информатики, радиофизики и физики конденсированного состояния : материалы V Междунар. науч.-практ. конф., Минск, 16–17 мая 2019 г. / МО Респ. Бел.; Ин-т приклад. физ. проблем Белорус. гос. ун-та ; редкол.: В. И. Попечиц [и др.]. – Минск, 2019. – С. 149–151.

Тезисы докладов

28. Анализ изображений популяций клеток в системах люминесцентной микроскопии / Е. В. Лисица, Н. Н. Яцков, В. В. Апанасович, В. В. Скакун // Генетика и биотехнология XXI века: проблемы, достижения, перспективы (к 50-летию ГНУ «Институт генетики и цитологии НАН Беларуси»): мат. II Междунар. науч. конф., Минск, 13–16 октября 2015 г. / Бел. о-во генетиков и селекционеров ; редкол.: А. В. Кильчевский [и др.]. – Минск: Ин-т генетики и цитологии НАН Беларуси, 2015. – С. 36.

29. Измеряемый коэффициент диффузии и плотность расположения опухолевых клеток в лимфатических узлах при лимфомах / С. А. Хоружик, С. В. Шиманец, И. С. Дулинец, О. Р. Анискевич, Н. М. Тризна, Т. В. Францкевич, Е. В. Лисица, Н. Н. Яцков, В. В. Скакун, Э. А. Жаврид // Евразийский онкологический журнал. – 2016. – Т. 4, № 2: : IX Съезд онкологов и радиологов стран СНГ и Евразии, Минск, 15–17 июня 2016 г. – С. 561.

30. Отбор значимых признаков методами иерархического кластерного анализа / Е. В. Лисица, В. В. Скакун, Н. Н. Яцков, В. В. Апанасович // Информационные технологии и системы 2017 (ИТС 2017) : материалы междунар. науч. конф., Минск, 25 октября 2017 г. / Редкол.: Л. Ю. Шилин [и др.]. – Минск: БГУИР, 2017. – С. 100–101.



РЕЗЮМЕ

Лисица Евгения Владимировна

**Анализ изображений популяций раковых клеток
в системах обработки данных люминесцентной микроскопии**

Ключевые слова: обработка изображений, классификация, кластеризация, программное обеспечение, моделирование люминесцентных изображений.

Цель работы: разработка моделей, алгоритмов и программных средств для анализа люминесцентных изображений раковых клеток, полученных в системах люминесцентной микроскопии при диагностике онкологических заболеваний.

Методы исследования и использованная аппаратура: имитационное моделирование, методы обработки изображений, теория вероятности.

Полученные результаты и их новизна: Разработана имитационная модель для воспроизведения люминесцентных изображений с двумя колониями клеток, которая включает в себя возможность воспроизведения интенсивности люминесценции в ядрах и цитоплазме, расположения в пространстве, учета искажений, возникающих в оптических узлах системы регистрации. Разработан алгоритм адаптивной сегментации ядер на люминесцентных изображениях раковых клеток, в основе которого лежит учет корреляции между пикселями соседних каналов за счет перекрытия спектров испускания и поглощения между цветовыми каналами в системе регистрации, используя метод главных компонент. Разработан алгоритм отбора информативных признаков для описания объектов на люминесцентных изображениях и с его помощью отобраны 13 наиболее информативных признаков. Выполнено сравнение методов классификации для решения разбиения объектов на группы. Разработано программное обеспечение, которое позволяет выполнять интерактивный анализ изображений и включает в себя автоматическую сегментацию изображений, построение имитационных изображений, возможность решения задач классификации и кластеризации, отображения данных в интерактивном режиме.

Рекомендации по использованию: Полученные результаты могут быть использованы при исследовании онкологических заболеваний, выявления злокачественных опухолей.

Область применения: Предложенные средства являются новыми инструментами для повышения эффективности анализа иммуногистохимических данных.

Лісіца Яўгенія Уладзіміраўна

Аналіз выяў папуляцый ракавых клетак у сістэмах апрацоўкі дадзеных люмінесцэнтнай мікраскапіі

Ключавыя словы: апрацоўка выяў, класіфікацыя, кластарызацыя, праграмае забеспячэнне, мадэляванне люмінесцэнтных выяў.

Мэта работы: распрацоўка мадэляў, алгарытмаў і праграмных сродкаў для аналізу люмінесцэнтных выяў ракавых клетак, атрыманых у сістэмах люмінесцэнтнай мікраскапіі.

Метады даследавання і скарыстаная апаратура: імітацыйнае мадэляванне, метады апрацоўкі выяў, тэорыя верагоднасці.

Атрыманыя вынікі і іх навізна: Распрацавана імітацыйная мадэль для мадэлявання люмінесцэнтных выяў з дзвюма калоніямі клетак, якая ўключае ў сябе магчымасць мадэлявання інтэнсіўнасці люмінесцэнцыі ў ядрах і цытаплазме, размяшчэння ў прасторы, уліку скажэнняў, якія ўзнікаюць у аптычных вузлах сістэмы рэгістрацыі. Распрацаваны алгарытм адаптыўнай сегментацыі ядраў на люмінесцэнтных выявах ракавых клетак, у аснове якога ляжыць аднаўленне інфармацыі за кошт перакрыцця спектраў выпускавання і паглынання паміж каляровымі каналамі ў сістэме рэгістрацыі, выкарыстоўваючы метады галоўных кампанент. Распрацаваны алгарытм адбору інфарматыўных прыкмет на люмінесцэнтных выявах і з яго дапамогай адабрана 13 самых інфарматыўных прыкмет. Выканана параўнанне метадаў класіфікацыі для вырашэння разбіцця аб'ектаў на групы. Распрацавана праграмае забеспячэнне, якое дазваляе выконваць інтэрактыўны аналіз выяў і ўключае ў сябе аўтаматычную сегментацыю выяў, пабудову імітацыйных выяў, магчымасць вырашэння задач класіфікацыі і кластарызацыі, адлюстравання дадзеных у інтэрактыўным рэжыме.

Рэкамендацыі па выкарыстанні: Атрыманыя вынікі могуць быць выкарыстаны пры даследаванні анкалагічных захворванняў, выяўлення злаякасных пухлін.

Галіна ўжывання: Прапанаваныя сродкі з'яўляюцца новымі інструментамі для павышэння эфектыўнасці аналізу імунагістахімічных дадзеных.

SUMMARY

Yauheniya Uladzimirauna Lisitsa
Image analysis of populations of cancer cells
in the data processing systems of fluorescent microscopy

Keywords: image processing, classification, clustering, software, modeling of fluorescent images.

Aim of study: to develop models, algorithms, and software for analyzing fluorescent images of cancer cells obtained in fluorescent microscopy systems.

Methods and used equipment: simulation, image processing techniques, probability theory.

Results of study and their novelty: The simulation model was developed to produce fluorescent images with two cell colonies. The model is able to reproduce the fluorescent intensity in the nuclei and cytoplasm, location in space. It takes into account optical distortions arising in the optical nodes of the registration system. The adaptive algorithm was developed to segment nuclei on fluorescent images of cancer cells. The algorithm is based on the restoration of information due to overlapping emission and absorption spectra between color channels in the registration system. It exploits the principal component method. An algorithm for selecting informative features for describing objects on fluorescent images has been developed and 13 most informative features have been selected by it. The comparison was made of classification methods to solve the splitting of objects into groups. The software was developed that allows to perform interactive image analysis and includes automatic image processing, image simulation to study the parameters of the algorithm, the ability to solve classification and clustering problems, and display data in an interactive mode.

Application guidelines: The results can be used in the study of cancer, the detection of malignant tumors.

Application area: The proposed tools are new tools to improve the efficiency of the analysis of immunohistochemical data.

