

ИССЛЕДОВАНИЕ ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ КОЛЛАГЕНА С ПРИМЕНЕНИЕМ МЕТОДОВ МНОГОФОТОННОЙ МИКРОСКОПИИ И МАШИННОГО ОБУЧЕНИЯ

©2019 г. Ю. В. КИСТЕНЕВ^{1,2}, Д. А. ВРАЖНОВ¹,
В. В. НИКОЛАЕВ¹, Е. А. САНДЫКОВА^{1,2},
Н. А. КРИВОВА¹

¹Томский государственный университет, Томск

²Сибирский государственный медицинский университет, Томск

I. Введение. II. Инструментальные методы визуализации пространственной структуры коллагена. III. Методы количественного анализа пространственной структуры коллагена. IV. Применение методов машинного обучения для количественного анализа пространственной структуры коллагена. V. Заключение.

I. ВВЕДЕНИЕ

Настоящий обзор посвящен анализу подходов и имеющихся результатов в области выявления информативных характеристик в экспериментальных данных инфракрасной многофотонной мультимодальной микроскопии коллагеновых структур в биоткани и разработке на этой основе алгоритмов автоматической классификации дезорганизации пространственной структуры коллагена при различных заболеваниях с использованием машинного обучения – разновидности методов искусственного интеллекта.

Коллагены – семейство фибриллярных белков, секретируемых клетками соединительной ткани. Коллагеновые белки составляют не

Принятые сокращения: ИК – инфракрасный; МРТ – магнито-резонансная томография; AGE – advanced glycation endproducts (конечные продукты гликирования); ANN – Artificial Neural Network (искусственная нейронная сеть); AUC – area under ROC curve (площадь под ROC-кривой); CNN – Convolutional

Окончание принятых сокращений см. на сл. стр.

Адрес для корреспонденции: yuk@iao.ru, Кистенев Юрий Владимирович.

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований – грант No.17-00-00186. Исследование выполнено при финансовой поддержке РФФИ и Администрации Томской области в рамках научного проекта № 18-42-703012.

менее 30% общей массы белков в теле человека [1]. Пространственная структура белков, в зависимости от последовательности аминокислот, может иметь вид α -цепи или β -складчатой структуры. Характерной особенностью коллагена является наличие протяженных 3-спиральных α -цепей (рис. 1а), называемых тропоколлагеном.

Молекулы коллагена в межклеточном матриксе (extracellular matrix, ECM) образуют полимеры, называемые фибриллами (рис. 1б). Фибриллы состоят из повторяющихся тропоколлагеновых структур, уложенных вдоль волокна в виде параллельных пучков по типу «голова к хвосту». В параллельных рядах молекулы тропоколлагена сдвинуты относительно друг друга на расстояние 64 нм. Этим объясняются характерные для фибрилл поперечные сшивки, которые повторяются с тем же периодом. Более толстые пучки фибрилл могут включать гликопротеины, протеогликаны [2].

Особенности 3-спиральных доменов в коллагене заключаются в следующем [2]: остатки аминокислот в α -цепях представлены однотипными, регулярно повторяющимися трипептидами (рис. 1б), где 1-я позиция занята остатком глицина (глицилом, – GLY), 2-я – как правило, пролином (пролилом, PRO), 3-я – остатком других аминокислот, наиболее часто – гидроксипролином (HYP) или гидроксизинном.

В настоящее время описано 28 типов коллагена, которые отличаются друг от друга по аминокислотной последовательности [2], в

Принятые сокращения, окончание

Neural Networks (сверточная нейронная сеть); CPA – collagen proportional area (относительная площадь коллагена); DTI – Diffusion tensor imaging (диффузионная тензорная визуализация); ECM – extracellular matrix (межклеточный матрикс); FLIM – fluorescence lifetime imaging (флуоресцентный имиджинг с временным разрешением); FOS – first order statistics (стагистики первого порядка); GLCM – Gray-Level Co-Occurrence Matrix (матрица смежности); MP – metalloproteinases (металлопротеиназы); MPM – multi photon microscopy (многофотонная микроскопия); PCA – Principal component analysis (метод главных компонент); ReLU – Rectified Linear Unit (усеченное линейное преобразование); ROC – Receiver operating characteristic (рабочая характеристика приёмника); ROS – Reactive oxygen species (активные формы кислорода); SIFT – Scale Invariant Feature Transform (масштабно инвариантное преобразование характеристик); SHG – second harmonic generation (генерация второй гармоники); SVM – Support vector machine (машина опорных векторов); TCSPC – Time Correlated Single Photon Counting (метод счета одиночных фотонов с корреляцией по времени); TDS – train dataset (обучающая выборка); TPF – two-photon fluorescence (флуоресценция, возбуждаемая двухфотонным поглощением); TPM – two photon microscopy (двухфотонная микроскопия).

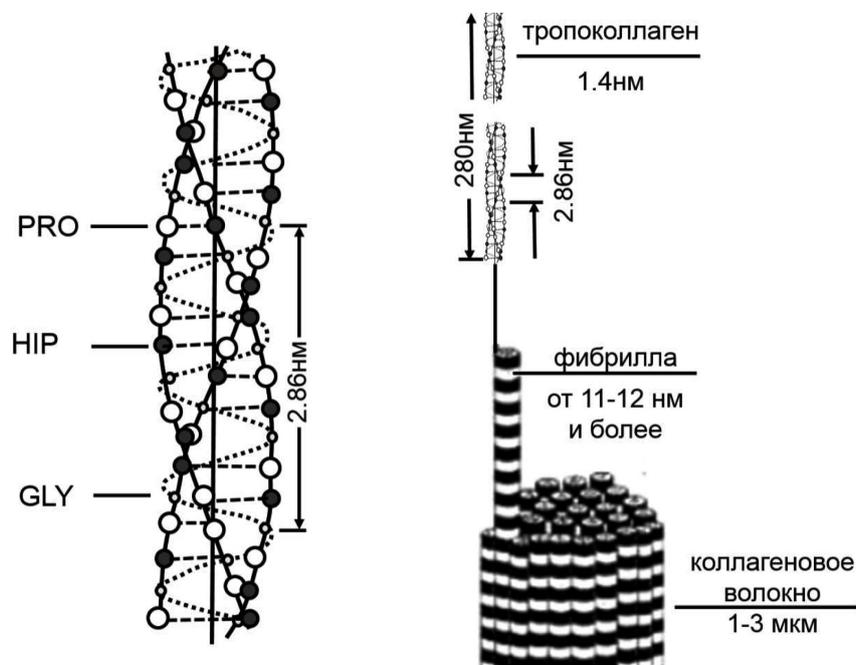


Рис. 1. Пространственная структура коллагеновых волокон: а) белок в виде α -цепи [4, 6], б) строение коллагена [5, 6] (адаптировано с изменениями из работы [6]).

частности, в состав дермы входят волокна, представленные в основном коллагенами фибриллярного типа.

Коллаген не только является основной частью ЕСМ, но и поддерживает функционирование клеток, включая адгезию, миграцию, дифференцировку, морфогенезис [3].

Синтез коллагена включает два этапа. На внутриклеточном этапе происходит трансляция и посттрансляционная модификация полипептидных цепей. Пептидные α -цепи коллагена синтезируются на полирибосомах, связанных с мембранами эндоплазматической сети. Из эндоплазматической сети молекулы синтезированного проколлагена перемещаются в аппарат Гольджи, где они включаются в секреторные пузырьки и в их составе секреторируются во внеклеточное пространство. Внеклеточный этап – модификация молекул проколлагена. В межклеточном пространстве от молекулы проколлагена отщепляются N- и C-концевые пептиды при участии протеолитических ферментов, и освобождается тройная спираль коллагена. Далее происходит процесс самосборки коллагеновых фибрилл, фиксиро-

ванных межмолекулярными ковалентными связями (сшивками), завершающий образование коллагеновых волокон [7].

Согласно исследованиям [8], деструктивные процессы в коллагене характеризуются изменением длины, толщины волокон, и их ориентацией в пространстве. Дезорганизация коллагена имеет ферментативные и неферментативные механизмы. Ферментативные механизмы осуществляются взаимодействием деполимеризующих коллагеназ и их ингибиторов, обеспечивающих регуляцию интенсивности каталитических процессов. Неферментативные механизмы связаны с образованием дополнительных межмолекулярных связей под действием конечных продуктов гликирования (Advanced Glycation End-products, AGE), ультрафиолетового облучения или механического давления (отек), а также в результате канцерогенеза. Различные регуляторные пути и механизмы деградации пересекаются друг с другом на разных этапах, а не работают независимо друг от друга [9].

Основным механизмом повреждения структуры коллагена является неферментативное гликирование [10]. Конечные продукты гликирования изменяют свойства коллагена, такие как потеря растворимости тройной спирали, увеличение его жесткости [11]. Гликирование коллагена усиливает образование и миграцию миофибробластов, участвует в развитии фиброза у больных сахарным диабетом [12]. Гликозилированный коллаген изменяет функции эндотелиальных клеток и может быть важным фактором в развитии атеросклеротических бляшек [13]. Он также может влиять на образование метастазов опухолевых клеток [10]. Деструкция коллагена в соединительной ткани наблюдается при артериосклерозе в результате химического воздействия глюкозы и ее метаболитов – гликоксинов (глиоксаля и метилглиоксаля) [14].

Деградация коллагена может быть обусловлена также действием активных форм кислорода (Reactive oxygen species, ROS) при оксидативном стрессе. ROS индуцируют синтез металлопротеиназ (metalloproteinases, MP) межклеточного вещества, которые, в свою очередь, вызывают разрушение фибробластов и подавление синтеза ЕСМ [15, 16]. Данный процесс характерен для различных тканей, включая сердечную мышцу [15, 17].

Развитие вторичной лимфедемы сопровождается изменением уровня MP [18], что также может вызывать деструкцию коллагена. Так, на изображениях ткани нижней конечности крысы при лимфедеме (рис. 2), полученных методом двухфотонной лазерной микроскопии, можно заметить наличие участков дезорганизации коллагеновых структур и их отсутствие в здоровой ткани [19].

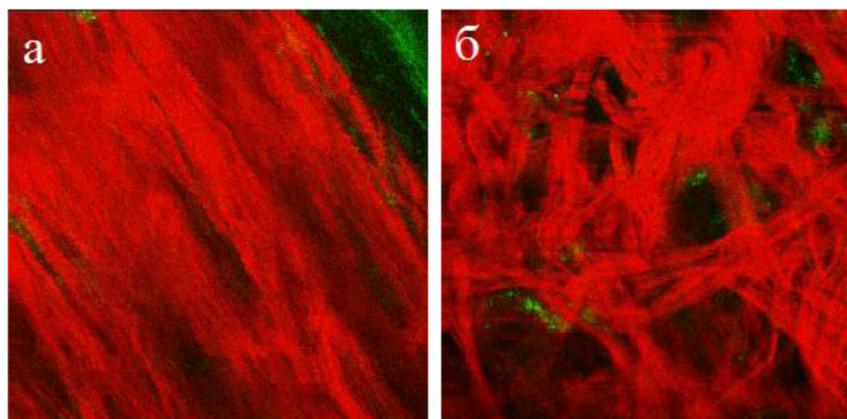


Рис. 2. Развитие лимфедемы в задней лапе крысы: здоровая ткань (а) и ткань с лимфедемой (б).

Гомеостаз ЕСМ важен для нормальной функции ткани, нарушение его регуляции тесно связано с фиброзом и онкологическими процессами. ЕСМ является основным компонентом микроокружения опухоли. Коллаген, образуя каркас ЕСМ, способствует инфильтрации опухоли, ангиогенезу, инвазии и миграции опухолевых клеток [20]. Показано, что аномальное пространственное распределение коллагена связано с быстрым ростом опухоли и устойчивостью к терапии. В частности, линейно структурированные коллагеновые волокна являются проводниками движения онкологических клеток, усиливая метастазирование. Плотные коллагеновые структуры затрудняют диффузию лекарств. Таким образом, подавление фиброза считается перспективной стратегией терапии рака, особенно при раке поджелудочной железы [21].

Гетерогенная организация коллагеновых волокон в перипапиллярной склере способствует снижению уровня деформации в решетчатой пластинке на границах каналов склеры и уменьшению кровоизлияния в диске зрительного нерва при глаукоме. Таким образом, количество и анизотропия коллагеновых волокон в склере являются важными факторами для поддержания соответствующих биомеханических свойств в решетчатой пластинке склеры [22].

На модели кожного лоскута была продемонстрирована ключевая роль модифицирования коллагена при старении [23]. Выявлено изменение механических свойств коллагеновых волокон корнеосклеральной оболочки с возрастом, что является одной из причин увеличения жесткости склеры и роговицы при старении [24].

Уменьшение синтеза коллагена фибробластами десны было обнаружено у пациентов с диабетом [25]. При ожирении в тканях наблюдается увеличение содержания коллагена 5-го типа и соответствующее изменение ЕСМ [3].

Таким образом, изменение пространственной структуры коллагена является важным диагностическим маркером большого числа заболеваний. Соответственно, представляют интерес методы прямой регистрации пространственной структуры коллагеновых волокон.

II. ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ ВИЗУАЛИЗАЦИИ ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ КОЛЛАГЕНА

Одним из основных технических требований при выборе метода визуализации является обеспечение достаточного пространственного разрешения для наблюдения исследуемых структур. Коллагеновые волокна (пучки фибрилл) имеют диаметр 0.5–12 мкм, отдельные фибриллы – 5–400 нм, коллагеновые фибриллы склеры – 25–230 нм, фибриллы роговицы глаза – 26–30 нм, миофибриллы – 1 мкм, фибриллы мышц – 5–15 нм [26]. Таким образом, оптические методы оптимальны для прямой визуализации пространственной структуры коллагеновых волокон.

Традиционно, исследование пространственной структуры коллагена в биологических тканях проводится с использованием гистохимии, иммуногистохимии, гибридации *in situ* (см., например, [27, 28]).

В работе [29] был проведен сравнительный анализ трех методов исследования пространственной ориентации коллагена хрящевой ткани.

Рутинным инструментом исследования *ex vivo* коллагена (например, в парафиновых блоках) является поляризационная световая микроскопия. Высоко ориентированные коллагеновые фибриллы обладают двулучепреломлением, таким образом, регистрация изменения поляризации света, прошедшего через парафиновые блоки с образцами ткани, позволяет косвенно оценивать пространственную организацию и количество коллагена в образце [30]. Сравнение проводилось с методом магнито-резонансной томографии (МРТ и многофотонной лазерной микроскопии (multi-photon microscopy, МРМ).

Количественное T_2 -МРТ картирование дает возможность неинвазивного *in vivo* биохимического анализа хряща, учитывая корреляцию времени T_2 -релаксации гиалинового хряща и дезорганизации структуры коллагена. В основе корреляции – зависимость содержания воды в ткани от пространственной структуры коллагена [31]. Данный метод

является непрямым для исследования структуры коллагена с учетом пространственного разрешения МРТ (в работе [29] использовался 3Т клинический томограф GE Healthcare с пространственным разрешением около 200 мкм).

Отметим, что для МРТ – картирования структуры коллагена также может быть использована диффузионная тензорная визуализация (Diffusion tensor imaging, DTI) – модальность визуализации, которая предоставляет информацию о пространственной анизотропии диффузии воды (скорость диффузии воды вдоль коллагеновых волокон выше, чем в поперечном направлении) [32]. Таким образом, тензор диффузии дает информацию об усредненной по объему 3D ориентации волокна. В работе [33] использована тензорная томография восприимчивости для анализа пространственной структуры коллагена. Был использован 9.4Т магнито-резонансный томограф для исследования коллагена в свиных суставных хрящах, пространственное разрешение (размер вокселя) достигало 100 мкм. Следует отметить, что высокопольные томографы уровня 9Т пока используются только для научных исследований.

В основе МРМ лежит многофотонное возбуждение оптических процессов в тканях ИК излучением с длиной волны в диапазоне 0.7-0.8 мкм (область одного из спектральных окон прозрачности кожи [34, 35]). При этом за счет поглощения сразу нескольких фотонов (чаще – двух) реализуется воздействие на ткань на длине волны в спектральной области 0.3–0.4 мкм. Возбуждение нелинейных эффектов возникает при высокой интенсивности излучения, которое достигается путем фокусировки пучка в локальной области среды, что повышает осевое разрешение метода (до единиц мкм), обеспечивая визуализацию пространственной структуры ткани с субклеточным разрешением.

Авторами [29] использовался оригинальный МРМ микроскоп на основе MaiTai Ti:sapphire лазера (SpectraPhysics, Mountain View, CA) и микроскопа Olympus IX-70.

Все три метода показали возможность количественных измерений дезорганизации коллагена, но с существенно различным пространственным разрешением, причем, лишь МРМ является прямым методом наблюдения структуры коллагена [29].

В последнее время для визуализации структуры биоткани *in vivo* и *in vitro* все чаще применяют технику двухфотонной микроскопии (two-photon microscopy, ТРМ) – разновидности МРМ. Двухфотонная микроскопия основана на использовании двух модальностей: флуоресценции ткани, возбуждаемой двухфотонным поглощением (two-

photon fluorescence, TPF) и генерации второй гармоники (second harmonic generation, SHG) отдельными биомолекулами.

Молекулы, обладающие нецентросимметричной структурой, являются особенно сильными излучателями SHG из-за наличия ненулевой восприимчивости второго порядка. Поскольку молекулы коллагена образуют тройную геликоидальную спираль, они не имеют центра симметрии, что обуславливает возможность их визуализации на основе сигнала SHG [36, 37]. SHG-визуализация является одним из самых надежных методов прямой визуализации коллагеновых структур биотканей [38], соответственно, сигнал второй гармоники может характеризовать деструкцию коллагеновой структуры [39–44]. Как правило, явление SHG не сопровождается фотообесцвечиванием, не фототоксично [36]. Визуализация TPF также является весьма информативным методом исследования биотканей, будучи весьма чувствительным к их составу [45].

Метод флуоресцентного имиджинга с временным разрешением (Fluorescence lifetime imaging, FLIM) дает возможность различать отдельные флуорофоры на временном интервале, соответствующем их времени жизни флуоресценции. Исследование автофлуоресценции тканей методом FLIM обеспечивает высокий контраст картирования эндогенных флуорофоров в тканях, позволяет исследовать взаимодействия белков и их конформационные превращения, распределение на клеточном уровне кислорода, показателя преломления биоткани [46].

Сочетание ТРМ и FLIM позволяет реализовать комбинацию модальностей – генерацию и прием гармоник, многофотонное возбуждение флуоресценции и ее регистрацию с временным разрешением, что повышает диагностическую информативность метода [40].

В данной области актуально развитие подходов количественной оценки взаимосвязей патологических процессов, выражающихся в деструкции коллагеновой структуры, и характеристик ТРМ изображений [8].

III. МЕТОДЫ КОЛИЧЕСТВЕННОГО АНАЛИЗА ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ КОЛЛАГЕНА ВЫДЕЛЕНИЕ ОБЛАСТЕЙ ИНТЕРЕСА И ИНФОРМАТИВНЫХ ПРИЗНАКОВ НА МРМ ИЗОБРАЖЕНИЯХ

Как отмечалось выше, дезорганизация пространственной структуры коллагена в тканях является биомаркером ряда заболеваний, соответственно при использовании прямых методов визуализации коллагена возникает задача поиска и количественного описания характерных структур на изображениях.

Автоматический поиск характерных структур на изображении базируется на их формализованном математическом описании. Стандартный подход к решению этой задачи основан на понятии текстуры. Следует отметить, что строгого определения текстуры не существует. Например, в биомедицине под текстурой понимается микро- и макроструктурные свойства тканей [47], в компьютерном зрении текстурой принято называть область гомогенных пикселей на изображении [48].

Известны различные подходы выделения текстур, включая анализ градиентов, например, Scale Invariant Feature Transform (SIFT) [49], гистограммы ориентированных градиентов (Histogram of Oriented Gradients) [50]; эджлеты (Edgelets) [51], цветовые коррелограммы [52], локальные бинарные шаблоны (Locally Binary Patterns) [53], методы на основе поворачивающихся пирамид (Steerable Pyramids) [54] и другие. Остановимся на наиболее часто используемых подходах выделения текстур.

Структурные методы опираются на представление текстуры в виде композиции визуальных примитивов. Примером данного подхода служат методы математической морфологии [55, 56]. Основными характеристиками бинарной морфологии служат структурный элемент, дилатация, эрозия, открытие, закрытие. Под структурным элементом понимается некое двоичное изображение произвольной геометрической формы. Дилатацией (наращиванием) называется результат логического ИЛИ структурного элемента и исходного изображения. Эрозия – результат логического И структурного элемента и всего изображения. Закрытием (замыканием) называется последовательное применение операции дилатации и эрозии к исходному изображению. Открытием (размыканием) называется последовательное применение операций эрозии и дилатации.

Статистическое описание количества оставшихся пикселей после применения морфологических операций со структурным элементом разного размера позволяет вычислить так называемую гранулометрию и рассматривать шероховатость изображения как текстурную характеристику изображения.

Для локализации текстур используется метод скользящего окна [57].

Спектральные методы описания текстур используют преобразования Фурье, Габора, вейвлет-преобразования [58–62]. Идея данных методов заключается в выявлении характерных особенностей изображения с применением интегральных фильтров с подобранными параметрами, и затем использовании их в качестве признаков. Каждый фильтр играет роль пространственного шаблона текстуры, свертка с которым исходного изображения дает отклик, вычисляя величину

которого, можно судить о схожести части изображения и шаблона. На практике вычисляются коэффициенты разложения исходного сигнала по базису (Фурье, Габора, материнским вейвлетам), а анализ этих коэффициентов позволяет установить корреляцию шаблона текстуры и части изображения.

Модельные методы описывают текстуру в виде ее математической модели, например, в виде модели фрактала [63, 64]. В практических приложениях фрактального анализа, объект, как правило, характеризуется размерностью покрытия. При этом исходное изображение рассматривается как трехмерная область, и ищется число кубов разного размера r , заполняющего данную область [65]. Например, размерность Минковского при покрытии изображения кубами размером r вычисляется по формуле

$$D = -\frac{\log(N(r))}{\log(r)},$$

где $N(r)$ – количество кубов размера r . Величина D хорошо описывает шероховатость исходной поверхности.

Статистические методы описания текстур используют свойства статистики распределения цветовых или серых оттенков пикселей в некоторой области изображения [8].

Статистики первого порядка (first order statistics, FOS) напрямую связаны с распределением интенсивности пикселей (гистограмма яркости) и не учитывают межпиксельные корреляции. Гистограмма яркости имеет вид [66]

$$H(g) = \frac{n_g}{N}, \quad g = 0, 1, \dots, G-1, \quad (1)$$

где N – число пикселей в изображении, G – число градаций яркости, n_g – число пикселей на изображении с определенным значением яркости.

Гистограмма яркости имеет смысл плотности вероятности и позволяет рассчитать статистические характеристики изображения. Примером FOS является среднее значение яркости изображения, стандартное отклонение (контраст), энтропия, статистические моменты. Гистограмма яркости может быть обобщена для цветных изображений.

Статистики второго порядка рассчитываются по вероятности наблюдения определенных значений пары пикселей изображения. Примером статистики второго порядка является матрица смежности. Матрица смежности (Gray-Level Co-Occurrence Matrix, GLCM)

была предложена в 1973 году Робертом Хараликом [67] и получила огромное распространение в задачах выделения текстур [68].

Рассмотрим построение GLCM $C_d(i, j)$ для изображения $I(x, y)$, представленного в p градациях серого. GLCM имеет размерность $p \times p$. В ней, на пересечении i -ой строки и j -го столбца указывается количество пар пикселей со значениями яркости i и j , находящихся на расстоянии не большем, чем заданное наперед значение d , выраженное через изменение координат $(\Delta x, \Delta y)$ по соответствующим полуосям:

$$C_d(i, j) = \sum_x \sum_y \begin{cases} 1, \text{ если } I(x, y) = i \text{ и } I(x + \Delta x, y + \Delta y) = j \\ 0, \text{ в противном случае} \end{cases} \quad (2)$$

Пары пикселей могут иметь различную ориентацию и для этого они квантизуются по величине угла 0–45–90–135 градусов.

GLCM содержит ряд характерных параметров изображения: обратный разностный момент (inverse difference moment), энергия, инерция, энтропия и корреляция, которые могут быть использованы в качестве дескрипторов изображения.

Обратный разностный момент количественно определяет степень однородности в интересующей области. Параметр энергии иногда упоминается как второй угловой момент или однородность матрицы смежности. Наименьшее значение параметр энергии принимает, когда все функции плотности вероятности равны, и нет доминирующих уровней яркости. Параметр инерции (также называемый вторым разностным моментом) чувствителен к большим различиям, возникающим внутри матрицы смежности: высококонтрастные области имеют высокое значение параметра инерции, а более однородные области – низкое. Корреляция количественно определяет зависимость уровней яркости пар пикселей от расстояния между ними. Низкое значение параметра корреляции означает, что уровни серого вообще не зависят друг от друга, т. е. в изображении нет регулярной структуры. Если корреляция высока, существует высокая вероятность того, что одна или несколько текстур повторяются. Энтропия характеризует отсутствие пространственной организации на изображении. Энтропия высока для грубой (крупнозернистой) текстуры, и мала, когда текстура более однородная и гладкая.

GLCM может быть обобщена на случай анализа цветных изображений.

Также существуют методы, которые позволяют не просто выделять определенные характеристики части изображения, но и комбинировать их. Идея состоит в использовании небольших частей

изображения для составления словаря и последующего кодирования исходного изображения по наличию слов из словаря. К таким методам можно отнести Bag-of-words [69], сверточные нейронные сети [70–73]. Например, в методе Bag-of-words берется набор изображений, которые разбиваются на части, каждая из которых преобразуется в вектор признаков (слово). Далее, полученные вектора признаков разбиваются на заданное экспертом количество классов, формирующих словарь. После чего вычисляется гистограмма частоты появления каждого вектора на изображении, которая служит основой дальнейшей классификации изображения для определения принадлежности некоторому классу текстур.

На практике в качестве признаков наиболее часто используются статистические и структурные методы выделения текстур в виду их гибкости настройки, робастности и эффективности численной реализации. Ниже, в таблице 1 представлен сравнительный анализ различных методов выделения текстур.

Важным этапом анализа изображений является выделение значимых признаков. Данный этап обработки информации может оказаться критически важным в случае использования методов автоматической селекции информативных признаков из большого объема экспериментальных данных, поскольку такие данные, как правило, обладают высокой степенью зашумленности. Здесь под шумом также понимаются неинформативные данные. При этом неинформативный компонент вектора-признаков с большим числовым значением может маскировать информативные компоненты с меньшими значениями [74]. Другим примером служат данные, обладающие сильной корреляцией между собой, что приводит к невозможности применения некоторых методов классификации (например, линейного дискриминантного анализа), поскольку ковариационная матрица для таких данных не будет положительно определенной.

В соответствии с общепринятой классификацией методы выделения значимых признаков делятся на непрерывные и дискретные [75–77]. К непрерывным относятся методы, основанные на преобразовании пространства признаков, например, метод главных компонент (principal component analysis, PCA); факторный анализ, Isomap, диффузионные карты, многослойные автоэнкодеры и др. К дискретным методам относят так называемые фильтры, т.е. алгоритмы, основанные на выборе подмножества исходного набора признаков (критерий Пирсона, информация Шэннона, дивергенция Кульбака-Лейблера, метод BORUTA [78]). Ансамблевые алгоритмы на основе деревьев решений также позволяют произвести оценку значимости признаков.

Таблица 1. Сравнительная характеристика компьютерных методов выделения текстур

Методы	Структурные	Спектральные	Модельные	Статистические	Остальные
Преимущества	Простота настройки	Хорошо выделяют ориентированные структуры	Хорошо описывают шероховатость	Хорошо выделяют ориентированные структуры	Универсальные методы, позволяют находить любые структуры
Недостатки	Мало подходят для выделения ориентированных структур	Вычислительная сложность	Мало подходят для выделения ориентированных структур	Большое количество признаков на основе статистик, трудно выбрать оптимальные	Трудно обучить, требуется большая база изображений

МЕТОДЫ ПОСТРОЕНИЯ ПРЕДИКТИВНЫХ МОДЕЛЕЙ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ

Машинное обучение – набор методов из области искусственного интеллекта, который позволяет построить прогнозную модель анализируемого явления на основе накопленных данных. Прогнозные модели реализуются в виде классификаторов, позволяющих принять то или иное решение по отношению к новому объекту на основе его описания в виде некоторого набора величин (вектора признаков) (см., например, [79]).

Классификация заключается в разбиении множества объектов на классы по заданному критерию. Объекты одного класса считаются эквивалентными с точки зрения критерия разбиения.

Различают классификацию с учителем (supervised learning) и без учителя (unsupervised learning). Обучение с учителем подразумевает обучение классификатора, когда доступен некоторый набор векторов признаков (*обучающая выборка*, train dataset, TDS), для которых известна их принадлежность к одному из классов.

Наиболее широко используемыми и хорошо зарекомендовавшими себя методами обучения с учителем в задачах опознавания образов являются машина опорных векторов (Support Vector Machine, SVM), искусственные нейронные сети (Artificial Neural Networks, ANN) и др.

Классификация при помощи SVM происходит путем построения гиперплоскости в пространстве признаков, которая наиболее оптимально разделяет данные из разных классов. Когда граница между классами имеет сложный вид, применяют нелинейное интегральное

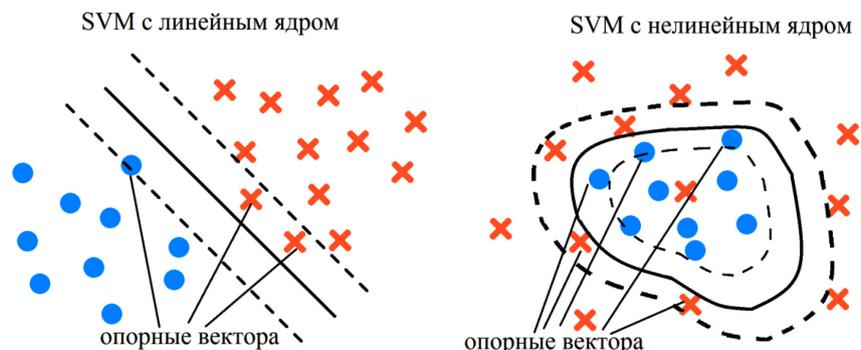


Рис. 3. Идея построения N -мерной гиперплоскости, оптимально разделяющей данные из двух разных классов: метод линейного SVM (а), вариант SVM с нелинейным ядром.

преобразование исходных данных с использованием различных ядер (kernel trick) [80] (рис. 3).

Искусственные нейронные сети, включая сверточные нейронные сети (Convolutional Neural Networks, CNN), представляет собой систему соединённых и взаимодействующих между собой искусственных нейронов. В нейроне входные значения суммируются и к сумме применяется нелинейная функция активации, например, сигмоида, гиперболический тангенс или усеченное линейное преобразование (Rectified Linear Unit, ReLU). Топология сети определяется связями между нейронами различных слоев (рис. 4), называемых весами. Целью обучения является подбор весов с целью минимизации ошибки между известным выходом нейронной сети и полученным значением на выходном нейроне. Для сверточных нейронных сетей входными данными являются изображения, веса заменяются на фильтры, с которыми производится свертка входных изображений, и параметры которых подбираются при обучении.

При обучении классификаторов существует опасность того, что он будет слишком хорошо подогнан под тренировочные данные, что приведет к существенным ошибкам при классификации новых данных. Эта проблема называется «перетренировкой» или «переобучением» классификатора. Принятие решения о качестве получившегося классификатора на основании проверки на тренировочных данных может привести к тому, что переобучение (если оно присутствует) не будет обнаружено.

Более адекватной является оценка качества классификации на *тестовом наборе* – множестве данных, размеченных по классам, но не

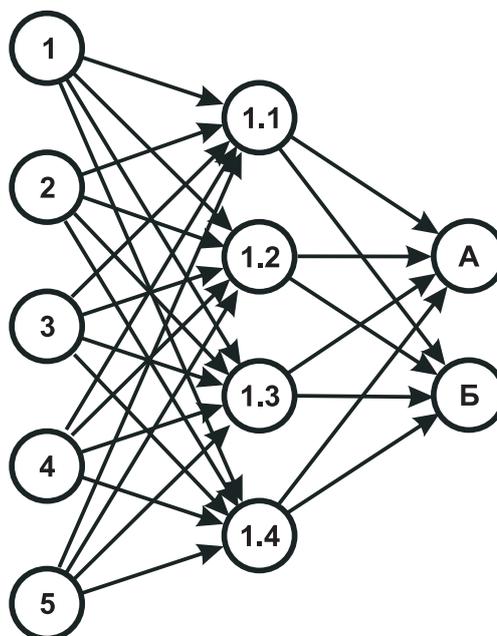


Рис. 4. Пример многослойного перцептрона; цифрами 1–5 обозначены входные нейроны, 1.1–1.4 – нейроны скрытого слоя, А и Б – выходные нейроны.

использовавшихся в процессе тренировки. Чаще всего при реализации данной идеи используются методы кросс-валидации или перекрестной проверки, основанные на манипуляциях с набором данных. Например, стратегия *Leave-one-out* подразумевает обучение на наборах, полученных путем последовательного удаления каждого элемента данных и последующее обучение классификатора на оставшемся множестве данных; при этом удаленный элемент используется при тестировании классификатора. Точность классификации в данном случае определяется средней точностью на каждом таком наборе. Существуют и другие модификации, подробнее о которых описано в работе [81], например, *Leave-several-out* (формирование тренировочных баз удалением нескольких элементов выборки), *Holdout* (разбиение случайным образом исходной выборки на тренировочную и тестовую).

Для избежания «переобучения» классификатора используются также методы, ограничивающие сложность модели [82], например, различные алгоритмы регуляризации. В большинстве случаев, регуляризация представляет собой дополнительное условие, реализующее штраф за сложность модели, например, большие значения коэффициентов в аппроксимирующих многочленах [83].

IV. ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ КОЛИЧЕСТВЕННОГО АНАЛИЗА ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ КОЛЛАГЕНА

Ниже представлены примеры выделения информативных признаков дезорганизации пространственной структуры коллагена на основе указанных методов и результатов дифференциации различных групп.

В работе [66] представлено решение задачи классификации SHG – изображений коллагена с использованием модельных, статистических методов и SVM. Признаки получают из описывающих характеристики пространственной статистики текстур изображения. Авторы учитывали ранее выявленную связь особенностей восприятия человеком различий изображений с их статистическими параметрами, описанными в разделе 2 данного обзора. Показано, что качество классификации изображений зависит от сложности статистики, в частности, перспективно использование GLCM, показывая результаты классификации, сравнимые с классификацией экспертом.

Узелки щитовидной железы обычно не имеют характерной симптоматики, но представляют собой клинические проявления широкого спектра различных патологий, в частности, ряда доброкачественных образований, таких как гиперпластические, коллоидные узелки или фолликулярные аденомы. Однако, около 10% узелков щитовидной железы связаны с развитием карциномы. Характерные изменения в пространственной дезорганизации коллагена потенциально могут служить ранним диагностическим маркером дифференциации подобных патологий. Для дифференциации пространственной структуры коллагена в капсулах вокруг фолликулярной аденомы щитовидной железы и узелках папиллярной карциномы был использован метод TRM в сочетании с текстурным анализом изображений [84].

SHG-изображения регистрировались с помощью конфокального сканирующего микроскопа Leica TCS-SP, для накачки использовался Ti:Sapphire лазер (Tsunami, Spectra Physics). Использовался диапазон сканирования 250×250 мкм² (512×512 пикселей), для уменьшения шума проводилось усреднение по 3 кадрам. При анализе изображений использовались три группы параметров: параметры оценки гистограммы яркости изображения (среднее значение, стандартное отклонение, эксцесс), параметры фрактального анализа (фрактальная размерность, лакунарность), параметры матрицы смежности (GLCM). Авторами показано, что из 18 рассмотренных 10 параметров позволяют различить доброкачественные и злокачественные образования. Это указывает на то, что существуют характерные специфические различия пространственной структуры коллагена для этих случаев.

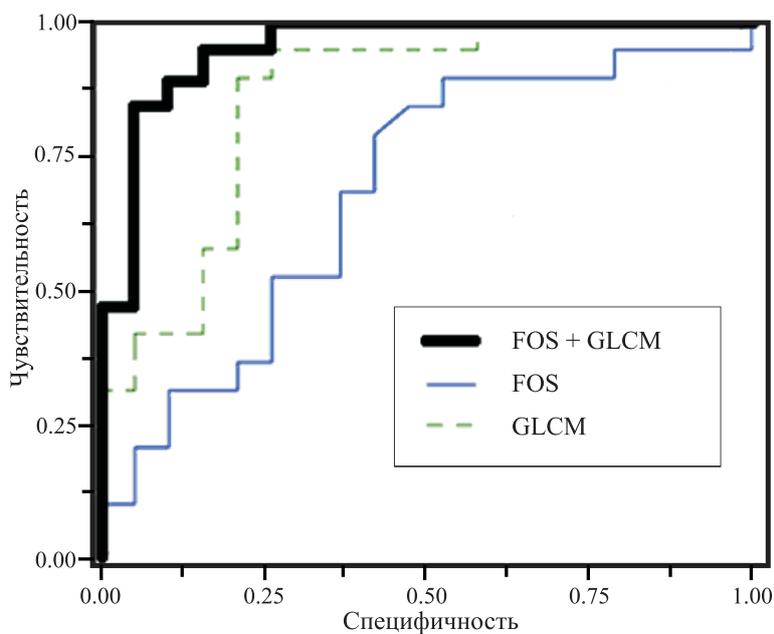


Рис. 5. ROC – кривая при использовании различных наборов признаков текстур коллагеновых волокон на SHG-изображениях ткани сердечной мышцы крыс с инфарктом миокарда, пролеченных с использованием стволовых клеток и без подобной терапии [8].

Работа [8] базируется на применении статистических методов и SVM. В ней представлен метод количественного описания изменений структуры коллагена на основе SHG – изображений. Для отработки метода были использованы ткани сердечной мышцы крыс с инфарктом миокарда и ткани атеросклеротических артерий кроликов.

В качестве признаков текстур коллагеновых волокон на SHG-изображениях ткани сердечной мышцы крыс с инфарктом миокарда, пролеченных с использованием стволовых клеток и без подобной терапии, использовались параметры FOS и GLCM. Классификация проводилась с использованием метода SVM. Для оценки качества классификации была построена кривая ошибок, называемая также ROC (receiver operating characteristic) – кривой. ROC – кривая графически характеризует зависимость доли верных положительных классификаций от доли ложных положительных классификаций при варьировании порога решающего правила. Количественную интерпретацию ROC – кривой даёт площадь под ней (area under

ROC curve, AUC) — площадь ограниченная ROC-кривой и осью доли ложных положительных классификаций. При использовании статистики первого порядка и матрицы смежности авторами [8] было получено значение $AUC = 0,95$ при дифференциации SHG-изображений инфарктной ткани. Точность классификатора (показатель AUC) снижалась до 0,86 и 0,72, если для дифференциации инфарктной ткани использовали только параметры GLCM или только параметры FOS, соответственно (см. рис. 5).

Данный подход был также использован для анализа текстур SHG-изображений коллагена в тканях атеросклеротической артерии. Анализировались образцы, полученные из просвета атеросклеротических артериальных сегментов, собранных у кроликов Watanabe heritable hyperlipidemic. Возраст кроликов составлял от 2 до 24 месяцев. Всего в работе было исследовано 414 SHG-изображений.

Исходя из характеристик структуры коллагеновых волокон, таких как форма, размер и структура, изображения были разделены экспертами на пять групп (А, В, С, D и E). Изображения группы А характеризовались наиболее тонкими, хорошо структурированными изогнутыми волокнами с определенным порядком выстраивания (ориентации). Изображения группы В имели более высокую плотность коллагена и более однородную ориентацию волокон по сравнению с группой А. Изображения групп С и D имели беспорядочно ориентированную сеть коротких волокон. Изображения группы E отличались длинными, прямыми и равномерно выстроенными (ориентированными) волокнами.

Был применен ранговый критерий Крускала-Валлиса (Kruskal-Wallis) для проверки статистически значимых различий между группами. Показано, что использование отдельного параметра FOS и GLCM не обеспечивает дифференцировку групп, за исключением группы А, которая имела параметры GLCM и некоторые параметры FOS текстур, существенно отличные от соответствующих характеристик других групп.

Авторами [8] также проведена многогрупповая классификация SHG-изображений коллагена атеросклеротических артериальных тканей с использованием дескрипторов параметров FOS и GLCM в качестве вектора признаков и нелинейного SVM. При попарном разделении групп разработанный классификатор обеспечил общую точность 90% при использовании всего набора параметров FOS и GLCM. Точность снижалась до 84% и 87% при использовании только параметров FOS или только параметров GLCM, соответственно (см. таблицу 2, [8]).

Таблица 2. Результаты SVM классификации для SHG-изображений коллагена атеросклеротических артерий на основе полного набора параметров FOS и GLCM [8]

Группа	Испытатель- ный набор	Точность, %	Чувствительность, %	Специфичность, %
A	6160	100	100	98
B	6520	73	79	100
C	6640	88	65	88
D	6360	90	86	94
E	4480	93	95	90
Итого	30160	90	91	92

Была также проведена оценка эффективности вышеизложенного подхода для классификации совокупности изображений при атеросклерозе и инфаркте миокарда. Авторы показали, что структуры коллагена артерии можно отделить от изображений структуры коллагена в тканях сердечной мышцы при инфаркте с точностью 99% при использовании всей совокупности параметров FOS или GLCM. Точность снижалась до 81% при использовании только параметров FOS, и до 96% – при использовании только параметров GLCM [8].

Таким образом, представленные результаты показывают, что SVM классификаторы обеспечивают достаточно высокую точность дифференцировки SHG-изображений структуры коллагена при различных заболеваниях с использованием вектора признаков, включающего полный набор параметров FOS и GLCM.

Авторами работы [85] были использованы структурные методы в совокупности со статистическими для описания текстур коллагена. В качестве классифицирующего критерия была использована оценка дисперсии 3D ориентации. Данный параметр количественно характеризует пространственную организацию волоконно-подобных элементов в некоторой области пространства.

При анализе изображений авторы [85] сначала выявляли на них структурные элементы, а затем рассчитывали статистику их ориентации в пространстве. Инвариантность ориентаций оценивалась по формуле

$$V_{3D} = 1 - \bar{R}_{3D}, \quad (3)$$

где

$$\bar{R}_{3D} = \sqrt{(\bar{C}_{3D}^c)^2 + (\bar{S}_{3D}^c)^2 + (\bar{Z}_{3D}^c)^2},$$

$$\bar{C}_{3D}^c = (1/n) \sum_{j=1}^n \sin \phi_j \cos \theta_j,$$

$$\bar{S}_{3D}^c = (1/n) \sum_{j=1}^n \sin \phi_j \sin \theta_j,$$

$$\bar{Z}_{3D}^c = (1/n) \sum_{j=1}^n \cos \phi_j.$$

Здесь n – общее число вокселей, которое используется для определения инвариантности ориентаций, θ , ϕ – азимутальный и полярный углы. Азимутальный угол связан с проекцией структур на плоскость XU . Для описания трехмерной ориентации авторы использовали проекцию структур на плоскость XZ путем введения дополнительных углов β , γ , аналогичных по смыслу углам θ , ϕ (рис. 6а). Связь между ними имеет вид:

$$\tan^2 \phi = 1 / \tan^2 \beta + 1 / \tan^2 \gamma.$$

На рис. 6б–6в показаны примеры структур с высокой и практически хаотической ориентацией волокон коллагена, полученные авторами обзора с помощью многофотонной микроскопии для здоровой ткани и ткани с лимфедемой. При этом более низкое значение дисперсии ориентации характеризует ярко выраженное пространственное выстраивание волокон, а более высокое значение соответствует случайной пространственной ориентации [85].

Данная характеристика была использована для описания пространственной организации коллагеновых волокон на основе SHG и TPF изображений. Изображения были получены с использованием конфокального микроскопа Leica TCS SP2. Анализ показал надежность восстановления значения дисперсии 3D ориентации исследуемых структур. Важно, что данный параметр не зависит от выбора системы координат. Предложенный подход может быть использован для анализа онкологических заболеваний, поскольку в опухоли наблюдается более выраженная ориентация коллагена [85].

Пространственная структура коллагеновых волокон, является ключевым фактором, определяющим упругие механические свойства ткани на макрокопическом уровне. Обычно, упругие свойства ткани исследуют с использованием деструктивных механических тестов.

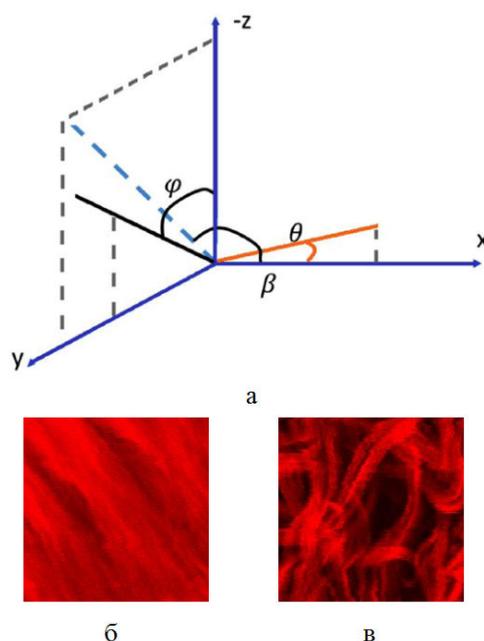


Рис. 6. К оценке дисперсии 3D ориентации; (а) – параметры, используемые при расчете дисперсии 3D ориентации [85], примеры структур с низкой (б) и высокой (в) величиной дисперсии 3D ориентации волокон коллагена, полученные авторами обзора с помощью многофотонной микроскопии для здоровой ткани и ткани с лимфедемой.

В работе [86] был предложен подход к анализу упругих свойств обработанной глутаральдегидом ткани бычьего перикарда на основе данных ТРМ с использованием структурных методов и сверточных нейронных сетей в виде так называемого глубинного обучения (Deep Learning). Это отрасль машинного обучения, использующая нейронные сети с большим числом обучаемых параметров.

Образцы ткани бычьего перикарда разрезали на образцы размером 20×20 мм, и четыре графитовых метки, ограничивающие квадрат размером приблизительно 2×2 мм², приклеивали к центральной области каждого квадрата для оптических измерений деформаций. Затем образцы размещали на двухосном анализаторе деформаций в кювете с водным 0,9%-ным раствором NaCl при 37 °C [87].

По завершении механических испытаний были записаны SHG – изображения образцов ткани в разгруженном состоянии. Для этого использовался инверсный конфокальный микроскоп Zeiss 710 NLO (Carl Zeiss Microscopy, LLC, Thornwood, NY, США). В процессе ска-

нирования образцы держали в водно-солевом растворе и покрывали стеклами для предотвращения высыхания. Сканировалась область образцов, ограниченная графитовыми метками, 2D-сканы изображения варьировались по глубине. Каждый 2D-скан содержал от 512×512 до 1024×1024 пикселей, для каждого образца количество сканов варьировалось таким образом, чтобы охватить всю толщину образца. В общей сложности были получены трехмерные SHG изображения 48 образцов ткани разных животных (размер от $512 \times 512 \times N$ до $1024 \times 1024 \times N$) и соответствующие данные механических испытаний. Здесь N – количество сканов.

Для анализа изображений авторы разработали сверточную нейронную сеть (convolutional neural network, CNN), состоящую из 5 скрытых слоев. CNN была спроектирована и обучена на базе данных SHG-изображений и результатов механических испытаний (диаграммы напряжений-деформаций) образцов для прогноза нелинейных анизотропных упругих свойств ткани на основе анализа SHG-изображений тканей без проведения механических испытаний. По оценкам авторов разработанная модель способна определять общую жесткость ткани с точностью 84%, и предсказывать форму анизотропной кривой напряжения-деформации со средней ошибкой регрессии от 0.021 до 0.031. Разработанный подход существенно сокращает время исследования упругих свойств ткани (до 10–30 минут) по сравнению с несколькими часами – с помощью традиционных механических испытаний.

В работе [37] авторами были использованы структурные методы для описания признаков. Регистрация процесса диффузии сигнала SHG в биоткани может быть реализована с использованием техники FLIM. Для регистрации временной зависимости SHG, в отличие от техники FLIM для регистрации автофлуоресценции в частотной области, необходим метод счета одиночных фотонов с корреляцией по времени (Time Correlated Single Photon Counting, TCSPC) [88]. Для реализации данного метода авторами данной работы были отобраны следующие образцы:

Сухожилия и кожа. У мужской особи крысы в возрасте шести недель было взято на исследование пяточное сухожилие.

Опухоль и сосуд сонной артерии. Семинедельных мышей использовали для воспроизведения опухолей путем подкожной инъекции клеток ЕМТ6 в правый фланг. Опухоли и сосуд сонной артерии удаляли, промывали 0,9% NaCl и погружали в фосфатный буферный солевой раствор при комнатной температуре.

Хрящ. Суставной хрящ был получен из сохранившихся областей бедренных мышечков пациентов (с их информированного согласия),

проходящих артропластику в отделении ортопедической хирургии (Университетская больница, Нанси, Франция). Хрящ отделяли от подкостной кости с помощью скальпеля. Цилиндрические эксплантаты (диаметром 5 мм) собирали с использованием стерильного биопсийного пуансона.

Бартонов студень (в пупочном канатике). Был получен из пуповины у доношенных новорожденных (с информированного согласия).

При помощи специальной сканирующей головки CLSM TCS SP2-AOBS (Leica Microsystems Mannheim, Германия), был реализован режим двухфотонной микроскопии на основе нелинейного оптического источника (MiraF900) и микроскопа MacroFluo™ Z6 APO A (Leica Microsystems Mannheim, Германия).

Времена затухания SHG – сигнала измерялись в направлении назад по отношению к лучу накачки с помощью системы подсчета единичных фотонов с временной корреляцией SPC-730 TCSPC Imaging module (Becker & Hickl, Berlin), соединенной со сканирующей головкой. Сигнал анализировался с помощью специализированной программы SPC Image (Becker & Hickl, Berlin). Для оценки распределения коллагена авторами был введен индекс, равный отношению времен жизни TPF к времени жизни SHG (TPF-SHG lifetime ratio). Результаты измерений данного индекса представлены на рис. 7.

Измерения показали, что время жизни SHG – сигнала от коллагена составляет 220 ± 100 пс, время жизни TPF > 500 пс. Данный подход, по мнению авторов, является высокоспецифичным по отношению к исследованию трехмерной структуры коллагена.

Подобный подход, основанный на оценке отношения содержания коллагена к эластину в ткани с использованием MPM, представлен в [89] и является разновидностью статистических методов выявления значимых признаков.

Оценка состояния пациентов с хроническим гепатитом С и повышенным содержанием трансаминаз часто проводится с использованием биопсии печени. В настоящее время, при анализе биоптата, как правило, используются перечисленные ниже полуколичественные методы диагностики стадии цирроза печени, основанные на оценке структурных изменений ткани [90].

Наиболее популярен индекс гистологической активности, предложенный Knodell и соавт. в 1981 г., который часто рассматривается в качестве эталона для полуколичественного (балльного) описания поражений печени при хроническом гепатите.

Индекс активности гепатита Шойера (Scheuer) был предложен в 1991 г. для балльной оценки хронических вирусных гепатитов, но, как и индекс Кнедела, применяется и для случаев невирусного

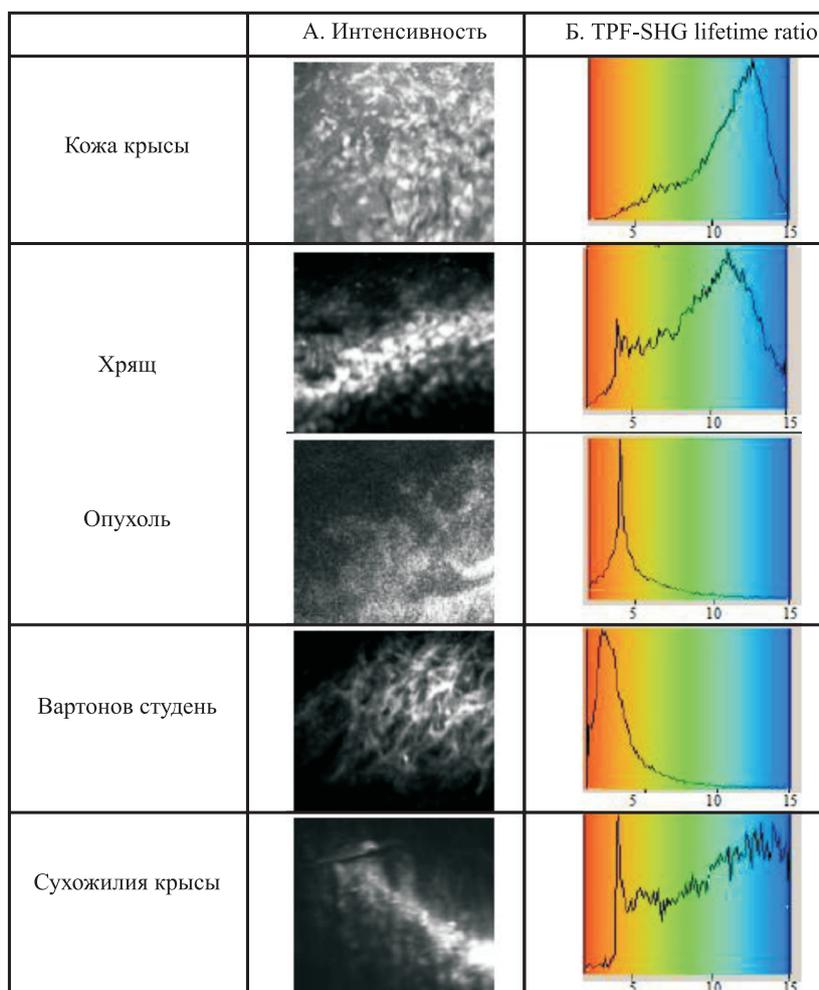


Рис.7. Измерение индекса TPF-SHG lifetime ratio для различных образцов биотканей (адаптировано с изменениями из работы [37])

гепатита. Основные преимущества этого метода по отношению к индексу Кнедела в том, что он более прост, реализует непрерывную шкалу, группирует перипортальные и портальные поражения в одну категорию.

Система подсчета баллов, введенная в 1994 году Ишаком (Ishak), способствовала использованию описательной терминологии для

оценки фиброза, классифицируя различные характеристики в логических переменных (присутствует/отсутствует).

Исследовательская группа Метавира (Metavir) в 1996 году предложила комплексную балльную систему гистологической оценки хронических заболеваний печени, вызванных гепатитом С. Предложенная шкала отражает комбинированную оценку фокального лобулярного некроза, портального воспаления, мелкоочагового и мостовидного некрозов.

Недостатком этих подходов является то, что стадия заболевания оценивается без количественного учета степени фиброза. Авторы работы [91] для количественной оценки степени цирроза печени предложили использовать параметр относительной площади коллагена (Collagen proportional area, CPA) на SHG – изображении. Методика оценки CPA состоит из трех этапов.

На первом этапе авторами использовались алгоритмы кластеризации для отделения фоновых пикселей от пикселей, принадлежащих ко всем другим областям изображения (ткани печени и другие типы тканей). По завершении этапа создается изображение с выделенными областями на нулевом фоне.

На втором этапе проводилось экспертное формирование векторов признаков для текстур, соответствующих ткани печени и других тканей (мышечная ткань, жировая ткань, кровеносный сосуд и т.п.). В результате, для каждой области изображения извлекается набор характеристик формы и цвета. Учитывались следующие параметры формы и цвета текстуры: площадь, эксцентриситет, диаметр, число Эйлера, отношение пикселей внутри области и на границе, длина большой и малой осей, длина периметра, доля пикселей в выпуклой оболочке текстуры, минимальная, максимальная и средняя интенсивность пикселей для каждого из каналов RGB в изображении. Таким образом, вектор признаков включал 18 характеристик (9 характеристик формы и 9 характеристик цвета).

Затем разрабатывался классификатор, позволяющий разделить ткань печени от других видов тканей. Параметр CPA рассчитывался по формуле

$$CPA = \frac{n_F}{n_F + n_H},$$

где n_F – число пикселей фиброзной ткани, n_H – число пикселей здоровой ткани печени.

Был проведен анализ 79 образцов биопсии печени, средняя ошибка оценки составила 1.31% с коэффициентом согласия корреляции, равным 0.923.

Авторы работы [92] использовали статистические методы и сверточные нейронные сети для дифференциальной диагностики доброкачественной онкоцитомы почек и злокачественной хромофобной почечной клеточной карциномы на основе анализа структуры коллагена на SHG – изображениях. На этапе предварительной обработки данных выделялись области с содержанием коллагена, сильно зашумленные изображения удалялись. Для оставшихся изображений проводилось выравнивание контраста, фильтрация шумов с использованием трех медианных фильтров с различным порогом обнуления яркости, затем – сглаживание с использованием гауссова фильтра. Для исключения зависимости результатов классификации от ориентации изображения, каждое изображение поворачивалось на угол 90° , 180° , 270° , а также выполнялась операция зеркального отражения.

Оценка качества классификации была проведена с разделением исходного набора на 10 подмножеств и дальнейшим использованием метода кросс-валидации путем исключения при обучении данных подмножеств по одному (Leave-one-out). В предыдущих исследованиях возможности дифференциальной диагностики этих же заболеваний печени с использованием SVM и валидации на основе метода Leave-one-out точность классификации составила 83.3%, т.е. в 83.3% случаев был получен правильный диагноз, чувствительность составила 75.0%, специфичность – 91.7%. Разработанная модель CNN обеспечила точность 68.7%, чувствительность 74.6% и специфичность 66.3% [91]. Для повышения качества диагностики авторы в дальнейшем планируют использовать образцы изображений большего размера и комбинировать SHG- изображения с изображениями TPF.

IV. ЗАКЛЮЧЕНИЕ

Изменение пространственной структуры коллагена является биомаркером большого числа заболеваний. Деструктивные изменения в волокнах коллагена характеризуются изменением длины, толщины волокон, и их ориентацией в пространстве.

Традиционно, анализ пространственной структуры коллагена проводится методами гистохимии, иммуногистохимии, гибридизации *in situ*, МРТ, рентгенографией.

В последнее время для исследования структуры биоткани все чаще применяют МРМ. Достоинствами МРМ визуализации является субклеточное пространственное разрешение и возможность использования совокупности модальностей: генерацию второй гармоники, автофлуоресценцию, включая флуоресцентный имиджинг с временным разрешением, что обеспечивает высокую информативность анализа.

С другой стороны высокое пространственное разрешение и комбинация модальностей приводит на практике к накоплению большого объема графических данных, эффективный содержательный анализ которых требует применения специальных компьютерных методов из области искусственного интеллекта.

Поскольку МРМ обеспечивает прямую регистрацию пространственного распределения коллагеновых волокон, первый шаг в применении компьютерных методов анализа связан с автоматическим выделением соответствующих структур на изображении. Стандартный подход к решению этой проблемы основан на понятии текстуры. Используемые на практике подходы выделения текстур отличаются способом их математического описания. Структурные методы опираются на представление текстуры в виде композиции визуальных примитивов, модельные методы описывают текстуру в виде ее математической модели, статистические методы используют свойства статистики распределения цветовых или серых оттенков пикселей в некоторой области изображения. Также применяются методы, основанные на выделении определенных характеристик части изображения для составления словаря и последующего кодирования исходного изображения по наличию слов из данного словаря. К таким методам можно отнести Bag-of-words, сверточные нейронные сети.

Ключевым этапом анализа изображений является выделение значимых признаков. Проблема заключается в том, что наличие большого количества неинформативных признаков может маскировать информативную часть накопленных данных. Это существенно ухудшает качество классификации. Для выделения информативных признаков используются непрерывные и дискретные методы. Преимущество непрерывных методов заключается в автоматическом преобразовании пространства признаков и выделении значимых признаков в этом пространстве на основе минимизации дисперсии данных. Также они позволяют проводить визуальный анализ проекций исходных объектов на различные гиперплоскости в пространстве информативных признаков. Недостатки состоят в том, что требуется дополнительный анализ эксперта, который позволяет подобрать комбинацию признаков, оптимальную с точки зрения.

Дискретные методы позволяют обрабатывать сильно зашумленные данные, учитывать взаимную корреляцию признаков, производить подбор различных комбинаций признаков, но требуют существенно больше вычислительных ресурсов.

Наличие значимых признаков позволяет построить прогнозную модель, позволяющую принять то или иное решение по отношению к новому объекту на основе его вектора признаков (провести классификацию). Классификация заключается в разбиении множества объектов

на классы по заданному критерию. Наиболее широко используемыми и хорошо зарекомендовавшими себя методами распознавания визуальных образов являются машина опорных векторов (Support Vector Machine), искусственные нейронные сети (Artificial Neural Networks).

При разработке классификаторов необходимо избегать его «переобучения». Для решения данной задачи могут быть использованы методы кросс-валидации, основанные на манипуляциях с набором данных. Примерами данного подхода являются стратегии Leave-one-out, Leave-several-out, Holdout. Также используются различные техники регуляризации моделей, чтобы не допускать их чрезмерного усложнения.

В настоящем обзоре представлен ряд работ, в которых классификация МРМ изображений структуры коллагена основана на использовании набора статистик первого и второго порядков в качестве вектора признаков, причем использование совокупности указанных характеристик позволяет достичь достаточно высокой точности классификации. Перспективно использование дисперсии 3D ориентации, а также относительной площади коллагена на изображении в качестве информативных характеристик пространственной дезорганизации структуры коллагена.

Представленные примеры показывают, что для создания предсказательных моделей в рамках рассматриваемой задачи также наиболее эффективен метод опорных векторов. Данный вывод также подтверждается нашими результатами при анализе спектров поглощения проб выдыхаемого воздуха для различных бронхолегочных заболеваний методами лазерной спектроскопии [93, 94]. Также перспективно использование сверточных нейронных сетей, однако эти методы требуют большей обучающей выборки по сравнению с методом опорных векторов.

В целом, как показывает представленный анализ, сочетание прямых методов визуализации дезорганизации структуры коллагена на субклеточном уровне (каковым является МРМ) с методами машинного обучения является эффективным подходом для разработки компьютерных методов дифференциальной диагностики. МРМ реализует набор оптических модальностей для регистрации большого объема графических данных высокого разрешения, а методы машинного обучения позволяют выделить высоко специфическую часть информации из накопленных данных и затем построить эффективные прогностические модели, которые обеспечат высокоточную классификацию новых данных.

Анализ литературы в области применения методов машинного обучения для количественной оценки дезорганизации коллагена в ткани показывает, что данное направление только начинает развиваться.

ЛИТЕРАТУРА

1. Andriotis, O.G., Chang, S.W., Vanleene, M., Howarth, P.H., Davies, D.E., Shefelbine, S.J., Buehler, M.J., Thurner, P.J. (2015) Structure–mechanics relationships of collagen fibrils in the osteogenesis imperfecta mouse model, *J. R. Soc. – Interface*, **12**, 20150701. <http://dx.doi.org/10.1098/rsif.2015.0701>.
2. Капулер О., Сельская Б., Галеева А., Камилов Ф. (2015), Метаболизм коллагеновых волокон на фоне возрастных изменений, *Врач*, № 8, 64–69.
3. Lin, De, Chun, T.-H., Kang, Li (2016) Adipose extracellular matrix remodelling in obesity and insulin resistance, *Biochem. Pharmacol.*, **119**, 8–16; doi: 10.1016/j.bcp.2016.05.005
4. <http://www.medsest.ru/info/physiology/kak-rabotayut-belki-denaturaciya-belka-urovni-organizacii-molekuly>.
5. http://www.dcp-recovery.ru/p/blog-page_2.html.
6. Симоненко Г.В., Тучин В.В. (2007) Основные сведения о строении различных типов биотканей. Дополнительные сведения из анатомии человека. Саратов: Саратовский государственный университет, 13 с.
7. Gelse, K., Pöschl, E., Aigner, T. (2003) Collagens – structure, function, and biosynthesis, *Adv. Drug Deliv. Rev.*, **55**, 1531–1546.
8. Mostaço-Guidolin, L.B., Ko, A.C.-T., Wang, F., Xiang, B., Hewko, M., Tian, G., Major, A., Shiomi, M., Sowa, M.G. (2013) Collagen morphology and texture analysis: from statistics to classification, *Sci. Rep.*, **3**, 2190.
9. Aziz, J., Shezali, Ha, Radzi, Z., Yahya, N.A., Abu Kassim, N.H., Czernuszka, J., Rahman M.T. (2016) Molecular Mechanisms of Stress-Responsive Changes in Collagen and Elastin Networks in Skin, *Skin Pharmacol. Physiol.*, **29**, 190–203; <https://doi.org/10.1159/000447017>.
10. Said, G., Guilbert, M., Millerot-Serrurot, E., Van Gulick, L., Terryn, Ch., Garnotel, R., Jeannesson, P. (2012) Impact of carbamylation and glycation of collagen type I on migration of HT1080 human fibrosarcoma cells, *Int. J. Oncol.*, **40**, 1797–1804.
11. Avery, N.C., Bailey, A.J. (2006) The effects of the Maillard reaction on the physical properties and cell interactions of collagen, *Pathol. Biol. (Paris)*, **54**, 387–395.
12. Yuen, A., Laschinger, C., Talior, I., Leev W., Chan, M., Birek, J., Young, E.W., Sivagurunathan, K., Won, E., Simmons, C.A., McCulloch C.A. (2010) Methylglyoxal-modified collagen promotes myofibroblast differentiation, *Matrix Biol.*, **29**, 537–548.
13. Kemeny, S.F., Figueroa, D.S., Andrews, A.M., Barbee, K.A., Clyne, A.M. (2011) Glycated collagen alters endothelial cell actin alignment and nitric oxide release in response to fluid shear stress, *J. Biomech.*, **44**, 1927–1935.
14. Titov, V.N., Khokhlova, N.V., Shiriaeva, I.K. (2013) Glucose, glyco-toxins, and protein glycation products: the role in pathogenesis, *Klin. Med. (Mosk.)*, **91**, 15–24.
15. Tu, Y., Quan, T. (2016) Oxidative Stress and Human Skin Connective Tissue Aging, *Cosmetics*, **3**, 1–12.
16. Jabłońska-Trypuć, A., Matejczyk, M., Rosochacki, S. (2016) Matrix metalloproteinases (MMPS), the main extracellular matrix (ECM) enzymes in collagen degradation, as a target for anticancer drugs, *Journal of Enzyme Inhibition and Medicinal Chemistry*, **31**, 177–183; doi: 10.3109/14756366.2016.1161620.
17. Siwik, D.A., Pagano, P.J., Colucci, W.S. (2001) Oxidative stress regulates collagen synthesis and matrix

- metalloproteinase activity in cardiac fibroblasts, *Am. J. Physiol. Cell Physiol.*, **280**, 53–60.
18. Rutkowski, J.M., Moya, M., Johannes, J., Goldman, J., Swartz, M.A. (2006) Secondary lymphedema in the mouse tail: Lymphatic hyperplasia, VEGF-C upregulation, and the protective role of MMP-9, *Microvasc. Res.*, **72**, 161–171; doi: 10.1016/j.mvr.2006.05.009.
 19. Николаев В.В., Курочкина А.С., Князькова А.И., Вражнов Д.А., Кистенев Ю.В. (2018) Изучение заболевания лимфедема при помощи многофотонной микроскопии высокого разрешения и технологии FLIM, *Перспективы развития фундаментальных наук*, **4**, Биология и фундаментальная медицина, 113–115.
 20. Fang, M., Yuan, J., Peng, C., Li, Y. (2014) Collagen as a double-edged sword in tumor progression, *Tumour Biol.*, **35**, 2871–2882; doi: 10.1007/s13277-013-1511-7.
 21. Polasek, M., Yang, Y., Schuhle, D.T., Yaseen, M.A., Kim, Y.R., Sung, Y.S., Guimaraes A.R., Caravan P. (2017) Molecular MR imaging of fibrosis in a mouse model of pancreatic cancer, *Scientific Reports*, **7**, 1–10.
 22. Zhang, L., Albon, J., Jones, H., Gouget, C.L.M., Ethier, C.R., Goh, J.C.H., Girard, M.J.A. (2015) Collagen Microstructural Factors Influencing Optic Nerve Head Biomechanics, *Invest. Ophthalmol. Vis. Sci.*, **56**, 2031–2042; doi:10.1167/iavs.14-15734.
 23. Pigeon, H., Zucchi, H., Rousset, F., Monnier, V.M., Asselineau, D. (2014) Skin aging by glycation: lessons from the reconstructed skin model, *Clin. Chem. Lab. Med.*, **52**, 169–174.
 24. Gogola, A., Jan, N.-J., Brazile, B., Lam, P., Lathrop, K.L., Chan, K.C., Sigal, I.A. (2018) Spatial patterns and age-related changes of the collagen crimp in the human cornea and sclera, *Invest. Ophthalmol. Vis. Sci.*, **59**, 2987–2998. <https://doi.org/10.1167/iavs.17-23474>.
 25. Navnita, S., Savita, S., Rithesh, K., Shivaprasad, B. (2017) Evaluation of collagen degradation in gingival epithelium and connective tissue of chronic periodontitis patients with and without diabetes mellitus International, *Journal of Scientific Research*, **6**, 33–35.
 26. Tuchin, V.V. (2015) Tissue Optics and Photonics: Light-Tissue Interaction [Review], *Journal of Biomedical Photonics & Eng.*, **1**, 98–135.
 27. Topping, G., Malda, J., Dawson, R., Upton, Z. (2006) Development and characterization of human skin equivalents and their potential application as a burn wound model, *Primary Intention*, **14**, 14–21.
 28. Oostendorp, C., Uijtdewilligen, P.J.E., Versteeg, E.M., Hafmans, Th.G., van den Bogaard, E.H., de Jonge, P.K.J.D., Pirayesh, A., von den Hoff, J.W., Reichmann, E., Daamen, W.F., van Kuppevelt, T.H. (2016) Visualisation of newly synthesized collagen *in vitro* and *in vivo*, *Scientific Reports*, **6**, 1–7; doi: 10.1038/srep18780.
 29. Ross, K.A., Williams, R.M., Schnabel, L.V., Mohammed, H.O., Potter, H.G., Bradica, G., Castiglione, E., Pownder, S.L., Satchell, P.W., Saska, R.A., Fortier, L.A. (2013) Comparison of Three Methods to Quantify Repair Cartilage Collagen Orientation, *Cartilage*, **4**, 111–120.
 30. Jan, N.-J., Lathrop, K., Sigal, I.A. (2017) Collagen architecture of the posterior pole: high-resolution wide field of view visualization and analysis using polarized light microscopy, *Invest. Ophthalmol. Vis. Sci.*, **58**, 735–744; doi: 10.1167/iavs.16-20772.
 31. Fullerton, G.D., Rahal, A. (2007) Collagen Structure: The Molecular Source of the Tendon Magic Angle Effect, *Journal of magnetic resonance imaging*, **25**, 345–361.

32. Ghazanfari, S., Driessen-Mol, A., Strijkers, G.J., Baaijens, F.P.T., Bouten, C.V.C. (2015) The Evolution of Collagen Fiber Orientation in Engineered Cardiovascular Tissues Visualized by Diffusion Tensor Imaging, *PLoS ONE*, **10**, 1–15; doi: 10.1371/journal.pone.0127847.
33. Johnson, G.A., Liu, Ch., Wei, H., Gibbs, E., Zhao, P., Wang, N., Cofer, G.P., Zhang, Y. (2017) Susceptibility Tensor Imaging and Tractography of Collagen Fibrils in the Articular Cartilage, *Magn. Reson. Med.*, **78**, 1683–1690; doi: 10.1002/mrm.26882.
34. Bashkatov, A.N., Genina, E.A., Kozintseva, M.D., Kochubei, V.I., Gorodkov, S.Yu., Tuchin, V.V. (2016) Optical Properties of Peritoneal Biological Tissues in the Spectral Range of 350–2500 nm, *Opt. and Spectrosc.*, **120**, 1–8. <https://doi.org/10.1134/S0030400X16010045>.
35. Yang, M.F., Tuchin, V.V., Yaroslavsky, A.N. (2009) Principles of Light-Skin Interactions. In: Baron E. (eds) *Light-Based Therapies for Skin of Color*. London: Springer, p. 1–44.
36. Roth, S., Freund, I. (1979) Second harmonic generation in collagen, *The Journal of Chemical Physics*, **70**, 1637–1643; doi: 10.1063/1.437677
37. Dumas, D., Werkmeister, E., Hupont, S., Huselstein, C., De Isla, N., Rousseau, M., Menu, P., Mainard, D. (2014) Non-Invasive Second Harmonic Generation (SHG) in Macroscopy (MacroSHG) as Bio-Diagnosis to Image Collagen Network Organization in Extracellular Matrix, *Engineering*, **6**, 485–490.
38. Tuchin, V.V. (2006) *Optical clearing of tissues and blood*. Bellingham: SPIE Press, 256 p.
39. Tseng, J.-Y., Ghazaryan, A.A., Lo, W., Chen, Y.-F., Hovhannisyan, V., Chen, Sh.-J., Tan, H.-Y., Dong, Ch.-Y. (2011) Multiphoton spectral microscopy for imaging and quantification of tissue glycation, *Biomed. Opt. Express*, **2**, 218–230.
40. Huang, S., Heikal, A.A., Webb, W.W. (2002) Two-photon fluorescence spectroscopy and microscopy of NAD(P)H and flavoprotein, *Biophys. J.*, **82**, 2811–2825.
41. Sdobnov, A., Darvin, M.E., Lademann, J., Tuchin, V. (2017) A comparative study of ex vivo skin optical clearing using two-photon microscopy, *J. Biophotonics*, **10**, 9, 1115–1123.
42. Park, C.Y., Marando, C.M., Liao, J.A., Lee, J.K., Kwon, J., Chuck, R.S. (2016) Details of the collagen and elastin architecture in the human limbal conjunctiva, Tenon’s capsule and sclera revealed by two-photon excited fluorescence microscopy, *Invest. Ophthalmol. Vis. Sci.*, **57**, 5602–5610; doi: 10.1167/iovs.16-19706.
43. Ambekar, R., Lau, T.-Y., Walsh, M., Bhargava, R., Toussaint, K.C.Jr. (2012) Quantifying collagen structure in breast biopsies using second-harmonic generation imaging, *Biomed. Opt. Express*, **3**, 2021–2035.
44. Golaraei, A., Kontenis, L., Cisek, R., Tokarz, D., Done, S.J., Wilson, B.C., Barzda, V. (2016) Changes of collagen ultrastructure in breast cancer tissue determined by second-harmonic generation double Stokes-Mueller polarimetric Microscopy, *Biomed. Opt. Express*, **7**, 4054–4068.
45. Bogdanov, A.A., Mazzanti, M.L. (2013) Fluorescent Macromolecular Sensors of Enzymatic Activity for In Vivo Imaging, *Progress in Molecular Biology and Translational Science*, **113**, 349–387.
46. Vishwanath, V., Frank, K.E., Elmet, C.A., Dauchot, P. J., Monnier, V.M. (1986) Glycation of Skin Collagen in Type I Diabetes Mellitus Correlation With Long-Term Complications, *Diabetes*, **35**, 916–921.
47. Depeursinge, A., Fageot, J., Al-Kadi, O.S. (2018) *Fundamentals of Texture*

- Processing for Biomedical Image Analysis: A General Definition and Problem Formulation. Biomedical Texture Analysis, 1–27.
48. Castellano, G., Bonihla, L., Li, L.M., Cendes, F. (2004) Texture analysis of medical images, *Clin. radiol.*, **59**, 1061–1069.
 49. Lowe, D.G. (1999) Object recognition from local scale-invariant features, *The proceedings of the seventh IEEE international conference on Computer vision*, **2**, 1150–1157.
 50. Dalal, N., Triggs, B. (2005) Histograms of oriented gradients for human detection, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, **1**, 886–893.
 51. Donoho, D.L., Huo, X. (1999) Combined image representation using edgelets and wavelets, *Wavelet applications in signal and image processing VII*, **3813**, 468–477.
 52. Huang, J., Kumar, R., Mitra, M., Zhu, W.-J., Zabih, R. (1997) Image indexing using color correlograms, *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 762–768.
 53. Guo, Z., Zhang, L., Zhang, D. (2010) A completed modeling of local binary pattern operator for texture classification, *IEEE Transactions on Image Processing*, **19**, 1657–1663.
 54. Simoncelli, E.P., Freeman, W.T. (1995) The steerable pyramid: A flexible architecture for multi-scale derivative computation, *Image Processing, Proceedings International Conference on IEEE*, **3**, 444–447.
 55. Haralick, R.M., Sternberg, S.R., Zhuang, X. (1987) Image analysis using mathematical morphology, *IEEE Transactions on pattern analysis and machine intelligence*, **4**, 532–550.
 56. Materka, A., Strzelecki, M. (1998) Texture analysis methods – a review, *Technical university of Lodz, Institute of Electronics, COST B11 report*, 9–11.
 57. Fletcher, N.D., Evans, A.N. (2005) Texture segmentation using area morphology local granulometries, *Mathematical Morphology: 40 Years On.* – Springer, 367–376.
 58. Bigün, J., Granlund, G.H., Wiklund, J. (1991) Multidimensional orientation estimation with applications to texture analysis and optical flow, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **8**, 775–790.
 59. Kemaq, Q. (2007) Two-dimensional windowed Fourier transform for fringe pattern analysis: principles, applications and implementations, *Optics and Lasers in Engineering*, **45**, 304–317.
 60. Arivazhagan, S., Ganesan, L. (2003) Texture classification using wavelet transform, *Pattern recognition letters*, **24**, 1513–1521.
 61. Chang, T., Kuo, C.C.J. (1993) Texture analysis and classification with tree-structured wavelet transform, *IEEE Transactions on image processing*, **2**, 429–441.
 62. Kistenev, Yu.V., Shapovalov, A.V., Borisov, A.V., Vrazhnov, D.A., Nikolaev, V.V., Nikiforova, O.Yu. (2015) Wavelet based de-noising of breath air absorption spectra profiles for improved classification by principal component analysis, *AIP Conference Proceedings* **1688**, 030010.
 63. Khobragade, P., Fan, J., Rupcich, F., Crotty, D.J., Schmidt, T.G. (2018) Application of fractal dimension for quantifying noise texture in computed tomography images, *Med. Phys.*, **45**, 3563–3573.
 64. Depeursinge, A., Al-Kadi, O.S., Mitchell, J.R. (2017) Biomedical Texture Analysis: Fundamentals, Tools and Challenges, Academic Press, 430 p.

65. Sankar, D., Thomas, T. (2010) Fractal features based on differential box counting method for the categorization of digital mammograms, *Journal of Computer Information Systems and Industrial Management Applications*, **2**, 11–19.
66. Bevk, M., Kononenko, I. (2002) A statistical approach to texture description of medical images: a preliminary study, *Proceedings of the 15th IEEE Symposium on Computer-Based Medical Systems*, **64**, 239–244.
67. Haralick, R.M., Shanmugam, K., Dinstein, I. (1973) Textural features for image classification, *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-3**, 610–621.
68. Nanni, L., Brahnam, Sh., Ghidoni, S., Menegatti, E., Barrier, T. (2013) Different approaches for extracting information from the co-occurrence matrix, *PloS one*, **8**, e83554.
69. Farinella, G.M., Moltisanti, M., Battiato, S. (2014) Classifying food images represented as bag of textons, *2014 IEEE International Conference on Image Processing (ICIP)*, 5212–5216.
70. Sahiner, B., Chan, H.P., Petrick, N., Wei, D., Helvie, M.A., Adler, D.D., Goodsitt, M.M. (1996) Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images, *IEEE Trans Med Imaging*, **15**, 598–610.
71. Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., Thrun, S. (2017) Dermatologist-level classification of skin cancer with deep neural networks, *Nature*, **542**, 115–118.
72. Tivive, F.H.C., Bouzerdoum, A. (2006) Texture classification using convolutional neural networks, *2006 IEEE Region 10 Conference (TENCON 2006)*, 1–4.
73. Anthimopoulos, M., Christodoulidis, S., Ebner, L., Mougiakakou, S. (2016) Lung pattern classification for interstitial lung diseases using a deep convolutional neural network, *IEEE Trans Med Imaging*, **35**, 1207–1216.
74. Kistenev, Yu.V., Borisov, A.V., Shapovalov, A.V., Nikiforova, O.Yu. (2015) Analysis of the component composition of exhaled air using laser spectroscopy and canonical correlation analysis, *Proc. SPIE 9680, 21st International Symposium Atmospheric and Ocean Optics: Atmospheric Physics*, **9680**, 96804C; doi: 10.1117/12.2205786.
75. Postma, E., Herik, H.J. Maaten, L.J.P. (2009) Dimensionality Reduction: A Comparative Review, *J. Mach. Learn. Res.*, 1–22.
76. Nunes, M.A., Prangle, D., Sisson, S.A. Blum, M.G.B. (2013) A Comparative Review of Dimension Reduction Methods in Approximate Bayesian Computation, *Statistical Science*. **28**, 189–208.
77. Angryk, R., Martens, P. Banda, J.M. (2012) Quantitative Comparison of Linear and Non-linear Dimensionality Reduction Techniques for Solar Image Archives, *Proceedings of the 25th International FLAIRS Conference (FLAIRS'12)*, 376–381.
78. Kursu, M.B., Rudnicki, W.R. (2010) Feature selection with the Boruta package, *J. Stat. Softw.*, **36**, 1–13.
79. Piatetsky-Shapiro, G., Matheus, C., Frawley, W. (1992) Knowledge discovery in databases: An overview, *AI Magazine*, **13**, 57–70.
80. Vrazhnov, D.A., Nikolaev, V.V., Shapovalov, A.V., Sandykova, E.A. (2018) The kernel construction for the biomedical data classification using support vector machine, *Proc. SPIE 10614, International Conference on Atomic and Molecular Pulsed Lasers XIII*, **10614**, 106141Y.
81. Arlot, S., Celisse, A. (2010) A survey of cross-validation procedures for model selection, *Statist. Surv.*, **4**, 40–79.

82. Hawkins, D.M. (2004) The problem of overfitting, *J. Chem. Inf. Comp. Sci.*, **44**, 1–12.
83. Boyd, S., Vandenberghe, L. (2004) Convex optimization. New York: Cambridge university press. 730 p.
84. Hristu, R., Eftimie, L.G., Stanciu, S.G., Tranca, D.E., Paun, B., Sajin, M., Stanciu, G.A. (2018) Quantitative second harmonic generation microscopy for the structural characterization of capsular collagen in thyroid neoplasms, *Biomed. Opt. Express*, **9**, 3923–3936.
85. Liu, Z., Pouli, D., Sood, D., Sundarakrishnan, A., Hui Mingalane, C.K., Arendt, L.M., Alonzo, C., Quinn, K.P., Kuperwasser, C., Zeng, L., Schnelldorfer, T., Kaplan, D.L., Georgakoudi, I. (2017) Automated quantification of three-dimensional organization of fiberlike structures in biological tissues, *Biomaterials*, **116**, 34–47.
86. Liang, L., Liu, M., Sun, W. (2017) A Deep Learning Approach to Estimate Chemically-Treated Collagenous Tissue Nonlinear Anisotropic Stress-Strain Responses from Microscopy Images, *Acta Biomaterialia*, **63**, 227–235.
87. Sacks, M.S. (2000) Biaxial Mechanical Evaluation of Planar Biological Materials, *Journal of Elasticity*, **61**, 199–246.
88. Беккер В., Желзов А., Чеславский В. (2015) Визуализация времени жизни флуоресценции с помощью многомерного TCSPC-метода: новые возможности в биомедицине, *Биофотоника*, **5/53**, 52–66.
89. Kistenev, Y.V., Nikolaev, V.V., Drozdova, A.K., Piyasova, E.E., Sandukova, E.A. (2018) Improvement of the multiphoton fluorescence microscopy images quality using digital filtration, *Proc. SPIE 10614, International Conference on Atomic and Molecular Pulsed Lasers XIII*, **10614**, 106141Q.
90. Mannan, R., Misra, V., Misra, S.P., Singh, P. A., Dwivedi, M. (2014) A Comparative Evaluation of Scoring Systems for Assessing Necro-Inflammatory Activity and Fibrosis in Liver Biopsies of Patients with Chronic Viral Hepatitis, *J. Clin. Diagn. Res.*, **8(8)**, FC08-FC12.
91. Tsipouras, M.G., Giannakeas, N., Tzallas, A.T., Tsianou, Z.E., Manousou, P., Hall, A., Tsoulos, I., Tsianos, E. (2017) A methodology for automated CPA extraction using liver biopsy image analysis and machine learning techniques, *Comput. Methods Programs Biomed.*, **140**, 61–68.
92. Judd, N., Smith, J., Jain, M., Mukherjee, M., Icaza, M., Gallagher, R., Szeligowski, R., Wu, B. (2018) A pilot study for distinguishing chromophobe renal cell carcinoma and oncocytoma using second harmonic generation imaging and convolutional neural network analysis of collagen fibrillar structure, *Proc. SPIE 10489, Optical Biopsy XVI: Toward Real-Time Spectroscopic Imaging and Diagnosis*, **10489**, 1048919.
93. Kistenev, Yu.V., Kuzmin, D.A., Vrazhnov, D.A., Borisov, A.V. (2016) Classification of patients with broncho-pulmonary diseases based on analysis of absorption spectra of exhaled air samples with SVM and neural network algorithm application, *Proc. SPIE 10035, 22nd International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics*, 1003507.
94. Kistenev, Yu.V., Borisov, A.V., Kuzmin, D.A., Penkova, O.V., Kostyukova, N.Y., Karapuzikov, A.A. (2017) Exhaled air analysis using wideband wave number tuning range infrared laser photoacoustic spectroscopy, *J. Biomed. Opt.*, **22**, 017002.