

Применение глубоких нейронных сетей для повышения точности диагностики ревматоидного артрита с использованием диффузионной оптической томографии*

Янцин Фэн, Д.Лайтер, Лэй Чжан, Ян Ван, Х.Дехани

Предложен набор моделей глубоких нейронных сетей для классификации ревматоидного артрита (РА) с использованием магистральной сети, сверточной нейронной сети и остаточной сети на основе данных диффузионной оптической томографии (ДОТ) с использованием излучения ближнего ИК диапазона, обеспечивающий раннюю диагностику патофизиологических изменений, вызванных воспалением. Для преодоления проблемы недостаточного количества доступных клинических ДОТ-изображений использована генерация искусственных изображений при помощи числовой модели пальца. Предложенные модели глубоких нейронных сетей были применены для автоматической классификации смоделированных ДОТ-изображений воспаленных и невоспаленных суставов; для повышения эффективности классификации использовано также трансферное обучение. Показано, что все три метода использования глубоких нейронных сетей улучшают диагностическую точность по сравнению с широко применяемым методом опорных векторов, особенно для баз данных с высокой вариабельностью между субъектами. В случаях моделирования тяжелого заболевания остаточная сеть достигла наивысшей точности (свыше 99%), а магистральные и сверточные нейронные сети обеспечивали точность до 99%. Однако, по мере уменьшения тяжести моделируемой болезни, указанная точность для остаточных сетей снижалась до 75.2%. Результаты показывают, что трансферное обучение может улучшить производительность глубоких нейронных сетей при классификации РА по данным ДОТ и усилить их потенциальные возможности в качестве компьютерного инструмента в диагностических системах ДОТ.

Ключевые слова: диффузионная оптическая томография, диагностика ревматоидного артрита, суставы пальцев, глубокие нейронные сети, классификация медицинских изображений.

1. Введение

Ревматоидный артрит (РА) – хроническое заболевание, связанное со значительной болью и приводящее к инвалидности [1], которое является самым распространенным типом воспалительного артрита и поражает от 0.5% до 1% населения мира [2]. Наиболее часто поражаются РА запястья, плюснефаланговые и проксимальные межфаланговые (PIP) суставы [1]. В течение трех-четырех месяцев после появления первых симптомов существует терапевтическое окно [3], за время которого агрессивная терапия приводит к улучшению отдаленных результатов у пациента [4]. Развитие технологий медицинской визуализации в последние десятилетия позволило более точно определять воспаление у пациентов с РА на ранних стадиях прогрессирования заболевания, например с помощью ультразвука [5, 6] или магнитно-резонансной томографии (МРТ) [7, 8], однако необходимость в опытных персонале приводит к высокой стоимости и ограничен-

ной доступности этих методов. Следовательно, существует необходимость в недорогих, неинвазивных и объективных методиках оценки, которые могут осуществляться неклиническим персоналом, что делает диффузионную оптическую томографию (ДОТ) привлекательной для диагностики и долгосрочного мониторинга пациентов с РА.

ДОТ – метод визуализации, в котором излучение ближнего ИК диапазона вводится в биологическую ткань, а затем рассеянные фотоны детектируются в нескольких местах на ее границе, что обеспечивает восстановление базового распределения оптических свойств ткани [9]. В последнее время системы ДОТ активно применяются во многих областях, включая обнаружение рака молочной железы, функциональную визуализацию головного мозга и диагностику артрита суставов [10–12]. А.Х.Хильшер и др. представили одноволновую спектральную ДОТ-систему, способную восстанавливать карты поглощения и рассеяния в суставах пальцев. На основании анализа основных статистических признаков на этих изображениях была продемонстрирована чувствительность и специфичность до 85% для классификации воспаленных суставов у пациентов с РА [12]. Л.Д.Монтехо и др. указали подходы для извлечения более сложных эвристических признаков из ДОТ-изображений суставов PIP и метод использования таких производных признаков для диагностики РА [13]. Они также представили всесторонний анализ методов классификации признаков, извлеченных из изображения ДОТ, в том числе методы ближайших k-соседей, линейного и квадратичного дискриминантного анализа, самоорганизующихся карт и метод опорных векторов (МОВ) [14].

Совсем недавно Д.Лайтер и др. представили мульти-спектральную систему ДОТ непрерывного действия для

* Перевод с англ. В.Л.Дербова.

Y.Feng. College of Computer Science, Sichuan University, Chengdu 610065, China; Imaging Lab of the School of Computer Science, University of Birmingham, Birmingham B152TT, UK

L.Zhang, Y.Wang. College of Computer Science, Sichuan University, Chengdu 610065, China

D.Lightner, H.Deighani. Imaging Lab of the School of Computer Science, University of Birmingham, Birmingham B152TT, UK; e-mail: H.Deighani@cs.bham.ac.uk

выявления патофизиологических изменений в воспаленных суставах пальцев при РА. Были получены восстановленные карты насыщения кислородом, общего гемоглобина, воды и амплитуды рассеяния у здоровых субъектов со значительно большей вариабельностью между субъектами по сравнению с вариабельностью, наблюдаемой у пальцев одного и того же субъекта [15]. Ранее классификация РА с использованием данных, полученных в непрерывном режиме на одной длине волны, показала слабую дискриминацию при использовании статистических характеристик [12]. При этом точность классификации изображений ДОТ, полученных в непрерывном режиме, была невысокой (чувствительность 64%, специфичность 55%). Провести различие между воспаленными и невоспаленными суставами в условиях шума, высокой вариабельности между субъектами [15], некорректной и недоопределенной природы изображения в стационарной одноволновой ДОТ – сложная задача, требующая решения. Таким образом, базовые статистические характеристики недостаточны для различения воспаленных и невоспаленных состояний [12], и, следовательно, необходимы новые методы выделения признаков из ДОТ-изображений РА для классификации заболеваний. Разработке таких методов и посвящена настоящая работа.

В последние годы глубокие нейронные сети (ГНС) [16, 17], глубокие сети доверия [18] и рекуррентные нейронные сети [19–22] успешно применялись для классификации изображений, распознавания речи, визуального сопровождения [23] и обработки естественного языка [24]. Типичные примеры такого применения в области классификации изображений – AlexNet [16], VGGNet [25] и GoogLeNet [26]. Эти модели ГНС достигли существенных успехов в задачах классификации изображений, однако все они обучаются на большом наборе маркированных данных. Хотя наборы медицинских изображений обычно намного меньше, чем широко используемые типичные базы данных, были предприняты попытки в области медицинской визуализации, вдохновленные успехом моделей ГНС в классификации изображений. Например, чтобы получить лучшее представление входных данных, были предложены заплаточные (patch-based) модели ГНС для классификации рака молочной железы [27, 28], в которых вариант сети, основанный на AlexNet, использовался для извлечения признаков и разделения новообразований молочной железы на доброкачественные и злокачественные по гистопатологическим изображениям.

Н.Таджбахш и др. исследовали эффективность моделей сверточной нейронной сети (СНС) в четырех различных приложениях медицинской визуализации в рамках трех специальностей (радиология, кардиология и гастроэнтерология), включающих трансферное обучение для классификации, обнаружения и сегментации [29]. Они обнаружили, что предварительно обученные СНС с адекватной тонкой настройкой превосходили СНС, обученные с нуля, или выполняли задачу так же, причем СНС с тонкой настройкой были более устойчивыми к размеру обучающих наборов, чем СНС, обученные с нуля.

Хотя предпринимаются определенные усилия в области компьютерной диагностики РА и применения некоторых моделей ГНС для классификации медицинских изображений, область применения моделей ГНС к изображениям ДОТ для диагностики РА остается практически не исследованной. В настоящей работе для классификации РА по изображениям ДОТ мы применили три метода глу-

бокого обучения: магистральные сети (Highway) [30], остаточные сети (ResNet) [31] и СНС. В частности, представлены основные концепции и архитектуры трех методов глубокого обучения. Оцениваются и анализируются преимущества и ограничения различных методов. Новизна и основной вклад этой работы заключается в следующем: 1) глубокие нейронные сети впервые адаптируются и модифицируются применительно к задаче классификации РА с использованием изображений ДОТ; 2) для улучшения характеристик моделей, когда клинических данных недостаточно, применяется трансферное обучение; 3) для преодоления нехватки клинических изображений на первом этапе работы системы ДОТ предложена численная модель пальца с признаками воспаления, генерирующая недостающие данные для исследования классификации РА.

Статья организована следующим образом. В разд.2 приведены основные концепции трех глубоких нейронных сетей и описаны модели, разработанные для классификации РА. В разд.3 представлены изображения ДОТ, а в разд.4 – результаты их использования. Заключение и выводы содержатся в разд.5.

2. Методы

В данном разделе представлены концепции трех подходов к использованию ГНС, применение которых оказалось наиболее успешным в распознавании изображений. Каждый из них имеет специфические преимущества и ограничения при классификации изображений. Кроме того, представлены модели ГНС, основанные на этих трех подходах к классификации РА, а также детали использования трансферного обучения для повышения точности классификации.

2.1. Магистральная нейронная сеть

Магистральные (Highway) нейронные сети [30], созданные на основе рекуррентных нейронных сетей с длин-

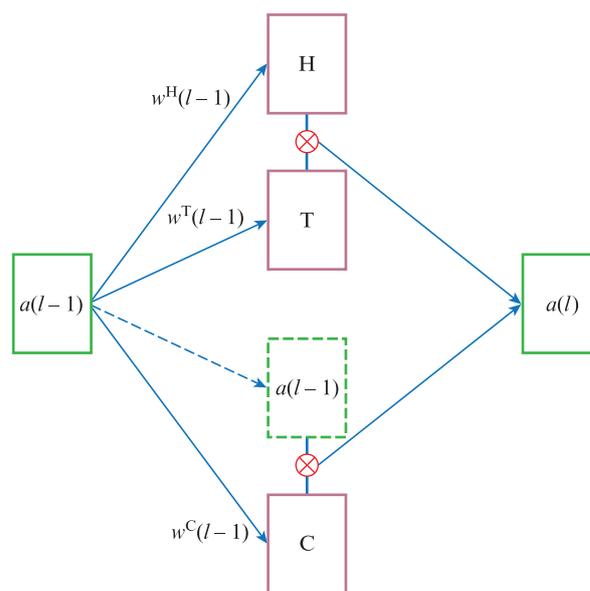


Рис.1. Строительный блок магистральной сети; Н – нелинейное преобразование. Два дополнительных нелинейных преобразования (Т и С) соответствуют вентилю преобразования Т и вентилю передачи С, $a(l)$ – выходные данные l -го слоя.

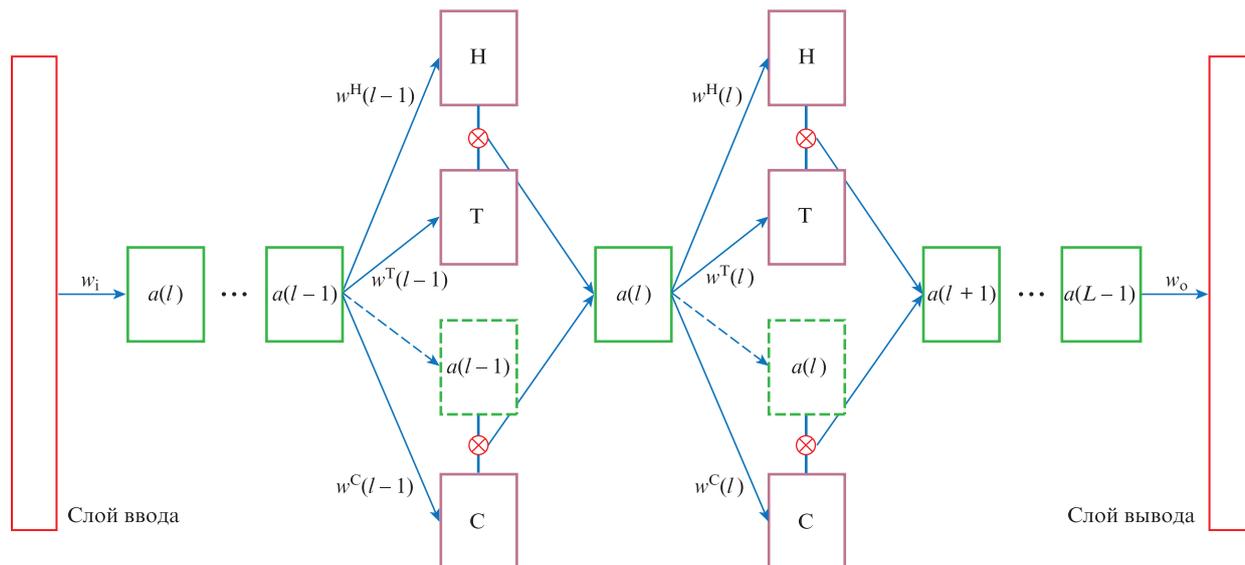


Рис.2. Архитектура магистральной сети прямого распространения с 18 скрытыми уровнями и 150 блоками на каждом уровне, предложенная для классификации РА, w_i – вес входного слоя, w_o – вес выходного слоя.

ной краткосрочной памятью (long short-term memory, LSTM), представляют собой методы построения сетей прямого распространения с сотнями или тысячами слоев. Они обучаются непосредственно с использованием стохастического градиентного спуска с различными функциями активации и приобретенными вентилями механизмами для регулирования потока информации. Вентильные механизмы позволяют нейронным сетям пропускать информацию через разные уровни информационных магистралей. Строительный блок магистральной сети показан на рис.1. Для магистральной нейронной сети определим $a_i(l)$ как выходные данные i -го блока ($i \in \{1, 2, \dots, N\}$) в l -м скрытом слое ($l \in \{1, 2, \dots, L - 2\}$) таким образом, что

$$a_i(l) = H(\text{net}_i^H(l))T(\text{net}_i^T(l)) + a_i(l-1)C(\text{net}_i^C(l)), \quad (1)$$

где N – размер слоя; L – общий уровень магистральной модели; H – нелинейное преобразование, за которым следует нелинейная функция активации; T – вентиль преобразования; C – вентиль переноса. Эти два вентиля определяют, какая часть выходных данных создается путем преобразования входных данных, а какая – путем их переноса соответственно; $\text{net}_i^H(l)$, $\text{net}_i^T(l)$ и $\text{net}_i^C(l)$ – полные входные данные нелинейного преобразования (H), вентиля преобразования (T) и вентиля переноса (C) i -го блока в l -м слое. Они выражаются как

$$\text{net}_i^H(l) = \sum [w_{ij}^H(l-1)a_j(l-1)], \quad (2)$$

$$\text{net}_i^T(l) = \sum [w_{ij}^T(l-1)a_j(l-1)], \quad (3)$$

$$\text{net}_i^C(l) = \sum [w_{ij}^C(l-1)a_j(l-1)]. \quad (4)$$

В настоящей работе применяется адаптированная для классификации РА магистральная сеть прямого распространения, которая состоит из $L - 2$ скрытых слоев (слоев магистрали), входного слоя и выходного слоя. Архитектура предлагаемой сети для РА показана на рис.2. Для данного обучающего набора $X = \{x(1), x(2), \dots, x(M)\}$ и тестового изображения, заданного как $T = \{t(1), t(2), \dots, t(K)\}$,

мы используем обучающий набор для оптимизации весов модели, а затем приобретенные веса для тестирования образцов из тестового набора.

В настоящей работе параметры сети экспериментально устанавливаются как $L = 20$ и $N = 150$. В число слоев предложенной модели входят входной слой и один выходной слой с функцией softmax, преобразующей вектор в нормированное на единицу распределение вероятностей с сохранением размерности, т. е. имеются 18 магистральных слоев. Процедура обучения и тестирования соответствует приведенному ниже алгоритму 1.

Алгоритм 1. Процедура обучения и тестирования магистральной сети, предлагаемой для классификации РА.

Ввод:

Набор данных обучения: $X = \{x(1), x(2), \dots, x(M)\}$.

Набор данных тестирования: $T = \{t(1), t(2), \dots, t(K)\}$.

Число итераций: I_t .

Вывод:

Прогноз для каждого тестируемого изображения:

$P = \{p(1), p(2), \dots, p(K)\}$.

// Инициализация

1. Инициализировать $\{w(l)^T, w(l)^H, w(l)^C\}_{i=1}^{(L-2)}$, w_i и w_o .

// Обучение магистральной сети

2. Для каждого $t \in [1, I_t]$ выполнить

3. Преобразование каждого обучающего изображения в вектор.

4. Прямой проход обучающих изображений.

5. Обновление $\{w(l)^T, w(l)^H, w(l)^C\}_{i=1}^{(L-2)}$, w_i и w_o .

6. Конец цикла.

7. Преобразование каждого тестирующего изображения в вектор.

8. Прямой проход тестирующих изображений для вычисления выходных данных сети: $P = \{p(1), p(2), \dots, p(K)\}$.

9. Выдача прогноза для тестируемых изображений:

$P = \{p(1), p(2), \dots, p(K)\}$.

2.2. Сверточные нейронные сети

СНС [32] – биологически инспирированные варианты полносвязных сетей, разработанные для распознавания зрительных образов непосредственно по двумерным изображениям. Распознавание достигается с помощью локальных связей и связанных весов с последующим пулингом какого-либо типа, что приводит к трансляции и сдвигу инвариантных объектов. СНС легче обучать, поскольку они имеют гораздо меньше параметров, чем полносвязные сети с таким же числом скрытых единиц. Кроме того, изображения имеют сильную двумерную локальную структуру: пиксели, которые находятся рядом друг с другом, имеют высокую корреляцию, тогда как в полносвязных сетях топология ввода полностью игнорируется. СНС хорошо осуществляют выделение локальных признаков, делая рецептивные поля скрытых единиц локальными.

СНС состоит из ряда слоев свертки и пулинга, за которыми могут следовать полностью подключенные слои, причем слой свертки состоит из нескольких обучаемых фильтров для выделения признаков. Каждый фильтр свернут с массивом входных данных, что создает двумерную карту признаков, которая определяет отклик этого фильтра в каждой пространственной позиции. Все карты признаков затем складываются стопкой для получения выходного массива данных слоя свертки. Пример того, как фильтр в слое свертки обрабатывает входные данные, показан на рис.3,а. Уровень пулинга используется для уменьшения пространственного размера представления, уменьшая, таким образом, число параметров и вычислений в сети. Он работает независимо на каждом срезе входных данных на установленной глубине и пространственно из-

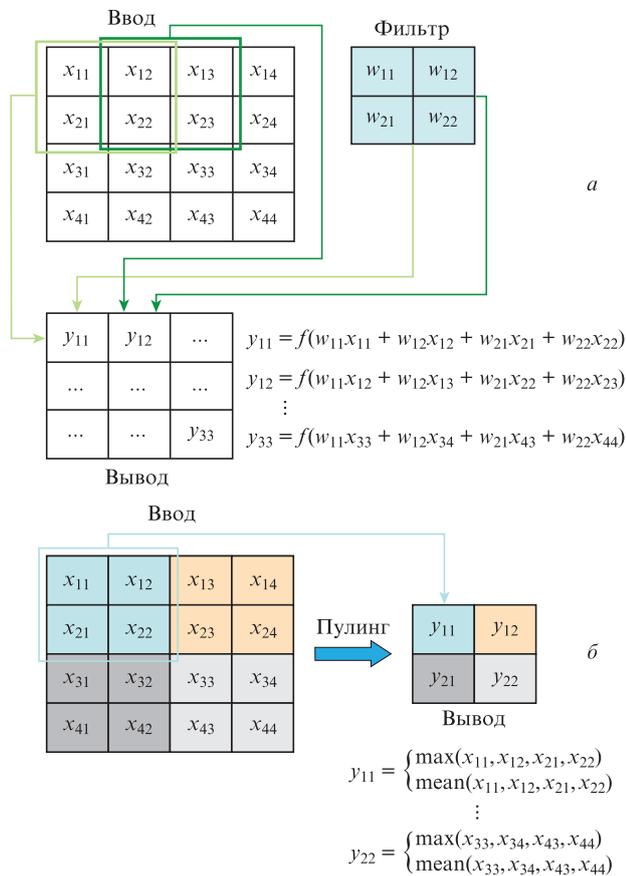


Рис.3. Операции свертки (а) и пулинга (б); f – функция активации.

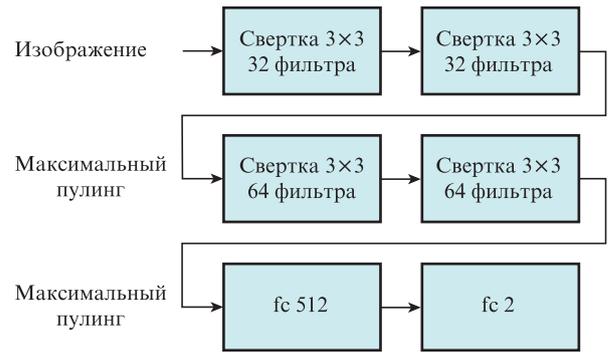


Рис.4. Архитектура сверточных нейронных сетей. Оператор максимального пулинга имеет размер 2×2 , а слои fc 512 и fc 2 – полностью подключенные слои с 512 выходными и двумя входными соединениями соответственно.

меняет размер этого среза, используя операцию «max» (для максимума) или «mean» (для усреднения). Пример операции пулинга показан на рис.3,б. Полностью подключенный слой состоит из нейронов, которые соединены со всеми подключениями в предыдущем слое и с выходными подключениями, рассчитанными с помощью матричного умножения с последующим смещением.

Для классификации РА используется проблемно-модифицированная сверточная сеть [32]. Архитектура модифицированной сверточной сети показана на рис.4. Она состоит из девяти уровней, включая входной. Второй уровень – уровень свертки, состоящий из 32 фильтров размером 3×3 для создания 32 карт признаков. За третьим уровнем, идентичным второму, следует слой максимального пулинга (размер 2×2 , страйд 2) для четвертого уровня, который дискретизирует с пониженной вдвое частотой каждый срез как по глубине, так по высоте, отбрасывая 75% подключений. Пятый и шестой уровни – уровни свертки, состоящие из 64 фильтров размером 3×3 и

Алгоритм 2. Процедура обучения и тестирования СНС для нашей классификации РА.

Ввод:

Набор данных обучения: $X = \{x(1), x(2), \dots, x(M)\}$.

Набор данных тестирования: $T = \{t(1), t(2), \dots, t(K)\}$.

Число итераций: I_t .

Вывод:

Прогноз для каждого тестируемого изображения:

$P = \{p(1), p(2), \dots, p(K)\}$.

// Инициализация

1. Инициализировать веса фильтров и полностью подключенных слоев.

// Обучение СНС

2. Для каждого $t \in [1, I_t]$ выполнить

3. Прямой проход по обучающим изображениям.

4. Обновление весов всех фильтров и полностью подключенных слоев.

5. Конец цикла.

6. Прямой проход по тестируемым изображениям для вычисления выходных данных сети: $P = \{p(1), p(2), \dots, p(K)\}$.

7. Выдача прогноза для тестируемых изображений:

$P = \{p(1), p(2), \dots, p(K)\}$.

служащие для создания 64 карт признаков. Седьмой уровень – уровень максимального пулинга с такими же установками, как и четвертый уровень. Оба последних уровня – полностью подключенные с 512 и 2 выходами соответственно. Функция активации всех уровней свертки и седьмого уровня – это функция линейного выпрямителя (ReLU) [33], а функция активации последнего слоя – это функция softmax. Когда изображение загружается в такую СНС, она относит его к одному из двух классов: с воспалением или без такового. Процедура обучения и тестирования СНС реализована в алгоритме 2.

2.3. Остаточные сети

Сети ResNet [31] реализуют остаточное обучение для решения проблемы деградации, когда точность насыщается, а затем быстро падает с увеличением глубины сети. Вместо стремления к тому, чтобы каждые несколько слоёв непосредственно соответствовали желаемому основному отображению, остаточное обучение явно позволяет этим слоям соответствовать остаточному отображению. Хотя обе формы способны аппроксимировать искомые функции с помощью теоремы универсального приближения [34], обучение по остаточному отображению намного проще, если искомая функция ближе к тождественному отображению, чем к нулевому. Как показано в [31], если добавленные слои могут быть построены как тождественные отображения, то более глубокая модель должна иметь не большую ошибку обучения, чем у ее аналога меньшей глубины. С этой точки зрения остаточное обучение позволяет очень глубоким сетям потенциально избежать проблемы деградации, и сети на его основе достигли очень высокой производительности при решении ряда визуальных задач [31].

Основной блок ResNet показан на рис.5, из которого видно, что входные данные x прежде всего преобразуются остаточным отображением с двумя слоями, затем сумма входных данных и результатов остаточного отображения берется в качестве выходных данных. Формально ядро остаточного алгоритма определяется как

$$y = \mathcal{F}(x) + x, \tag{5}$$

где x и y – рассматриваемые векторы входных и выходных данных, а функция $\mathcal{F}(x)$ означает остаточное отображение. Поскольку на рис.5 имеются два слоя, расчет остаточного отображения осуществляется по формуле $\mathcal{F}(x) = w(2)\sigma(w(1)x)$, в которой σ – функция линейного выпрямителя (ReLU) [33]; $w(1)$ и $w(2)$ – матрицы весов пер-

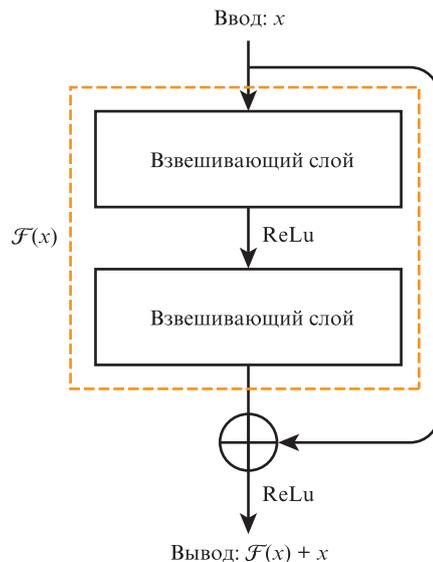


Рис.5. Строительный блок ResNet [31]. Действие внутри блока, окруженного штриховой линией, представляет собой процесс остаточного отображения.

вого и второго слоев; смещения опущены для упрощения обозначений. Наконец, второе нелинейное отображение σ выполняется над суммой над суммой y для формирования вектора выходных данных блока.

Видно, что размерности x и \mathcal{F} в формуле (5) должны быть равны. Если это не так, можно выполнить линейное проектирование P вектора x для согласования размерностей:

$$y = \mathcal{F}(x) + Px. \tag{6}$$

Более того, использование приведенных формулировок для полностью подключенных слоев сделано для простоты, и они применимы также к слоям свертки.

В нашей работе путем объединения блоков остаточного обучения можно построить глубокую остаточную сеть, чтобы извлечь отличительные признаки из необработанных изображений. Архитектура ResNet для нашей задачи показана на рис.6. В частности, вдохновленные философией VGGnets авторы ResNet построили ее со слоями свертки, которые в основном имеют фильтры 3×3 и следуют двум простым правилам проектирования: 1) для одного и того же размера выходной карты признаков слои имеют одинаковое количество фильтров; 2) если размер карты признаков уменьшен вдвое, число фильтров

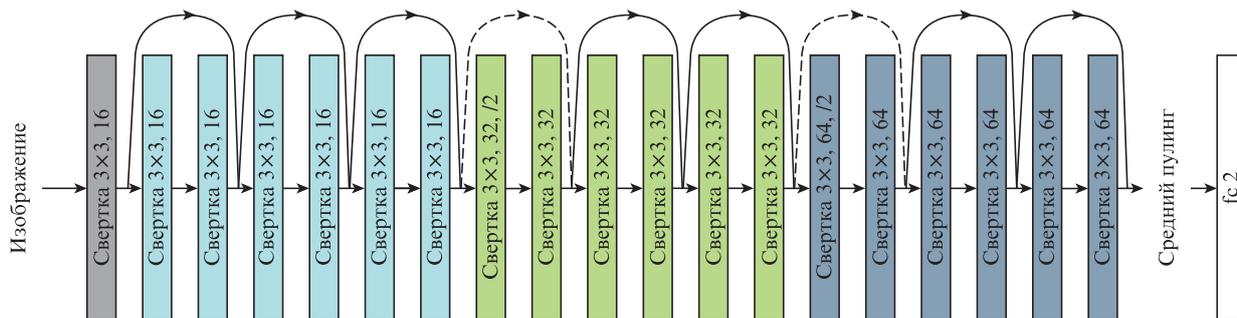


Рис.6. Архитектура остаточной сети с 20 слоями, предназначенной для классификации РА. Штриховые обходы увеличивают размерность. Данные входного изображения проходят множество блоков остаточной сети, затем средний пулинг (размер пула 2×2) и полностью подключенный выходной слой с активационной функцией softmax.

удваивается, чтобы сохранить временную сложность, приходящуюся на один слой. Нисходящая выборка выполняется непосредственно уровнями свертки, имеющими шаг, равный двум. Сеть заканчивается слоем глобального среднего пулинга и полностью подключенным слоем g -way с softmax, где g – число классов изображений в задаче классификации изображений на наличие или отсутствие воспаления. Процедура обучения и тестирования ResNet соответствует алгоритму 3.

Алгоритм 3. Процедура обучения и тестирования ResNet для нашей классификации РА.

Ввод:

Набор данных обучения: $X = \{x(1), x(2), \dots, x(M)\}$.

Набор данных тестирования: $T = \{t(1), t(2), \dots, t(K)\}$.

Число итераций: I_t .

Вывод:

Прогноз для каждого тестируемого изображения:

$P = \{p(1), p(2), \dots, p(K)\}$.

// Инициализация

1. Инициализировать веса фильтров и полностью подключенных слоев.

// Обучение ResNet

2. Для каждого $t \in [1, I_t]$ выполнить

3. Прямой проход по обучающим изображениям.

4. Обновление весов всех фильтров и полностью подключенных слоев.

5. Конец цикла.

6. Прямой проход по тестируемым изображениям для вычисления выходных данных сети: $P = \{p(1), p(2), \dots, p(K)\}$.

7. Выдача прогноза для тестируемых изображений:

$P = \{p(1), p(2), \dots, p(K)\}$.

2.4. Трансферное обучение

В области машинного обучения трансферное обучение – это способность системы распознавать и применять знания и навыки, полученные в ходе решения предыдущих задач, к новым задачам. Оно исследует, как модели могут учиться на одной задаче и применять полученные навыки к другой аналогичной задаче. В последние десятилетия методы трансферного обучения применялись для классификации текста [35–37], распознавания лиц [38, 39], распознавания речи [40, 41] и обучения с подкреплением [42].

В задачах медицинской диагностики получение образцов медицинских изображений является дорогостоящим и занимает много времени, особенно тогда, когда требуются точные метки истинности изображений. Генерация смоделированных изображений, похожих на клинические данные, и использование трансферного обучения могут быть хорошим способом улучшения производительности модели глубокого обучения. Для этого можно генерировать искусственные данные и использовать их для предварительного обучения ГНС, т.е. магистральных нейронных сетей, ResNet и СНС, причем параметры сети могут быть точно подстроены под реальные клинические данные. Если сгенерированные и реальные клинические изображения имеют схожие черты, ГНС потенциально могут применять знания, полученные на основе данных

имитационного моделирования, для последующего повышения своей эффективности на реальных клинических данных.

2.5. Сравнение с традиционными алгоритмами машинного обучения

Существует много традиционных алгоритмов машинного обучения, использованных для классификации медицинских изображений, например метод опорных векторов (МОВ) [43], дерево решений (ДР) [44] и классификатор случайных лесов (КСЛ) [45]. МОВ строит гиперплоскость в многомерном пространстве для разделения разных классов. Основная идея МОВ – найти максимальную маргинальную гиперплоскость, которая делит выборки данных на правильные классы. ДР – это непараметрический метод обучения под наблюдением. Его цель состоит в том, чтобы создать модель, которая прогнозирует значение целевой переменной, изучая простые правила принятия решений, выведенные из функций данных. КСЛ – метаоценщик, который производит подгонку по нескольким деревьям решений для разных подвыборок набора данных и использует усреднение для повышения точности прогнозирования и контроля сверхточности. Однако эти традиционные методы требуют человеческих знаний при извлечении признаков для достижения высокой точности в задачах классификации медицинских изображений. Глубокие же нейронные сети могут автоматически достигать большой мощности и гибкости, обучаясь представлению медицинских изображений в виде вложенной иерархии понятий.

3. Клинические изображения пальца

С марта по август 2018 г. было проведено пилотное исследование, участники которого были набраны через амбулаторную клинику исследовательской лаборатории Бирмингемского университета (Больница королевы Елизаветы, Бирмингем). Этическое одобрение было получено в рамках исследовательского проекта «Прогнозирование результатов у пациентов с воспалительным артритом» (RRK4678), реализуемого в университетских клиниках Бирмингема, участвующие в котором добровольцы предварительно давали письменное информированное согласие. Спецификации и протоколы сбора данных мультиспектральной бесконтактной системы ДОТ были изложены ранее [15].

Краткое описание процедуры получения изображения таково: пациент клал руку на платформу системы формирования оптического изображения пальца на пропускание с помощью ПЗС-камеры с воздушным охлаждением, в то время как широкополосный свет от точечного источника вводился в сустав с противоположной стороны. Изображения записывались при 14 положениях источника вдоль прямой линии сагиттального направления пальца с повторением на пяти длинах волн (650, 710, 730, 830 и 930 нм), выделяемых набором спектральных фильтров на поворотном диске, который размещался перед объективом камеры. С помощью программного пакета NIRFAST [46] для суставов пальца восстанавливались трехмерные томографические карты кислородного насыщения (StO_2), общего гемоглобина (tHb), содержания воды (H_2O), мощности (Sp) и амплитуды рассеяния (SA).

В настоящей работе данные были получены от 13 здоровых и больных РА добровольцев. Изображения вклю-

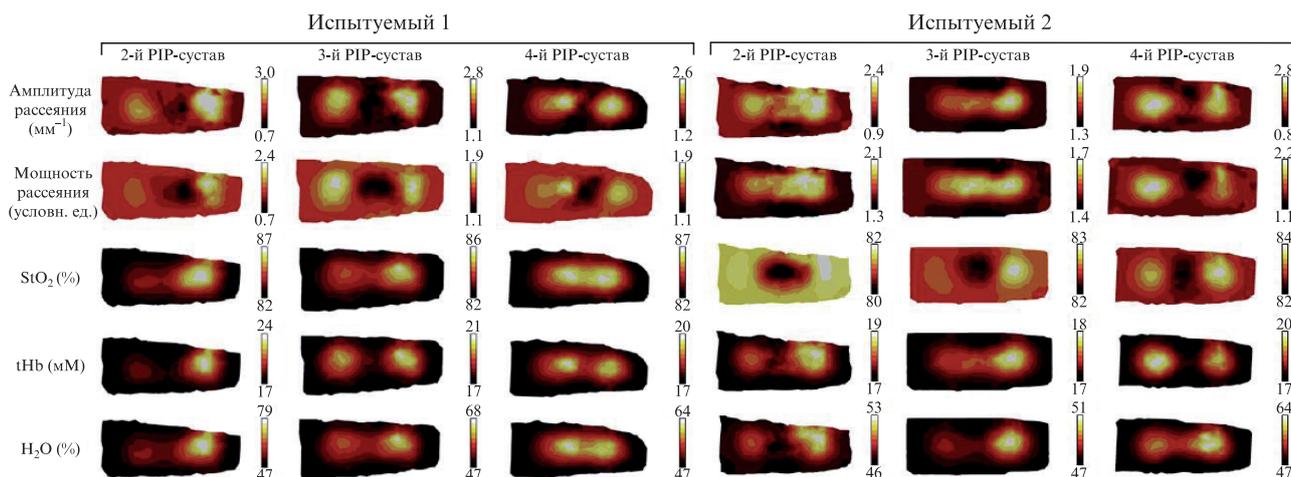


Рис.7. Центральные поперечные срезы реконструированных изображений патофизиологических параметров трех PIP-суставов двух здоровых испытуемых.

чали 2-й, 3-й и 4-й суставы, а для некоторых участников – также 5-й PIP сустав, на каждой руке, что в сумме давало 88 суставов пальцев. Типичный пример здорового PIP-сустава приведен на рис.7, где двумерные срезы вдоль центральной трансверсальной плоскости сустава соответствуют реконструированным картам клинически важных параметров StO₂, tHb, SP, SA и H₂O. К известным патофизиологическим изменениям воспаленного сустава относятся сниженная StO₂ (гипоксия) [47] и повышение tHb (синовиальный ангиогенез) [48] по сравнению со здоро-

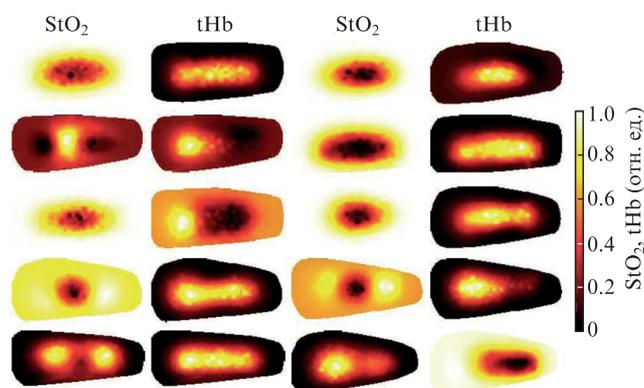


Рис.9. Клинические изображения, иллюстрирующие слабую дискриминацию. Первые два столбца – нормированные изображения StO₂ и tHb здоровых пальцев, каждая строка соответствует одному и тому же пальцу. Последние два столбца – нормированные изображения StO₂ and tHb больных пальцев, в каждой строке снова один и тот же палец.

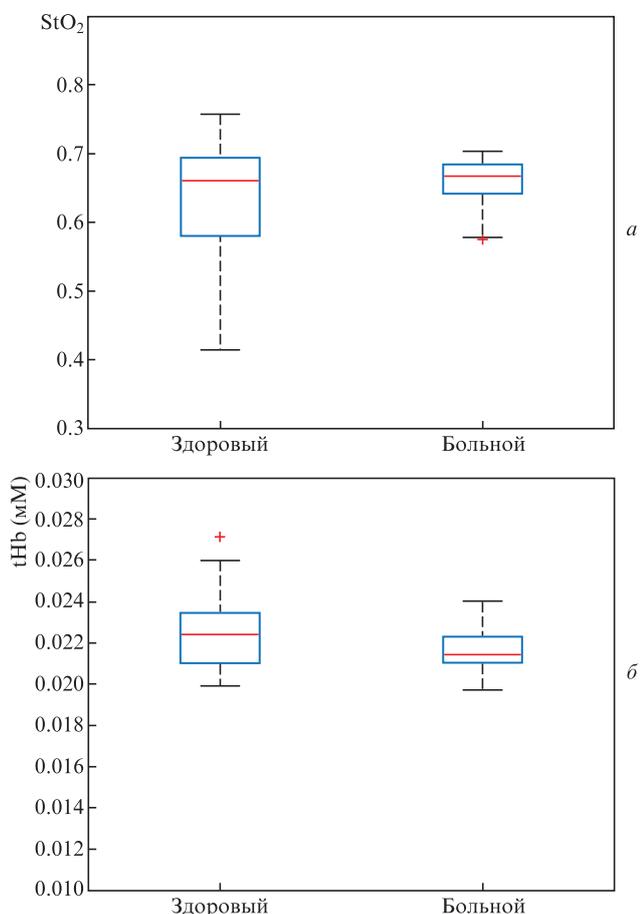


Рис.8. Распределения StO₂ и tHb, построенные по изображениям 88 пальцев (74 без воспаления, 14 с воспалением).

выми суставами. Поэтому для диагностики воспаленных и невоспаленных суставов в нашей работе использовались изображения StO₂ и tHb.

Классификация изображений StO₂ и tHb является амбициозной задачей из-за значительной их вариабельности от субъекта к субъекту, что согласуется с результатами работы Лайтера и др. [15]. Для иллюстрации дисперсии признаков на этих изображениях для данных от здоровых и больных испытуемых на рис.8 показаны графики восстановленных распределений StO₂ и tHb для всех 88 пальцев. Высокая вариабельность между испытуемыми ведет к ослаблению различия между больными и здоровыми; указанное ослабление дискриминации на рис.9 иллюстрируется примером нормализованных изображений StO₂ и tHb больного сустава. Все изображения на рис.9 относятся к центральному срезу каждого PIP-сустава и нормированы на собственное максимально значение.

4. Имитационные эксперименты

Описанная выше мультиспектральная система визуализации воспаленного сустава на основе стационарной ДОТ предназначена для восстановления оптических параметров в виде пространственных карт каждого пальца,

которые являются входными данными для описанных методов классификации. Процесс сбора клинических данных занимает много времени, поэтому для изучения поведения моделей ГНС был сгенерирован набор данных из имитационных моделей, который состоял из полученных экспериментально образцов как здоровых, так и пораженных РА пальцев. Кроме того, эксперименты для изучения поведения алгоритмов ГНС проводились именно на модельных данных, поскольку уровнем их дискриминации можно управлять, а точная классификация известна. Без использования этой методологии ГНС неработоспособны из-за нехватки данных.

4.1. Модельное изображение сустава пальца

Для создания набора данных имитационного моделирования была разработана двумерная модель сустава пальца (рис.10), состоящая из тканей кожи, кости, сустава и мышцы со значениями StO₂, tHb и H₂O для разных пальцев, взятыми из работ [49, 50] (см. табл.1). Эти значения были заданы на сетке метода конечных элементов, имевшей 2747 узлов и 5280 линейных тетраэдрических элементов. В модели был создан набор из восьми источников под суставом пальца и восьми детекторов поверх него, разделенных расстоянием 2.86 мм. Для каждой модели в области сустава задавался набор значений StO₂ и tHb, зависящий от того, является сустав здоровым (невоспаленным) или больным (воспаленным). Данные генерировались для каждой имитационной модели пальца с добавлением 1% гауссова шума. По этим данным изображения StO₂ и tHb восстанавливались с помощью итерационной процедуры Левенберга – Марквардта. Численное моделирование проводилось с применением находящегося в открытом доступе программного пакета NIRFAST, реализующего метод конечных элементов для моделирования распространения света [46]. Процедура создания изображений StO₂ и tHb описывается алгоритмом 4.

Табл.1. Параметры различных тканей пальца.

Ткань пальца	StO ₂	tHb (мМ/л)	Вода
Кожа	0.75	0.06	0.50
Мышца	0.80	0.10	0.50
Кость	0.80	0.08	0.40
Сустав	Def	Def	0.50

Примечание: StO₂ и вода измерены в долях. Для сустава StO₂ и tHb определялись (обозначение Def) исходя из уровней гипоксии и ангиогенеза здоровых и больных суставов. Для разных тканей пальца SA рассчитывались на основе данных [51].

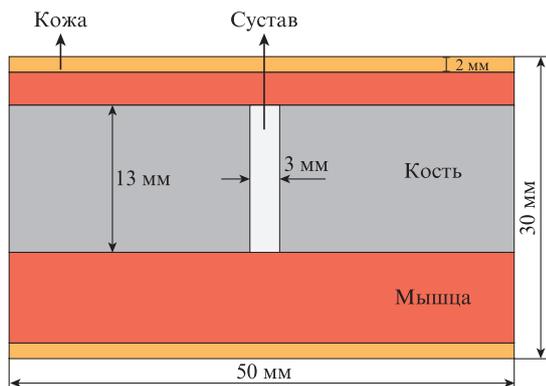


Рис.10. Двумерная модель пальца.

Алгоритм 4. Процедура генерации изображений StO₂ и tHb.

Ввод:
 Значение StO₂.
 Значение tHb.
 Уровень шума.

Вывод:
 Одно изображение StO₂ и одно изображение tHb.

// Чтобы создать одно изображение здорового пальца, следует:
 1. Сгенерировать физически реалистичную двумерную модель на основе значений StO₂, tHb и других параметров, приведенных в табл.1.
 2. Добавить к модели заданный процент гауссова шума.
 3. Реконструировать изображение из модели.
 4. Выдать одно изображение StO₂ и одно изображение tHb.

4.2. База данных имитационного моделирования

Для эксперимента были сгенерированы четыре разных базы данных имитационного моделирования, каждая из которых содержала 800 имитаций здоровых пальцев и 800 имитаций больных, по два изображения (StO₂ и tHb) на каждый. Следовательно, всего в каждую базу данных имитационного моделирования входило 1600 StO₂-изображений и 1600 tHb-изображений. На рис.11 представлено распределение значений StO₂ и tHb для каждой базы данных имитационного моделирования, из которых восстанавливались эти изображения.

Все реконструированные изображения нормированы согласно условию

$$z_i = (x_i - \min P) / (\max P - \min P),$$

где z_i – нормированное изображение, получаемое из изображения x_i; maxP и minP – максимальное и минимальное значения в базе данных P. Четыре базы данных контролировались на разных уровнях статистической дискриминации по коэффициенту перекрытия ξ, определяемому как мера согласия статистических распределений и точечной оценки перекрытия двух нормальных плотностей распределений. Большой коэффициент перекрытия означает большую неразличимость здоровых и больных суставов.

Среднее значение StO₂ для здоровых суставов выше, чем для больных, а среднее значение tHb для здоровых суставов ниже, чем для пораженных РА. Базы данных 1 (hdls) и 2 (hdhs) имеют одинаковые средние значения StO₂ и tHb, однако стандартное отклонение для базы данных 1 меньше, чем для базы 2. Для баз данных 3 (ldls) и 4 (ldhs) среднее значение StO₂ и tHb одно и то же, но база данных 3 имеет меньшее стандартное отклонение, чем база данных 4. Различие между средними значениями StO₂ и tHb для больных и здоровых суставов для баз данных 3 и 4 меньше, чем для баз данных 1 и 2. Кроме того, все четыре базы данных имитационного моделирования контролировались по коэффициенту перекрытия на разных уровнях сложности (большой коэффициент перекрытия ведет к ослаблению дискриминации). Действительно, на данной стадии неясно, насколько различными эти популяции (т.е. больные и здоровые) будут при исследованиях на больших выборках, а эффективность моделей ГНС при различ-

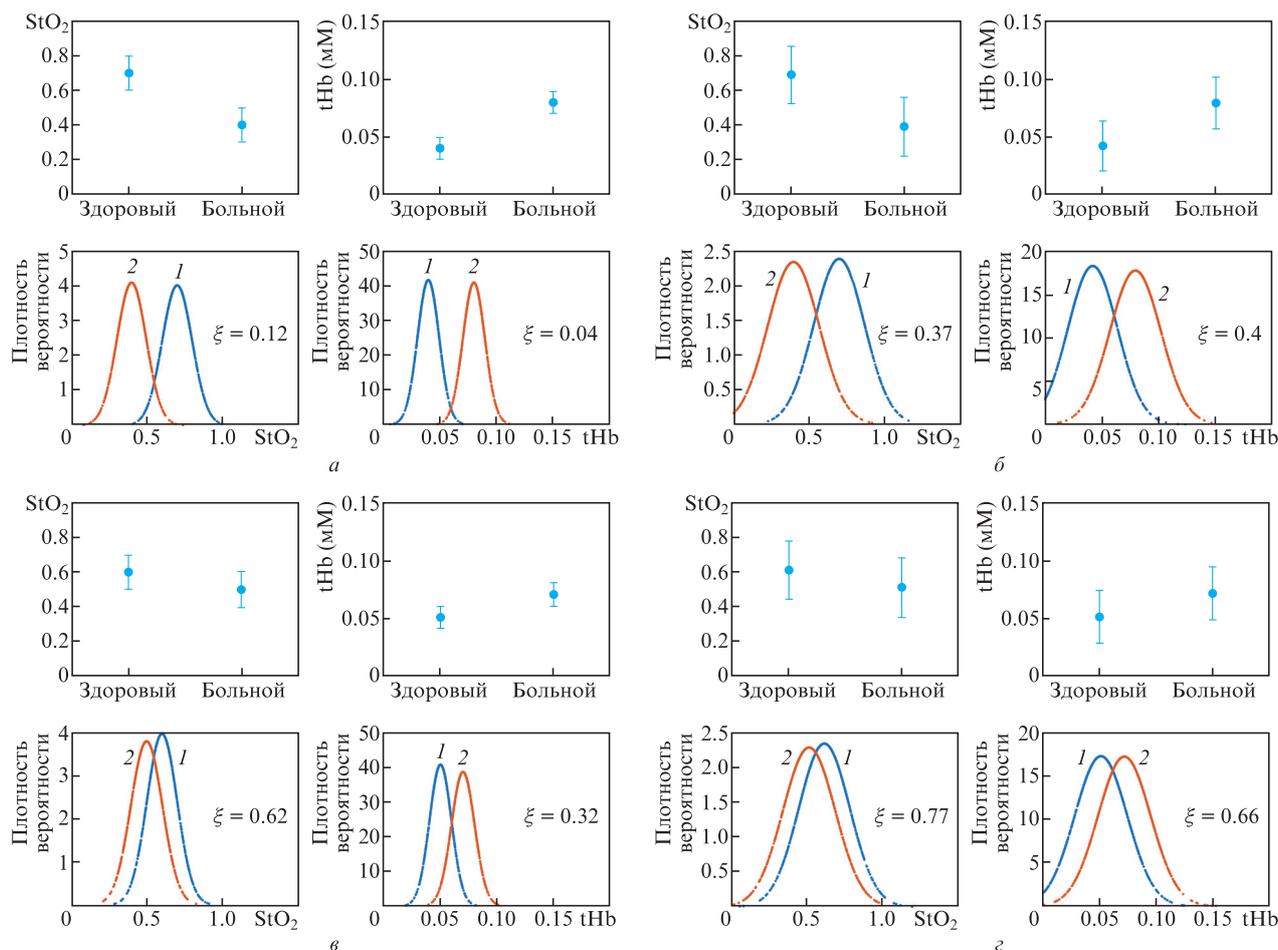


Рис.11. Четыре распределения исходных параметров для генерирования баз данных имитационного моделирования; 1 – здоровый сустав, 2 – больной; ξ – коэффициент перекрытия.

ных распределениях популяций нуждается в изучении. Установочные данные для четырех баз сведены в табл.2.

При классификации изображений на основе этих четырех баз данных имитационного моделирования необходимо решить две проблемы.

1. Слабая дискриминация. Все четыре базы данных основаны на перекрывающихся параметрах тканей, что ведет к слабой дискриминации между больными и здоровыми суставами. Подробности иллюстрирует рис.11, где, например, имеется перекрытие значений 37% (StO_2) и 40% (tHb). Это означает, что существуют изображения StO_2 и tHb с одинаковыми распределениями, но относящиеся к разным классам (больные и здоровые).

2. Шумы. Все имитационные данные имеют 1% гауссова шума, которые введен для обеспечения более реалистичного сценария; пример показан на рис.12. Очевидно,

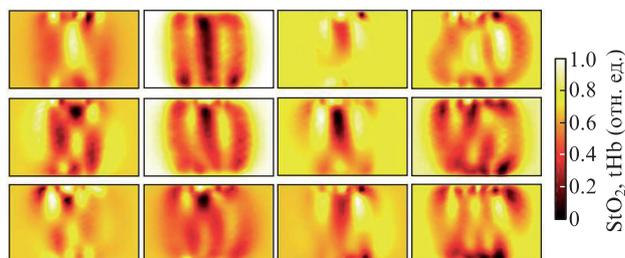


Рис.12. Примеры изображений из базы данных имитационного моделирования 1 (hdls) демонстрирующие артефакт. Первые два столбца – нормированные изображения StO_2 и tHb здоровых пальцев, в каждой строке один и тот же палец, последние два столбца – нормированные изображения StO_2 и tHb больных пальцев, в каждой строке также один и тот же палец.

что появятся артефакты изображения как для больных, так и для здоровых суставов, что приведет к их слабой дискриминации.

4.3. Результаты экспериментов и их анализ

В первом эксперименте новое изображение генерируется для каждого пальца простым сложением изображений StO_2 и tHb, что трактуется как введение нового параметра – оптического показателя. Затем исследуется его влияние на классификацию РА путем использования разных изображений, таких как только StO_2 или только tHb либо $StO_2 + tHb$, для ввода в нейронные сети. В результа-

Табл.2. Установочные данные для четырех баз данных.

Базы данных	Здоровые суставы (StO_2)	Больные суставы (StO_2)	Здоровые суставы (tHb)	Больные суставы (tHb)
1 (hdls)	0.70 ± 0.10	0.40 ± 0.10	0.04 ± 0.01	0.08 ± 0.01
2 (hdhs)	0.70 ± 0.17	0.40 ± 0.17	0.04 ± 0.023	0.08 ± 0.023
3 (ldls)	0.60 ± 0.10	0.50 ± 0.10	0.05 ± 0.01	0.07 ± 0.01
4 (ldhs)	0.60 ± 0.17	0.50 ± 0.17	0.05 ± 0.023	0.07 ± 0.023

Примечание: hd и ld – большое и малое различия между средними значениями для больных и здоровых суставов; hs и ls – большое и малое стандартные отклонения для больных и здоровых суставов.

те получены средняя точность и стандартное отклонение пятикратной перекрестной валидации, представленные в табл.3 для магистральной сети, ResNet, СНС, а также для метода опорных векторов (МОВ) [43], дерева решений (ДР) [44] и классификатора случайных лесов (КСЛ) [45], которые взяты за базовый уровень (тестировались многоклассовый линейный метод МОВ со стратегией «один против всех» и параметром штрафа 1.0 и метод КСЛ с 10 деревьями).

Результаты показывают, что все четыре метода работают хорошо на изображениях из базы данных 1. В частности, все они достигают средней точности классификации свыше 95%. Наивысшей средней точности (98.6%) ResNet достигает на изображениях StO₂ и на изображениях tHb (98.9%), а магистральная сеть и ResNet лучше всего работают на изображениях StO₂ + tHb (средняя точность 99.3%). Для всех шести алгоритмов с ростом перекрытия классов здоровых и больных суставов рабочие характеристики всех методов ухудшаются, особенно для МОВ и ДР. С точки зрения алгоритмов, метод ResNet превосходит три других метода для изображений из всех четырех баз данных. Преимущество ResNet по сравнению с тремя другими методами увеличивается с ростом перекрытия между образцами классов больных и здоровых. Три модели глубокого обучения нейронных сетей, магистральная, ResNet и СНС, позволяют получать наилучшие результаты на изображениях StO₂ + tHb и лучшие результаты на изображениях tHb по сравнению с изображениями StO₂.

Во втором эксперименте для глубоких нейронных сетей изучалось влияние трансферного обучения с использованием всех четырех баз данных, которые имели различное перекрытие. В частности, весовые параметры, полученные на базе данных 1, использованы в качестве параметров инициализации сетей при решении классификационных задач на базах данных 2–4.

Для получения статистических результатов, приведенных в табл.4, использовался метод пятикратной кросс-валидации. Результаты показывают, что все три модели глубоких нейронных сетей улучшают работу при использовании трансферного обучения. Например, ResNet достигает наивысшей средней точности 90.4% на изображениях

Табл.4. Средние точности (%) и стандартные отклонения для трех глубоких нейронных сетей при трансферном обучении.

Номер базы данных	Вводимое изображение	Магистраль	ResNet	СНС
2	StO ₂	88.6 ± 1.0	90.4 ± 1.5	88.1 ± 2.3
	tHb	88.6 ± 1.1	90.1 ± 1.7	88.5 ± 1.3
	StO ₂ + tHb	91.9 ± 1.4 [†]	92.8 ± 0.9 [†]	90.7 ± 1.5
3	StO ₂	82.6 ± 1.1	88.1 ± 3.4	81.8 ± 2.6
	tHb	87.9 ± 1.9	88.6 ± 1.4	86.5 ± 1.4
	StO ₂ + tHb	87.6 ± 1.5	90.3 ± 2.5	87.8 ± 1.8
4	StO ₂	69.3 ± 1.4	74.8 ± 2.5 [†]	69.4 ± 2.9
	tHb	73.2 ± 1.7 [†]	75.9 ± 4.8	74.1 ± 2.1 [†]
	StO ₂ + tHb	73.9 ± 1.4	77.3 ± 4.5	75.8 ± 1.9 [†]

[†] Существенное отклонение точности метода от точности в табл.3 на уровне 0.05 по критерию Уилкоксона.

StO₂ в базе данных 2, а также наилучшим образом работает на изображениях tHb и StO₂ + tHb со средней точностью 90.1% и 92.8%. Показано, что трансферное обучение может улучшить качество обработки ДОТ-изображений. Это подчеркивает преимущества использования трансферного обучения при моделировании для предсказания состояния сустава в условиях дефицита экспериментальных данных.

5. Заключение

Для достижения высокой точности диагностики воспаления и РА следует опираться не только на различимость исходных данных (малые изменения внутри каждого класса и большое различие между классами), но и на точный и объективный алгоритм классификации. В настоящей работе представлен метод создания имитационных ДОТ-изображений здоровых и больных суставов пальцев с тремя моделями ГНС для классификации. Цель состояла в отнесении каждого сустава РИР к классу воспаленных из-за РА или невоспаленных на основе анализа признаков изображения, автоматически выделяемых моделями ГНС. В частности, были исследованы три новейшие

Табл.3. Средние точности (%) и стандартные отклонения для трех методов глубокого обучения, а также МОВ, дерева решений (ДР) и классификатора случайных лесов (КСЛ).

Номер базы данных	Вводимое изображение	Магистраль	ResNet	СНС	МОВ	ДР	КСЛ
1	StO ₂	97.6 ± 1.8	98.6 ± 0.6	95.3 ± 1.9 [†]	96.8 ± 1.4 [†]	90.6 ± 1.5 [†]	92.9 ± 1.8 [†]
	tHb	98.8 ± 0.5	98.9 ± 0.6	98.7 ± 0.8	97.9 ± 0.7	96.3 ± 1.0 [†]	97.3 ± 1.4
	StO ₂ + tHb	99.3 ± 0.5	99.3 ± 0.4	99.1 ± 0.7	98.0 ± 0.8 [†]	97.1 ± 0.8 [†]	98.3 ± 1.0
2	StO ₂	88.3 ± 1.8	88.8 ± 1.9	87.8 ± 2.3	79.1 ± 10.7 [†]	79.3 ± 2.6 [†]	84.6 ± 2.7
	tHb	88.4 ± 0.8	89.1 ± 1.8	88.4 ± 1.6	72.9 ± 17.9 [†]	77.7 ± 1.7 [†]	83.3 ± 1.3 [†]
	StO ₂ + tHb	89.4 ± 1.6	90.2 ± 1.4	89.0 ± 1.6	79.4 ± 14.4 [†]	80.6 ± 1.4 [†]	87.3 ± 1.6 [†]
3	StO ₂	80.1 ± 6.0 [†]	86.3 ± 3.2	78.8 ± 5.1 [†]	77.3 ± 5.5 [†]	66.9 ± 3.3 [†]	72.2 ± 4.1 [†]
	tHb	86.3 ± 1.8	87.0 ± 4.4	86.3 ± 2.0	83.4 ± 1.2	75.9 ± 2.7 [†]	82.1 ± 1.8
	StO ₂ + tHb	86.8 ± 2.0	88.4 ± 2.0	86.6 ± 1.9	75.9 ± 13.7 [†]	77.1 ± 1.4 [†]	82.1 ± 2.0 [†]
4	StO ₂	68.6 ± 2.6	70.1 ± 3.3	68.6 ± 2.7	54.2 ± 3.6 [†]	57.4 ± 2.7 [†]	61.6 ± 1.4 [†]
	tHb	70.5 ± 1.5	73.8 ± 5.3	70.4 ± 2.7	59.8 ± 5.1 [†]	57.6 ± 2.8 [†]	64.8 ± 2.1 [†]
	StO ₂ + tHb	71.4 ± 2.7	75.2 ± 8.8	71.6 ± 1.7	55.0 ± 8.6 [†]	57.1 ± 3.6 [†]	65.1 ± 1.7 [†]

[†] Существенное отклонение точности метода от точности ResNet на уровне 0.05 по критерию Уилкоксона.

модели ГНС: магистральная нейронная сеть (Highway), СНС, и ResNet, для извлечения отличительных признаков из ДОТ-изображений с целью улучшения точности диагностики. Чтобы лучше понять алгоритмы классификации, выделены четыре базы данных, контролируемые по коэффициентам перекрытия изображений StO₂ и tHb. Затем изучена эффективность моделей ГНС с точки зрения повышения точности диагностики в исследованиях больных артритом.

Впервые показано, что восстановленные ДОТ-изображения StO₂ и tHb больных и здоровых суставов пальцев демонстрируют высокую вариабельность от одного пациента к другому. Основная трудность состояла в том, что: 1) изображения здоровых и больных пальцев имеют сходную структуру, что может привести к слабой дискриминации между ними; 2) степень перекрытия между различными субъектами в клинических условиях неизвестна; 3) для детальной оценки алгоритмов на основе ГНС подробных клинических данных недостаточно. Таким образом, модель сустава пальца была использована для создания баз данных, которые позволили исследовать влияние слабой дискриминации в случае компьютерных методов диагностики.

Ранее некоторые исследователи с целью распознавания данных визуализации в пространственно-частотной области использовали традиционные методы, например МОВ и регрессор случайных лесов [52, 53]. Для понимания ДОТ-изображений пальца применялся также метод МОВ [14]. В настоящей работе для диагностики ДОТ-изображений пальцев мы использовали методы глубокого обучения. Кроме того, необработанные ДОТ-изображения трудно интерпретировать, поэтому для исследования возможностей тестируемых методов нами была разработана имитационная модель для управления уровнями сложности задач классификации РА. Такая модель помогает понять, какие алгоритмы работают лучше других в разных условиях.

В экспериментах по моделированию показано, что все три модели, ГНС, МОВ и КСЛ, работают лучше на изображениях StO₂ + tHb, чем на изображениях StO₂ или tHb в отдельности. Это указывает на то, что измерения как насыщения кислородом (показатель гипоксии), так и содержания крови (васкуляризация) несут информацию о различиях, а их объединение дает лучшие результаты при обнаружении РА. Хотя использование других восстанавливаемых данных, таких как содержание воды и рассеяние, должно стать предметом будущих исследований, экспериментальные результаты также показывают, что все три модели ГНС превосходят традиционные методы машинного обучения (МОВ, ДР и КСЛ). С увеличением перекрытия в распределении StO₂ и tHb точность моделей ГНС оказывается выше, чем у традиционных методов машинного обучения, что указывает на большую надежность моделей на основе ГНС по сравнению с традиционными методами. Наконец, для повышения точности классификации использовалось трансферное обучение, результаты которого показали, что оно может улучшить точность классификации, следовательно в будущем должно тестироваться на клинических данных.

Были представлены результаты классификации трех разных моделей ГНС, демонстрирующие их способность выявлять воспаленные и невоспаленные суставы пальцев с целью диагностики РА. Анализ представленных результатов показывает, что модели ГНС являются надежными

при использовании ДОТ даже тогда, когда восстановленные оптические карты демонстрируют большое перекрытие между двумя различными функциональными состояниями и, как правило, менее дифференцированы для разных классов. Эти результаты подчеркивают возможность использования моделей ГНС в качестве компьютерного инструмента в диагностических системах ДОТ и требуют более масштабных перспективных исследований, чтобы окончательно продемонстрировать конечную клиническую полезность этого подхода.

Работа поддержана Национальным фондом естественных наук Китая (гранты 61772353 и 61332002), Исследовательской группой Фонда молодежных научных и технологических инноваций провинции Сычуань (грант 2016TD0018), Образовательным фондом Фок Ин Тунга (грант 151068), а также EPSRC в рамках программы обучения Центра докторантуры Sci-Phy-4-Health (EP/L016346/1).

1. Majithia V., Geraci S.A. *Am. J. Med.*, **120** (11), 936 (2007).
2. Helmick C.G., Felson D.T., Lawrence R.C., Gabriel S., Hirsch R., Kwoh C.K., Liang M.H., Kremers H.M., Mayes M.D., Merkel P.A., Pillemer S.R., Reveille J.D., Stone J.H., Workgrp N.A.D. *Arthritis Rheum.*, **58** (1), 15 (2008).
3. Nell V.P.K., Machold K.P., Eberl G., Stamm T.A., Uffmann M., Smolen J.S. *Rheumatology*, **43** (7), 906 (2004).
4. Landewe R.B.M., Boers M., Verhoeven A.C., Westhovens R., van de Laar M.A.F.J., Markuse H.M., van Denderen J.C., Westedt M.L., Peeters A.J., Dijkmans B.A.C., Jacobs P., Boonen A., van der Heijde D.M.F.M., van der Linden S. *Arthritis Rheum.*, **46** (2), 347 (2002).
5. Scheel A.K., Hermann K.G.A., Ohrndorf S., Werner C., Schirmer C., Detert J., Bollow M., Hamm B., Muller G.A., Burmester G.R., Backhaus M. *Ann. Rheum. Dis.*, **65** (5), 595 (2006).
6. Wakefield R.J., O'Connor P.J., Conaghan P.G., McGonagle D., Hensor E.M.A., Gibbon W.W., Brown C., Emery P. *Arthrit. Rheum. Arthr.*, **57** (7), 1158 (2007).
7. Klarlund M., Ostergaard M., Jensen K.E., Madsen J.L., Skjodt H., Lorenzen I., Grp T. *Ann. Rheum. Dis.*, **59** (7), 521 (2000).
8. Haavardsholm E.A., Boyesen P., Ostergaard M., Schildvold A., Kvien T.K. *Ann. Rheum. Dis.*, **67** (6), 794 (2008).
9. Durduran T., Choe R., Baker W.B., Yodh A.G. *Rep. Prog. Phys.*, **73** (7), 076701 (2010).
10. Tromberg B.J., Pogue B.W., Paulsen K.D., Yodh A.G., Boas D.A., Cerussi A.E. *Med. Phys.*, **35** (6), 2443 (2008).
11. Giacometti P., Diamond S.G. *Bioanalysis Adv. Mat.*, **3**, 57 (2013).
12. Hielscher A.H., Kim H.K., Montejo L.D., Blaschke S., Netz U.J., Zwaka P.A., Illing G., Muller G.A., Beuthan J. *IEEE Trans. Med. Imaging*, **30** (10), 1725 (2011).
13. Montejo L.D., Jia J.F., Kim H.K., Netz U.J., Blaschke S., Muller G.A., Hielscher A.H. *J. Biomed. Opt.*, **18** (7), 076001 (2013).
14. Montejo L.D., Jia J.F., Kim H.K., Netz U.J., Blaschke S., Muller G.A., Hielscher A.H. *J. Biomed. Opt.*, **18** (7), 076002 (2013).
15. Lighter D., Hughes J., Styles I., Filer A., Dehghani H. *Biomed. Opt. Express*, **9** (4), 1445 (2018).
16. Krizhevsky A., Sutskever I., Hinton G.E. *Commun. ACM*, **60** (6), 84 (2017).
17. Chen Y.Y., Zhang L., Yi Z. *Inform. Sci.*, **424**, 27 (2018).
18. Mohamed A.R., Dahl G.E., Hinton G. *IEEE Trans. Audio Speech*, **20** (1), 14 (2012).
19. Zhang L., Yi Z. *Chaos Soliton Fract.*, **33** (3), 979 (2007).
20. Zhang L., Yi Z. *IEEE Trans. Neural Networks*, **22** (7), 1021 (2011).
21. Zhang L., Yi Z., Amari S. *IEEE Trans. Neur. Net. Lear.*, **29** (11), 5242 (2018).
22. Zhang L., Yi Z., Yu J.L. *IEEE Trans. Neural Networks*, **19** (1), 158 (2008).
23. Wang L.T., Zhang L., Yi Z. *IEEE Trans Cybernetics*, **47** (10), 3172 (2017).
24. LeCun Y., Bengio Y., Hinton G. *Nature*, **521** (7553), 436 (2015).
25. Simonyan K., Zisserman A. arXiv preprint arXiv:1409.1556 (2014).
26. Szegedy C., Liu W., Jia Y.Q., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. *Proc. IEEE CVPR* (Boston, 2015, pp 1–9).

27. Spanhol F.A., Oliveira L.S., Petitjean C., Heutte L. *Proc. IEEE IJCNN* (Vancouver, 2016, pp 2560–2567).
28. Spanhol F.A., Cavalin P.R., Oliveira L.S., Petitjean C., Heutte L. *IEEE Trans. Syst. Man, Cyb.*, 1868 (2017).
29. Tajbakhsh N., Shin J.Y., Gurudu S.R., Hurst R.T., Kendall C.B., Gotway M.B., Liang J.M. *IEEE Trans. Med. Imaging*, **35** (5), 1299 (2016).
30. Srivastava R.K., Greff K., Schmidhuber J. arXiv preprint arXiv:1505.00387 (2015).
31. He K.M., Zhang X.Y., Ren S.Q., Sun J. *Proc. IEEE CVPR* (Las Vegas, 2016, 770–778).
32. Lecun Y., Bottou L., Bengio Y., Haffner P. *Proc. IEEE*, **86** (11), 2278 (1998).
33. Nair V., Hinton G.E. *Proc. 27th Int. Conf. on Machine Learning (ICML-10)* (Haifa, 2010, pp 807–814).
34. Hornik K. *Neural Networks*, **4** (2), 251 (1991).
35. Quattoni A., Collins M., Darrell T. *Proc. IEEE CVPR* (Anchorage, 2008, pp 2300–2307).
36. Oquab M., Bottou L., Laptev I., Sivic J. *Proc. IEEE CVPR* (Columbus, 2014, pp 1717–1724).
37. Zhu Y., Chen Y., Lu Z., Pan S.J., Xue G.R., Yu Y., Yang Q. *Proc. 25th Conf. on Artificial Intelligence* (San Francisco, 2011, p 4057).
38. Ahmed A., Yu K., Xu W., Gong Y.H., Xing E. *Lect. Notes Comput. Sci.*, **5304**, 69 (2008).
39. Cao X.D., Wipf D., Wen F., Duan G.Q., Sun J. *Proc. IEEE ICCV* (Sydney, 2013, pp 3208–3215).
40. Deng J., Zhang Z.X., Marchi E., Schuller B. *Proc. IEEE ACII* (Geneva, 2013, pp 511–516).
41. Huang J.T., Li J.Y., Yu D., Deng L., Gong Y.F. *Proc. IEEE ICA SSP* (Vancouver, 2013, pp 7304–7308).
42. Taylor M.E., Stone P. *J. Mach. Learn. Res.*, **10**, 1633 (2009).
43. Vapnik V.N., Vapnik V. *Statistical Learning Theory* (New York: Wiley, 1998).
44. Safavian S.R., Landgrebe D. *IEEE Trans. Syst. Man Cyb.*, **21** (3), 660 (1991).
45. Breiman L. *Mach. Learn.*, **45** (1), 5 (2001).
46. Dehghani H., Eames M.E., Yalavarthy P.K., Davis S.C., Srinivasan S., Carpenter C.M., Pogue B.W., Paulsen K.D. *Commun. Numer. Meth. Eng.*, **25** (6), 711 (2009).
47. Ng C.T., Biniecka M., Kennedy A., McCormick J., FitzGerald O., Bresnihan B., Buggy D., Taylor C.T., O'Sullivan J., Fearon U., Veale D.J. *Ann. Rheum. Dis.*, **69** (7), 1389 (2010).
48. Falchuk K.H., Goetzl E.J., Kulka J.P. *Am. J. Med.*, **49** (2), 223 (1970).
49. Ash H.E., Unsworth A. *Proc. Instn. Mech Engrs*, **211H** (5), 377 (1997).
50. Yuan Z., Zhang Q.Z., Sobel E.S., Jiang H.B. *Biomed. Opt. Express*, **1** (1), 74 (2010).
51. Klose A.D., Hielscher A.H. *Med. Phys.*, **26** (8), 1698 (1999).
52. Rowland R., Ponticorvo A., Baldado M., Kennedy G.T., Burmeister D.M., Christy R.J., Bernal N.P., Durkin A.J. *J. Biomed. Opt.*, **24** (5), 056007 (2019).
53. Panigrahi S., Gioux S. *J. Biomed. Opt.*, **24** (7), 071606 (2019).