

Быстрая оптическая идентификация гемолиза и липемии в образцах сыворотки крови: компьютерное зрение и спектроскопия диффузного отражения

© Г.М. Денисенко^{1,2}, Р.Р. Фитагдинов^{3,4,5}, Б.П. Якимов^{2,6,7}, А.А. Бирюков^{7,8}, Ю.А. Шитова⁵,
Е.Н. Керунту⁷, А.С. Шкода⁷, Е.А. Ширшин^{2,6}

¹ Институт регенеративной медицины, Первый Московский государственный медицинский университет им. И.М. Сеченова,

119991 Москва, Россия

² Лаборатория клинической биофотоники, научно-технический парк Биомедицины, Первый Московский государственный медицинский университет им. И.М. Сеченова, 119991 Москва, Россия

³ Московский физико-технический институт, 141701 Долгопрудный, Россия

⁴ Институт ядерных исследований РАН, 117312 Москва, Россия

⁵ ООО „Медеум“, 127566 Москва, Россия

⁶ Московский государственный университет им. М.В. Ломоносова, физический факультет, 119234 Москва, Россия

⁷ Городская клиническая больница № 67 им. Л.А. Ворохобова ДЗМ, 123423 Москва, Россия

⁸ Лаборатория по поддержке принятия врачебных решений на базе технологий искусственного интеллекта, Первый Московский государственный медицинский университет им. И.М. Сеченова, 119991 Москва, Россия

e-mail: eshirshin@gmail.com

Поступила в редакцию 12.05.2023 г.

В окончательной редакции 18.08.2023 г.

Принята к публикации 28.09.2023 г.

Исследована аналитическая способность двух подходов, основанных на 1) спектроскопии диффузного отражения сыворотки крови в видимом и ближнем ИК диапазоне (500–1000 nm), и 2) компьютерном зрении — в классификации образцов сыворотки крови на норму, пригодную для анализа, и образцы с гемолизом и липемией. Установлены высокая чувствительность и специфичность (более 97%) спектроскопии диффузного отражения в классификации образцов сыворотки крови, однако этот метод требует подвода измерительного зонда к образцу. Методы компьютерного зрения позволили определять пригодность образца для дальнейшего анализа с меньшей точностью классификации, но в более сложных условиях, в частности при движении по конвейерной линии в клиничко-диагностической лаборатории. Преимуществом исследованных методик, помимо высокой точности преаналитической классификации, является простота их технической реализации и возможность характеризовать образцы без дополнительного пробоотбора сыворотки крови, что говорит об их перспективности в качестве методов преаналитического анализа образцов сыворотки крови.

Ключевые слова: спектроскопия диффузного отражения, компьютерное зрение, липемия, гемолиз, сыворотка крови, преаналитика.

DOI: 10.61011/OS.2023.09.56618.5025-23

Введение

Биохимический анализ крови является одним из рутинных анализов любой клиничко-диагностической лаборатории. Подавляющая часть биохимических анализов крови проводится на сыворотке крови после седиментации эритроцитов путем центрифугирования образца — лаборант или автоматизированная система осуществляет пробоотбор сыворотки из центрифугированного образца венозной крови для ее смешивания с активным

агентом и последующего считывания концентраций интересующих аналитов, определение которых часто связано с использованием фотометрического подхода [1,2]. Корректность проведенного результата биохимического анализа крови является критической как при постановке диагноза, так при выписке пациента [3,4].

Большинство ошибок лабораторного тестирования возникает на преаналитической фазе всего процесса [5]. Основной причиной ошибочных результатов при проведении биохимических анализов (как ложноположитель-

ных, так и ложноотрицательных) является повышенная концентрация гемоглобина в образце (гемолиз), избыточное содержание липидов (липемия) или избыточное содержание билирубина и его продуктов деградации (иктеричность), за счет изменения оптических свойств сыворотки крови в сравнении с нормой. Чаще всего наличие гемолиза, липемии или иктеричности в сыворотке связано либо с процедурными факторами — некорректной процедурой забора крови (например, не натощак или неаккуратным забором, приводящим к гемолизу) [6–9], некорректными условиями транспортировки образцов [10], — либо с индивидуальными проблемами конкретного пациента — неправильной диетой [11] или наличием заболеваний, приводящих к липемии и гемолизу [12].

В среднем количество образцов, непригодных для анализа, колеблется в диапазоне 3–5% [13,14]. Однако для некоторых отделений или групп пациентов количество проб с гемолизом, иктеричностью и липемией может достигать десятков процентов — например у образцов детской крови до трех лет эмпирически наблюдается самый высокий уровень гемолиза 11.6% и липемии 1.6% [15], а в отделениях неотложной помощи наблюдаемый процент непригодных для анализа образцов колеблется от 6.8 до 19.8% [16]. При этом основной причиной непригодности образцов как для амбулаторных, так и для стационарных отделений, является гемолиз — на долю гемолизированных образцов сыворотки крови приходится 40–70% всего брака [17], процент ошибочных образцов, связанных с липемией, можно оценить как 10–20% от общего количества ошибочных [15,18].

Гемолиз, липемия и иктеричность могут быть легко обнаружены визуально за счет изменения оптических свойств сыворотки крови. Так, гемоглобин обладает характерными линиями поглощения в сине-зеленой области спектра (полоса Soret 410 nm и Q-полоса в диапазоне 520–580 nm) и визуально окрашивает сыворотку крови в красный цвет. Повышенная концентрация липидов приводит к образованию липидных капель, приводящих к помутнению сыворотки крови, иктеричная сыворотка может быть обнаружена по характерному желтому оттенку, за счет поглощения света билирубином (полоса поглощения на 430–480 nm). Оценка оптических свойств сыворотки крови — обнаружение гемолиза, иктеричности и липемии может быть осуществлено визуально лаборантом, однако такой способ является субъективным и слабо скоррелирован со стандартизованными автоматическими преаналитическими методами, поэтому все чаще предпочтение отдается последним [19–21].

Автоматизированные методы преаналитической классификации образцов, в свою очередь, можно разделить на методы, требующие использования реагентов — например с использованием цианметгемоглобинового метода в оценке концентрации гемоглобина — и „нативных“ спектрофотометрических методов, не требующих дополнительной пробоподготовки [22–25]. Так, в

преаналитическом анализаторе ADVIA2400 (Siemens) используется две длины волны — 571 и 596 nm для оценки содержания гемоглобина, 658 и 694 nm для оценки степени липемии и 478 и 505 nm для оценки содержания билирубина [26].

Однако не всегда в лабораториях со средней и низкой пропускной способностью установлены спектрофотометрические преаналитические анализаторы. В таком случае может быть востребовано отдельное решение по предварительной оценке степени гемолиза и липемии в образце на преаналитическом этапе. В частности, может быть востребовано решение, сочетающее в себе возможность визуального инспектирования образца лаборантом и одновременное быстрое его измерение на наличие гемолиза, иктеричности или липемии с помощью стандартизованного сертифицированного прибора так, чтобы лаборант мог легко в руках подносить пробирку перед сканирующим устройством без дополнительного вскрытия пробирки и пробоотбора.

Поскольку гемолиз и липемия могут быть оценены визуально, для бесконтактной оценки липемии и гемолиза в ряде работ предлагается использовать методы компьютерного зрения. Так, для выявления наличия гемолиза и липемии применялись как „классические алгоритмы“ машинного обучения и компьютерного зрения [27], так и методы глубокого обучения [28,29]. Данные методы показывают высокие метрики чувствительности и специфичности в выделении степеней гемолиза и липемии [28].

В настоящей работе мы исследуем возможность преаналитической классификации образцов среди образцов нормы, гемолиза и липемии с помощью двух оптических методик: метода спектроскопии диффузного отражения, в котором анализируется сигнал отраженного от сыворотки крови отклика на различных длинах волн в видимом и ближнем ИК диапазоне, и методов компьютерного зрения (анализа изображений пробирок сыворотки крови). Помимо преимуществ, связанного с отсутствием необходимости пробоподготовки, данные методы просты в технической реализации, могут быть использованы как для полуавтоматической работы, когда оценка с помощью метода производится лаборантом для каждого образца, что подходит для работ малых лабораторий, так и в автоматическом режиме — например, при использовании метода для поточного обнаружения гемолиза и липемии на конвейерной линии в крупной клинической лаборатории. Полученные результаты говорят о перспективности обоих подходов в выявлении гемолиза и липемии в образцах.

1. Материалы и методы

1.1. Исследуемые образцы

Оценка аналитической способности методов спектроскопии диффузного отражения и компьютерного зрения в задаче преаналитической классификации сыворотки

(на „норму“, пригодную для анализа, „липемию“ и „гемолиз“) производилась на образцах венозной крови, отобранных в пробирки с объемом 4.9 ml с разделительным гелем (S-Monovette, Германия), центрифугированных для разделения эритроцитарного сгустка от сыворотки крови.

Все измерения и работа с образцами проводились в сертифицированной аккредитованной клинико-диагностической лаборатории (ИСО 15189:2012) городской клинической больницы № 67 им. Л.А. Ворохобова (Москва). Исследования были одобрены локальным этическим комитетом ГKB № 67.

Для оценки разрабатываемых методик среди всех проб отбирались образцы сыворотки крови с концентрацией гемоглобина в сыворотке более 0.5 g/L, что соответствует наиболее „мягкой“ отсечке по уровню гемоглобина при наличии гемолиза (включает все полуколичественные категории, часто используемые для характеристики гемолиза „+“ — „++++“ [15]), и образцы с липемией с концентрацией липидов выше 1.25 g/L, что также соответствует „мягкой“ отсечке по липемии (включает все полуколичественные категории наличия липемии „+“ — „++++“ [15]), образцы нормы выбирались так, чтобы концентрация липидов и гемоглобина составляла менее 1.25 и 0.5 g/L соответственно. Верификация концентрации гемоглобина и липемии осуществлялась по данным независимого спектрофотометрического исследования с помощью измерения оптической плотности сыворотки с помощью спектрофотометра Perkin-Elmer-25 в диапазоне 300–800 nm. Для этого не более 1 ml супернатанта центрифугированной крови отбиралось в кварцевую кювету с оптическим путем 2 mm, и измерялась оптическая плотность образца. Концентрация гемоглобина определялась по линии поглощения оксигемоглобина на 572 nm, концентрация липидов производилась по длине волны 660 nm в сравнении с оптической плотностью липофундина с концентрацией 1 g/L. Измерения с помощью методов спектроскопии диффузного отражения и компьютерного зрения производились в исходных пробирках без дополнительного пробоотбора в исходных пробирках, использованных для забора крови.

Всего для исследования с помощью спектроскопии диффузного отражения, а также с помощью алгоритмов компьютерного зрения было отобрано 113 образцов с центрифугированной кровью: 33 образца с сывороткой, пригодной для дальнейшего биохимического анализа („норма“), 59 образцов с различной степенью липемии („липемия“), 21 образец с различной степенью гемолиза („гемолиз“). Образцы с центрифугированной кровью хранились в холодильной камере при температуре 4°C. Измерения спектров диффузно рассеянного света и получение изображений для применения методов компьютерного зрения в мануальном режиме проводились в течение 72 h после отбора крови. Это позволяло говорить о высокой стабильности образцов цельной крови и возможности дальнейшего гематологического анализа [30].

Для тестирования алгоритмов компьютерного зрения в „поточном“ режиме была собрана отдельная статистика по 150 образцам различных классов, о чем будет указано отдельно ниже.

1.2. Измерения спектров диффузного отражения

Для измерения диффузного отражения сыворотки крови, находящейся в пробирке, была разработана специальная двухволоконная схема детектирования диффузного отражения. В ней излучение от галогеновой лампы мощностью 15 W (длины волн излучения 450–2000 nm) с помощью системы двух выпуклых линз коллимировалось на торец осветительного многомодового кремниевое волокна (IPG Photonics, Россия, диаметр сердцевины 550 μm , числовая апертура 0.22), другой торец которого прикладывался к области измерения на исследуемом образце. Второе идентичное оптическое волокно одним из торцов располагалось на расстоянии $d = 0.2 \text{ mm}$ от края осветительного волокна и использовалось для сбора диффузно-отраженного сигнала, который далее передавался на щель спектрометра (YSM-8101, Yixist, КНР), детектирующего излучение в спектральном диапазоне 200–1100 nm со спектральным разрешением $\sim 10 \text{ nm}$.

Детектирование сигнала диффузного отражения от образцов сыворотки крови производилось следующим образом: торцы осветительного и собирающего волокон прижимались перпендикулярно стенке пластиковой пробирки образца напротив области, соответствующей сыворотки крови, после чего производилась запись спектра отраженного сигнала.

Для расчета спектра диффузного отражения $R(\lambda)$, помимо интенсивности сигнала от образца сыворотки крови $I(\lambda)$, измерялся отклик от референсного белого образца, в качестве которого использовался стандарт Spectralon (Labsphere, США) с коэффициентом отражения 99% в видимом и инфракрасном диапазонах $I_{\text{ref}}(\lambda)$, а также фоновый сигнал $I_{bg}(\lambda)$, учитывающий шум детектора. Далее коэффициент диффузного отражения на различных длинах волн рассчитывался по формуле

$$R(\lambda) = \frac{I(\lambda) - I_{bg}(\lambda)}{I_{\text{ref}}(\lambda) - I_{bg}(\lambda)}, \quad (1)$$

а затем диффузное отражение преобразовывалось в эффективную оптическую плотность $\text{OD}(\lambda)$, которая далее использовалась для характеристики образцов:

$$\text{OD}(\lambda) = -\ln R(\lambda). \quad (2)$$

Поскольку интенсивность детектируемого сигнала в диапазоне менее 500 nm была существенно ниже в сравнении с сигналом в диапазоне 500–1000 nm, для дальнейшего анализа использовался сигнал диффузного отражения в спектральном диапазоне 500–1000 nm.

1.3. Фотосъемка образцов в стандартизованных условиях с помощью камеры

Для первичной оценки аналитической способности методов компьютерного зрения в классификации образцов центрифугированной крови по классам „норма“, „гемолиз“ и „липемия“ производилась фотосъемка образцов в фиксированных условиях освещения.

Для этого образец с центрифугированной венозной кровью размещался в специализированном штативе, расположенном на расстоянии ~ 20 см от цифровой фотокамеры, снимающей RGB-изображения с цифровым разрешением 1920×1080 пикселей (Logitech C922 Pro, Китай). Фотографии осуществлялись в специализированном коробе, изолированном от внешней засветки с включенным LED-источником белого света с освещенностью 4880 lm и цветовой температурой 6220 K. Образец размещался в штативе так, чтобы в поле зрения камеры попадала область, соответствующая сыворотке крови, для возможности оценки ее цветовых характеристик.

1.4. Фотосъемка образцов в режиме работы на конвейерной линии

В рамках работы нами была разработана установка для преаналитической классификации образцов с помощью методов компьютерного зрения в высокопоточном режиме в случае, когда образцы сыворотки крови едут по специализированному конвейеру в клинико-диагностической лаборатории.

Установка состояла из четырех цифровых камер, снимающих изображение в разрешении 1920×1080 пикселей (Logitech C922 Pro, КНР), расположенных на расстоянии ~ 30 см друг от друга под углом 90° , синхронно получающих изображения для одной и той же геометрической области с различных сторон с частотой получения кадров 15 Hz. Над креплениями для камер располагался LED-источник белого освещения (4880 lm, цветовой температурой 6220 K) для устранения неоднородной засветки области. Изображения, получаемые с фотокамер, попадали на управляющий компьютер, на котором производилась запись и обработка изображений.

Камеры разработанного устройства располагались так, что одна из них была сонаправлена движению образца по конвейерной линии, вторая — противонаправлена, а две другие камеры располагались перпендикулярно движению образца по конвейерной линии.

Всего с помощью указанного прототипа прибора была собрана статистика по 150 образцам каждого класса — „норма“, „гемолиз“ и „липемия“. Для обучения модели отобрано по 740 изображений, на которых присутствовала пробирка каждого класса, а также для увеличения робастности модели добавлено 76 изображений, на которых пробирка отсутствовала.

1.5. Разметка изображений для классификации и детектирования пробирок

Для разработки алгоритма детектирования и классификации пробирок для отобранных изображений производилась ручная разметка изображений на платформе Supervise.ly [31]. Процесс ручной разметки данных в supervise.ly представлял собой выделение необходимой части пробирки на кадре в рамку. В данном случае рамкой выступал прямоугольник от конца держателя пробирки до верхнего уровня сыворотки крови. Стороны рамки плотно прилегали к границам объекта. Результатом разметки являлся набор данных, содержащий информацию о названии фотографии, ее размерах, классе и координатах границ объекта на детектируемом изображении.

В случае, если область изображения, связанная с сывороткой крови, была закрыта штрих-кодом, наклеенным на образце, выделялась только область, содержащая штрих-код, а образцу в соответствие ставился класс „область скрыта“.

1.6. Процедура построения классификаторов и оценка точности классификаторов

Для оценки аналитической способности методов спектроскопии диффузного отражения и компьютерного зрения в режиме с фиксированными условиями и в „поточном“ режиме при перемещении пробирки по конвейеру был разработан ряд классифицирующих моделей, предсказывающих принадлежность образца классу „норма“, „гемолиз“ и „липемия“. Построение моделей и их анализ производился с помощью специально разработанных скриптов на языке Python 3 с использованием библиотек NumPy, Pandas, Matplotlib, PyTorch, Ultralytics, Scikit-Learn [32–37]. Подробнее приведем описание процедур построения и оценки качества классификации каждого из алгоритмов.

1.6.1. Построение и оценка качества классификации по данным спектроскопии диффузного отражения.

Обучение модели классификации образцов на классы „норма“, „липемия“ и „гемолиз“ производилось на 70% от общего числа наблюдений (спектров оптической плотности), остальные 30% измерений использовались для оценки точности (для тестирования) модели, при этом для получения погрешностей метрик качества классификации генерировалось 30 случайных разбиений базы данных на обучающую и тестовую выборки.

В качестве признаков, используемых для классификации образцов, были использованы амплитуды разложения спектров оптической плотности (OD) на три первые главные компоненты, построенные по данным разложения спектров из тренировочной выборки. Для

классификации предобработанных рассматривались алгоритмы случайного леса с числом деревьев, равным 100, без ограничения по глубине и минимальному числу объектов в листе дерева, а также метод опорных векторов с мягким зазором и гауссовым ядром с коэффициентом, подбираемым на основе точности на валидационной выборке. Была использована имплементация метода главных компонент, случайного леса и опорных векторов из библиотеки scikit-learn [37].

1.6.2. Построение и оценка качества классификации с помощью компьютерного зрения при фиксированных положениях образца. На изображениях, полученных в стандартизованных условиях освещения при фиксированных положениях пробирки, выделялись области, соответствующие сыворотке крови на изображении. Для указанных областей были рассчитаны средние значения интенсивности в R (красном), G (зеленом) и B (синем) цветовых каналах цифрового изображения, фиксирующих интенсивности излучения в диапазонах соответственно 600–680, 520–600 и 430–500 nm, и выступавших далее в качестве предикторов для предсказания класса пробирки.

Для классификации образцов аналогично пункту 2.6.1 рассматривались методы случайного леса и опорных векторов. Эмпирически было выявлено, что лучшей конфигурацией обладает модель на основе метода опорных векторов с мягким зазором, использующая гауссово ядро (RBF) с коэффициентом 0.33. Для оценки метрик использовался метод кросс-валидации. Тренировочный и валидационный набор разбивались в пропорциях 70% к 30% соответственно. Процедура обучения повторялась 100 раз с предварительным перемешиванием и отбором новых данных для тренировочного и валидационного наборов.

1.6.3. Оценка качества детектирования модели компьютерного зрения для классификации образцов на конвейерной линии. Для работы в высокопоточном режиме в случае движения образца по конвейерной линии требовалась разработка модели, которая осуществляет поиск объекта на изображении (выделяет область, соответствующую объекту на изображении) и его классификацию, т.е. решает задачу детектирования. Для решения данной задачи использовалась модель YOLO v5 [36].

Качество поиска объекта на изображении оценивалось с помощью метрики средней точности (mAP) при пересечении истинных и предсказанных областей объекта более 50% (intersection over union > 50%). Эта метрика отражает долю корректно распознанных объектов к общему числу срабатываний детектора в случае, если площадь пересечения истинных областей и областей предсказаний составляет более 50%.

Помимо классов „норма“, „гемолиз“ и „липемия“ в модель в качестве классов были добавлены классы

„область скрыта“, в том случае если область сыворотки крови закрыта штрих-кодом, наклеенным на стенке пробирки, а также „фон“, в том случае если объект отсутствовал на изображении.

Для обучения использовалось 80% выборки, для ее валидации 20%. Архитектура YOLO v5 была инициализирована с весами, предобученными на COCO [38] датасете для улучшения сходимости, после чего дообучена на тренировочном наборе данных с помощью метода стохастического градиентного спуска. Перед применением модели изображения дополнительно преобразовывались до размера 980 × 980 пикселей, модель дообучалась на данных 50 эпох с размером пакета, равным 2.

1.6.4. Расчет метрик качества классификации.

Классификаторы характеризовались с помощью ROC-кривых зависимости чувствительности и специфичности классификатора от уровня порогового значения присвоения положительной метки класса. Чувствительность и специфичность идентификации одновременно гемолиза и липемии рассчитывались следующим образом: образцам с гемолизом и липемией приписывался положительный класс, а образцам из класса „норма“ приписывался отрицательный класс, после чего чувствительность и специфичность оценивались стандартным для задачи бинарной классификации образом. Расчет метрик производился с помощью встроенных реализаций библиотек scikit-learn [37].

2. Результаты

2.1. Оценка спектроскопии диффузного отражения в преаналитической классификации сыворотки крови

В первую очередь была исследована возможность использования спектроскопии диффузного отражения в задаче идентификации гемолиза и липемии в образцах сыворотки крови в схеме на отражение — таким образом, что можно проводить диагностику без дополнительного вскрытия пробирки, исключительно по диффузно отраженному сигналу. Для этого была использована схема детектирования, в которой два оптических волокна (диаметр сердцевины 550 μm), расположенных на расстоянии друг от друга 0.2 mm, подводились к визуально наблюдаемому участку сыворотки крови. При этом по одному из волокон заводилось излучение широкополосного источника излучения (галогенная лампа, излучение в диапазоне 450–2000 nm), а отраженный сигнал затем собирался другим волокном, совмещенным со спектрометром, детектирующим отклик в диапазоне 200–1000 nm (см. описание установки в разд. 2.2). На рис. 1,а продемонстрирована принципиальная схема измерений спектров диффузного отражения от участка сыворотки крови с помощью двухволоконной схемы.

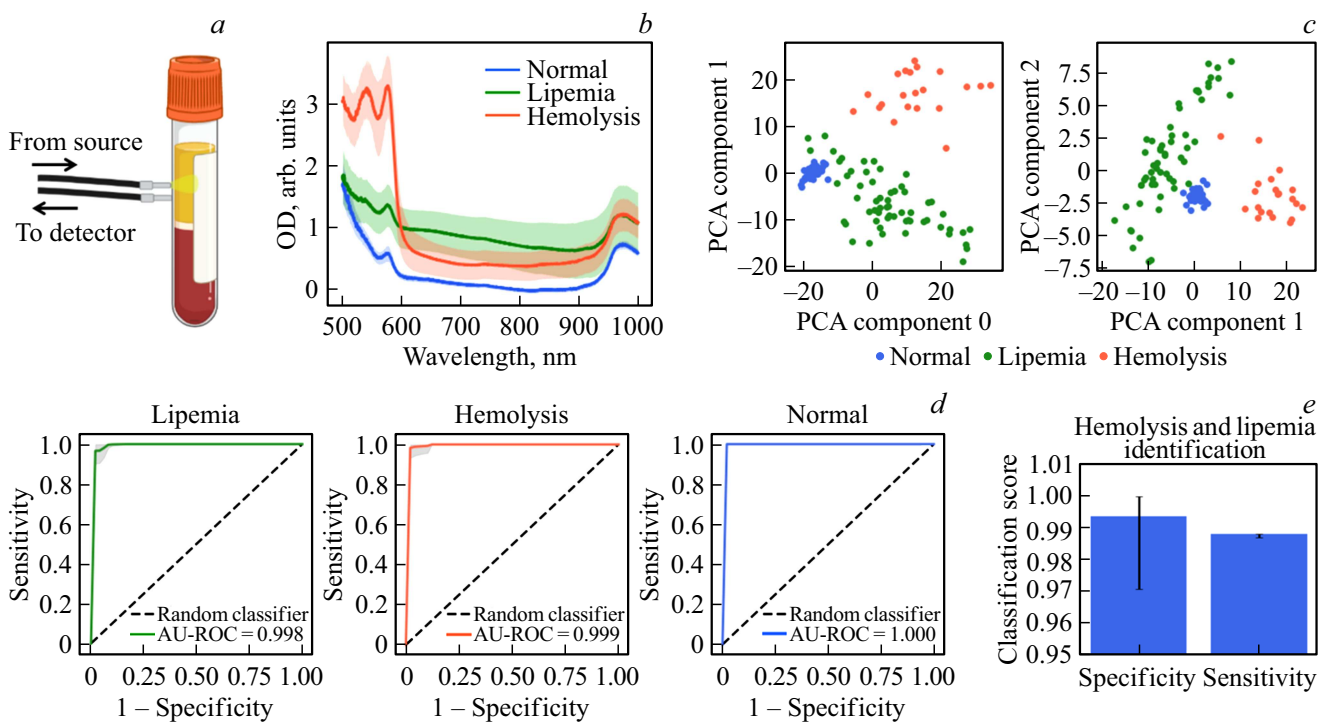


Рис. 1. (a) Схематическое изображение способа измерения диффузного отражения с помощью двухволоконной схемы от образца с центрифугированной сыворотки крови. (b) Среднее и внутригрупповое среднееквдратическое отклонение для спектров эффективной оптической плотности (OD) сыворотки крови с выраженной липемией („липемия“), образцов с гемолизом и образцов, пригодных для биохимического анализа („норма“). (c) Результат применения метода главных компонент (МГК) для понижения размерности спектров оптической плотности образцов сыворотки крови (отражены амплитуды первых трех главных компонент). (d) ROC-кривые для классификации липемии, гемолиза и нормы, оцененные с помощью кросс-валидации. (e) Значения специфичности и чувствительности в идентификации липемии и гемолиза в образцах с помощью построенного классификатора по данным кросс-валидации.

Для оценки классификационной чувствительности такого подхода нами были отобраны образцы с наличием гемолиза и липемии: минимальная концентрация гемоглобина в образцах составляла 0.7 g/L, минимальная концентрация липидов в образцах с липемией 1.3 g/L, что соответствует низкому уровню гемолиза и липемии в образцах сыворотки крови, согласно рекомендациям преаналитической лабораторной диагностики.

Для всех образцов был измерен спектр диффузного отражения, из которого рассчитывалась эффективная оптическая плотность (optical density, OD) как отрицательный логарифм коэффициента отражения (формула (2)). На рис. 1, b представлены усредненные внутри класса средние спектры эффективной оптической плотности для образцов нормы, липемии и гемолиза. Как видно, образцы нормы обладали низкими значениями OD в диапазоне 600–900 nm в отличие от образцов с липемией, для которых абсолютные значения OD в ближнем ИК диапазоне были существенно выше. Такое поведение оптической плотности связано с повышенным коэффициентом рассеяния образцов с выраженной липемией за счет наличия жировых капель в растворе (рис. 1, b). В образцах с гемолизом при этом не наблюдалось высоких значений оптической плотности в ИК диапазоне,

однако в видимой области наблюдались линии, характерные для поглощения оксигемоглобина в диапазоне 500–600 nm. Детектирование сигнала диффузного отражения на длинах волн менее 500 nm было ограничено низкой интенсивностью источника излучения в данной спектральной области и высоким поглощением гемоглобина для ряда образцов. Подобные особенности в спектрах эффективной оптической плотности позволяют легко разделить образцы по классам гемолиза, липемии и нормы. Нами было обнаружено, что при применении метода главных компонент (МГК) к спектрам оптической плотности в плоскостях, соответствующих первым трем компонентам, образцы различных классов были хорошо разделимы (рис. 1, c).

Для количественной оценки точности классификации образцов на норму, липемию и гемолиз нами был построен простейший алгоритм классификации, который принимал на вход спектры эффективной оптической плотности в диапазоне 500–1000 nm, которые затем преобразовывались с помощью МГК так, что вместо исходных спектральных данных в качестве признакового описания образцов использовались амплитуды первых трех главных компонент спектров оптической плотности. После понижения размерности для классификации

образцов по их спектральным признакам (амплитудам трех главных компонент) использовались модель случайного леса, либо модель на основе метода опорных векторов. Наилучшей по результатам оценки точности на валидационной выборке показала себя модель на основе случайного леса, поэтому далее будут представлены результаты для указанного алгоритма.

Точность классификации оценивалась стандартным способом: 70% объектов из всей выборки использовались нами для обучения модели (как подбора главных компонент, так и для обучения алгоритма классификации), а на 30% оценивалась точность работы классификатора. Причем для оценки погрешности метрик классификации использовалось 30 случайных разбиений данных на тренировочную и тестовую выборки.

Распространенными метриками, которые часто используются для решения задачи классификации, являются метрики чувствительности и специфичности — доли корректно распознанных истинно положительных образцов и доли корректно распознанных истинно отрицательных образцов. Однако значения чувствительности и специфичности могут зависеть от порогового значения, по которому вероятность принадлежности к конкретному классу переводится в значение предсказанной метки класса. В связи с этим прибегают к оценке классификатора с помощью ROC-кривой — зависимости чувствительности классификатора от специфичности при всех значениях указанного порога классификации. При этом в качестве „интегральной оценки“ классификатора может выступать площадь под ROC-кривой (AUC-ROC). Площадь под ROC-кривой по сути можно интерпретировать как среднее значение чувствительности классификации при всех уровнях специфичности. На рис. 1, *d* представлены ROC-кривые в задаче идентификации нормы, липемии и гемолиза — как видно, точность классификации для всех классов близка к единице.

Мы также оценили чувствительность и специфичность в обнаружении гемолиза и липемии среди всех образцов (рис. 1, *e*). Было установлено, что указанный классификатор может обладать чувствительностью, равной 99% при специфичности более 99.4% (по данным кросс-валидации). Полученные результаты говорят о высокой перспективности используемого метода в идентификации гемолиза и липемии в сыворотке крови. Предположительно этот подход также может быть использован для количественной и полуколичественной характеристики образцов: определения диапазона концентраций гемоглобина и липидов в сыворотке крови.

Так как было обнаружено, что для классификации достаточно трех параметров диффузного отражения (амплитуд трех главных компонент), мы решили исследовать возможность классификации образцов по данным их отражения с помощью алгоритмов компьютерного зрения. В первую очередь были проведены эксперименты в „идеальных условиях“, на изображениях, где пробирка была зафиксирована и находилась в постоянных условиях освещенности для оценки точности

классификации, а затем исследована точность модели компьютерного зрения „в реальных условиях“ — в задаче классификации образцов, движущихся по конвейерной линии в клинично-диагностической лаборатории.

2.2. Оценка точности преаналитической классификации образцов с помощью компьютерного зрения: модельные условия

Для оценки точности преаналитической классификации пробирок для образцов центрифугированной сыворотки крови, измеренной с помощью спектроскопии диффузного отражения (разд. 2.1), было предложено исследовать их цветовой контраст путем получения изображений при фиксированных условиях освещенности. Помимо фиксированных условий освещенности на полученных изображениях пробирка всегда располагалась в одной и той же области изображения, что существенно облегчало выделение области, соответствующей сыворотке крови. Подобные эксперименты были нацелены на то, чтобы получить „верхний предел“ точности преаналитической классификации образцов с помощью компьютерного зрения по цветовым характеристикам сыворотки крови, в условиях, когда все внешние условия — в первую очередь засветка и положение пробирки — зафиксированы. Пример получаемого изображения с пробиркой центрифугированной крови вместе с примером разметки (рис. 2, *a*).

На рис. 2, *b* представлены примеры выделенных областей сыворотки крови различных классов („норма“, „липемия“, и „гемолиз“). После описанной процедуры сегментации областей сыворотки крови для них рассчитывались средние значения интенсивностей в синем (B), зеленом (G) и красном (R) каналах изображения. На рис. 2, *c* показаны диаграммы рассеяния для средних значений интенсивности в различных цветовых каналах для областей сыворотки крови. Как видно, области точек, соответствующие различным классам (норма, гемолиз и липемия), не перекрываются — следовательно, можно ожидать высокой точности преаналитической классификации по средним значениям интенсивностей в R, G и B каналах изображения.

Далее нами были построены несколько алгоритмов классификации, на основе случайного леса и метода опорных векторов, принимающих в качестве признаков средние значения интенсивности в каналах R, G и B, предсказывающих принадлежность образца к классу „норма“, „липемия“ и „гемолиз“. Эмпирически было установлено, что наилучшей по точности на валидационной выборке оказалась модель на основе метода машин опорных векторов с мягким зазором и гауссовым ядром. Для этого классификатора нами были оценены чувствительность и специфичность классификации для всех уровней порога классификации (ROC-кривые, рис. 2, *d*). Значения площади под ROC-кривой для трех классов варьировались в диапазоне 0.96–0.99 (значение

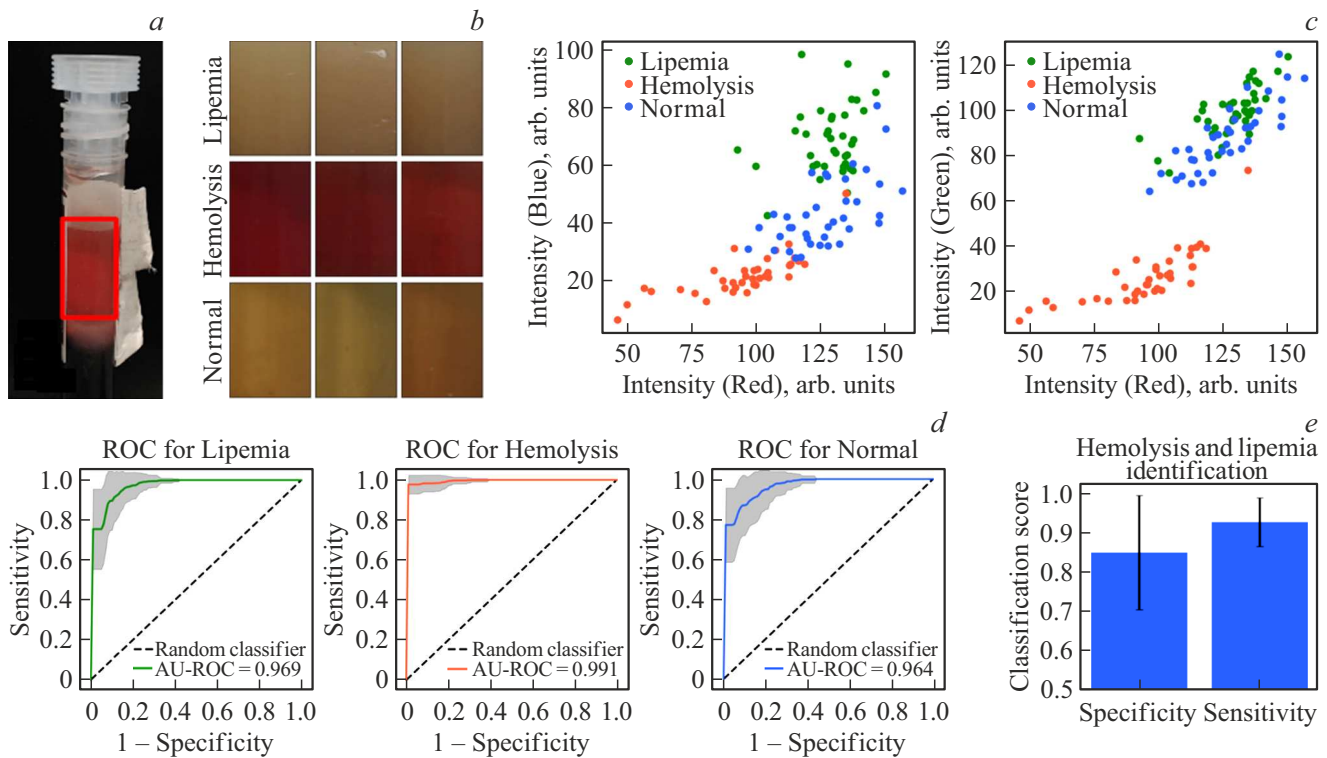


Рис. 2. (a) Пример изображения пробирки с сывороткой крови с выраженным гемолизом. Красным прямоугольником выделена область, соответствующая сыворотке крови. (b) Примеры выделенных областей сыворотки крови для образцов, принадлежащих к различным классам (норма, липемия, гемолиз). (c) Диаграммы рассеяния для средних значений интенсивности сигнала в R, G и B каналах изображения для сегментированных участков сыворотки крови образцов. (d) ROC-кривые классификации липемии, гемолиза и нормы по данным RGB значений сыворотки крови с помощью классификатора на основе метода опорных векторов. (e) Специфичность и чувствительность идентификации гемолиза или липемии. Порог классификации выбран так, чтобы максимизировать чувствительность классификации.

площади под ROC-кривой равно единице соответствует „идеальному“ классификатору). Также нами была оценена специфичность и чувствительность обнаружения гемолиза и липемии с помощью цветовых характеристик сыворотки крови: максимальная чувствительность составила $90.9 \pm 7.5\%$ при специфичности обнаружения гемолиза и липемии $84 \pm 15\%$. Как видно, на различных разбиениях наблюдается существенный разброс по чувствительности и специфичности — это может быть связано с наличием „выбросов“ в обучающей выборке в некоторых разбиениях. При этом в среднем значения чувствительности и специфичности преаналитической классификации с помощью компьютерного зрения также высоки, однако уступают значениям метрик, получаемых для метода спектроскопии диффузного отражения.

2.3. Оценка точности преаналитической классификации образцов с помощью компьютерного зрения: тестирование в поточном режиме

Преимуществом съемки и анализа изображений является то, что требования к положению объекта относительно фотокамеры могут быть не столь значительными,

как в других методах, в частности спектроскопии диффузного отражения, когда для получения качественных результатов зонд должен подводиться строго к участку с анализируемой сывороткой крови. Благодаря этому методы компьютерного зрения могут быть применены в сложных случаях, например, когда пробирка располагается в различных местах на изображении или движется относительно фотокамеры. В частности, такая постановка задачи может быть актуальной для крупных лабораторий, в которых образец движется по конвейерной ленте: располагая преаналитический анализатор на конвейерной линии, можно дополнительно проверять образцы на наличие гемолиза и липемии и не проводить анализ для непригодных образцов.

Для применения подходов компьютерного зрения в такой постановке задачи нами была разработана установка, которая располагалась на конвейерной линии. Установка состояла из четырех цифровых фотокамер, расположенных на расстоянии ~ 30 см друг от друга под углом 90° , синхронно снимающих одну и ту же геометрическую область с различных сторон с частотой кадров 20 fr/s. Наличие четырех камер было необходимо из-за того, что часть пробирки с образцом может быть закрыта наклейкой с идентификационным штрих-кодом образ-

ца, скрывающим „полезную“ часть сыворотки крови в области зрения одной камеры. Камеры располагались непосредственно в близости от дополнительного LED-источника белого освещения, фиксирующего условия засветки образцов, движущихся по конвейерной линии. Изображения с фотокамер попадали на управляющий компьютер, который производил обработку. Фотография разработанного модуля при установке на участке конвейерной линии приведена на рис. 3, *a*. Камеры разработанного устройства располагались так, что две из камер смотрели в сторону направления движения пробирки и в противоположную сторону, а две другие камеры располагались перпендикулярно движению образца по конвейерной линии. Примеры изображений, фиксируемых с помощью камеры, направленной по движению конвейера и перпендикулярно движению конвейера, показаны на рис. 3, *b*. С помощью указанного прототипа прибора была набрана выборка с образцами центрифугированной крови.

Для образцов сыворотки крови детектировались изображения с четырех фотокамер, после чего отбирались несколько кадров для каждой пробирки, которые были использованы для создания моделей компьютерного зрения для преаналитической классификации пробирок. Всего для обучения разработанного алгоритма было собрано 3 776 изображений с различных углов зрения — набор данных был сбалансирован так, чтобы в обучающей выборке присутствовало 185 изображений каждого класса (норма, липемия и гемолиз) с четырех ракурсов, 185 изображений пробирок, на которых область сыворотки крови скрыта идентификационным штрих-кодом (дополнительный класс „область скрыта“), и дополнительно 76 фотографий без пробирок для улучшения точности работы модели компьютерного зрения. Для обучения модели собранная база данных изображений пробирок была размечена: выделены области изображения, которую занимает пробирка (прямоугольники на рис. 3, *b*), а также отмечен класс пробирки. Примеры сегментированных областей, принадлежащих различным классам, представлены на рис. 3, *c*.

Подробнее опишем процедуру разработки модели компьютерного зрения. В имеющейся конфигурации детектирования пробирка с образцом могла располагаться в произвольном месте изображения, поэтому для обнаружения и преаналитической классификации пробирки требовалась разработка дополнительных моделей „ищущих“ и классифицирующих пробирку на изображении. Эта задача могла быть решена в два этапа: сначала поиск объекта (пробирки) на изображении, а затем ее анализ, например анализ цвета сыворотки крови, как это было продемонстрировано выше в разд. 3.2. Однако более робастным подходом, хорошо зарекомендовавшим себя в подобных задачах, является подход с одностадийным детектированием, когда для поиска и классификации объекта на изображении используется одна модель. Нами был выбран последний — одностадийный подход к

детектированию объекта, реализуемый с помощью сверточной нейронной сети с архитектурой YOLOv5 [36], распознающей указанные 4 класса („норма“, „гемолиз“, „липемия“ и „область скрыта“) и указывающей положения объектов в виде прямоугольников на изображении. При этом изображения, соответствующие одной и той же пробирке, не попадали одновременно в тренировочную и тестовую выборки. Для валидации использовалось 20% изображений, получаемых с конвейерной линии, остальные 80% использовались для обучения.

Качество обнаружения объекта на изображении вычислялось с помощью часто используемой метрики средней точности классификации („mean Average Precision“, mAP) [39]. Эта метрика отражает долю корректно распознанных объектов к общему числу срабатываний детектора в правильной области. Оценка модели для области пересечения истинных областей и областей предсказаний равной 50% (IoU = 0.5) составляла $mAP_{IoU=0.5} = 0.984$. Таким образом, можно сделать вывод, что модель достаточно точно определяла местоположение пробирки на фотографиях.

Для оценки точности классификации параметров модели мы следили за средними значениями точности и чувствительности классификатора. Было установлено, что среднее значение точности классификации составляет 0.964, при уровне чувствительности определения класса 0.972. На рис. 3, *d* показана матрица ошибок классификатора, в которой отмечено распределение истинных меток классов по классам, предсказанным с помощью построенной модели. При этом пороги предсказаний настраивались таким образом, чтобы наиболее чувствительно обнаруживать гемолиз и липемии в образцах.

Из матрицы ошибок видно, что при такой настройке алгоритма модель с единичной точностью идентифицирует наличие гемолиза в образцах, с точностью 0.94 идентифицирует наличие липемии, при этом иногда (в 23% случаев) модель может ложноположительно идентифицировать гемолиз в образцах нормы. В случае, когда область сыворотки крови не видна на изображении (класс „область скрыта“), никогда не распознается как один из классов гемолиз, липемия или норма. В случаях же, когда на разметке не была выделена пробирка (класс „фон“), модель одинаково относилась объекты ко всем классам, однако такие изображения, согласно метрике $mAP = 0.984$, встречались менее чем в 2% случаев и были связаны с наличием пробирок на изображении, находящихся на существенном удалении от камеры.

Мы также оценили общую специфичность и чувствительность в идентификации гемолиза и липемии среди тех образцов, для которых была видна сыворотка крови (т.е. только среди трех классов (нормы, липемии и гемолиза) — чувствительность идентификации гемолиза и липемии составила 97% при специфичности на уровне 74% (рис. 3, *e*). Низкие уровни специфичности в первую очередь связаны с тем, что алгоритм ошибочно приписывал класс „гемолиз“ образцам с нормальной

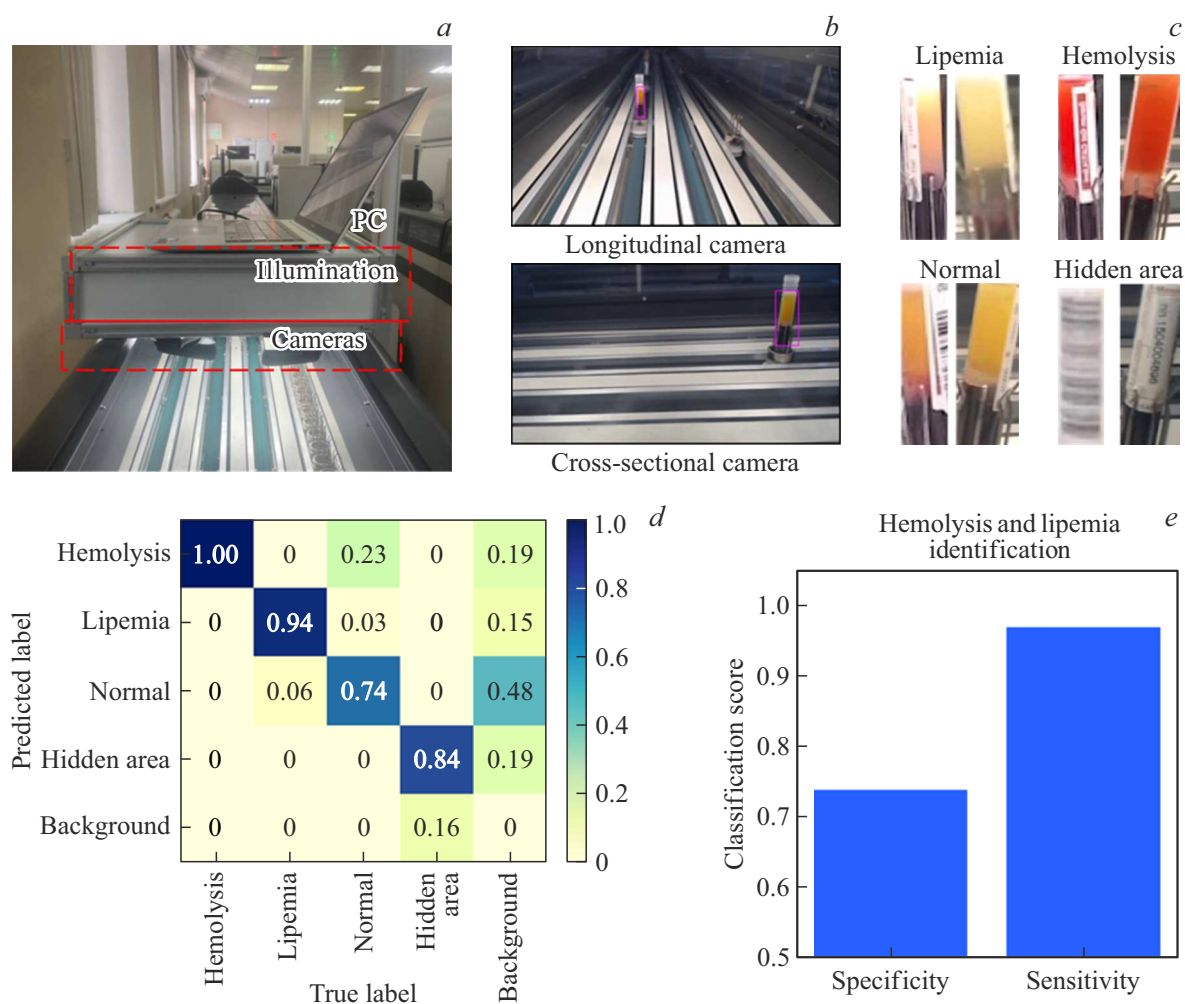


Рис. 3. (а) Фото разработанного прототипа установки, осуществляющего фотосъемку образцов на конвейерной линии клинко-диагностической лаборатории. (б) Примеры изображений, получаемых с камер устройства, направленных вдоль направления движения пробирки (сверху) и перпендикулярно направлению движения пробирки (снизу). (с) Примеры сегментированных областей изображений с различными классами („норма“, „гемолиз“, „липемия“ и „область скрыта“). (d) Матрица ошибок для задачи преаналитической классификации с помощью построенной модели компьютерного зрения. Числом отражена доля объектов с соответствующим истинным значением класса. (е) Специфичность и чувствительность в идентификации гемолиза и липемии среди образцов.

сывороткой крови. Дальнейшее улучшение алгоритма может быть проведено за счет увеличения статистики наблюдений, оптимизации параметров алгоритма, а также расширения обучающей выборки с помощью изображений с различным положением камеры при различных условиях освещенности.

3. Обсуждение результатов

В настоящей работе была проведена оценка двух методов преаналитической классификации сыворотки венозной крови — метода спектроскопии диффузного отражения и метода компьютерного зрения — в задаче классификации образцов на классы „норма“, „липемия“ и „гемолиз“. Преимуществом этих методик является их бесконтактность — измерения могут быть проведены

напрямую через пластиковый корпус пробирки, в случае если хорошо различима область, соответствующая сыворотке крови.

Наиболее точным среди представленных методик, оказалась спектроскопия диффузного отражения. На данных 113 образцов сыворотки крови с различными уровнями гемолиза и липемии, а также образцам нормы была получена высокая чувствительность более 99% при среднем уровне специфичности 99.4% классификатора, использующего в качестве признаков спектры эффективной оптической плотности (отрицательного логарифма диффузного отражения) в диапазоне 500–1000 нм. При этом укажем, что для предварительного понижения размерности исходных спектральных данных был использован метод главных компонент, преобразующих исходные спектры оптической плотности в амплитуды первых трех

главных компонент — т.е. эффективно, построенный классификатор использовал всего три „спектральных“ признака в указанном диапазоне длин волн для классификации образцов. Следовательно, в перспективе вместо детектирования диффузного отражения в широком спектральном диапазоне может быть выделено небольшое количество длин волн (3–4 спектральных полосы), на которых необходимо вести измерение коэффициента диффузного отражения для успешной классификации образцов, что устраняет необходимость использования спектрометра в данной измерительной схеме. Дополнительное использование более коротковолнового диапазона (400–500 nm) может повысить устойчивость метода в отношении определения образцов с повышенной концентрацией гемоглобина, а также в выявлении повышенной концентрации билирубина по его полосе поглощения с положением максимума на ~ 460 nm. Также обратим внимание, что разделение образцов на „норму“, „липемия“ и „гемолиз“ производилось нами с использованием низких порогов концентрации липидов (1.25 g/L) и гемоглобина (0.5 g/L) в сыворотке крови — с использованием нескольких порогов по концентрации липидов и гемоглобина возможно решение задачи полуквалиметрической характеристики степени липемии и гемолиза в образце.

Однако измерение с помощью предложенной волоконной методики измерения спектров диффузного отражения требует подведения зонда непосредственно к области сыворотки крови образца, что является сложным для автоматизации. Исходя из этого, по мнению авторов, нишевым применением данного метода является его использование в малых лабораториях, где автоматизированный преаналитический этап отсутствует, и верификация образцов производится вручную. Применения метода спектроскопии диффузного отражения в таком случае устраняет субъективность и человеческий фактор в оценке свойств сыворотки крови.

Далее, в работе нами была оценена точность преаналитической классификации образцов сыворотки крови с помощью методов компьютерного зрения. В разд. 3.2 представлены результаты возможности классификации образцов на классы „норма“, „липемия“ и „гемолиз“ когда положение пробирки и уровень освещенности пробирки полностью зафиксированы, а в разд. 3.3 представлены результаты тестирования более сложной модели, осуществляющей одновременно поиск образца и его классификацию, когда пробирка движется по конвейерной линии в медицинской лаборатории.

Было установлено, что в первом случае (фиксированное положение образца и условий освещения) точную классификацию (чувствительность обнаружения гемолиза и липемии $91 \pm 7.5\%$ при специфичности $84 \pm 15\%$) можно провести по средним значениям интенсивностей в каналах R, G, B для областей изображения, соответствующих сыворотке крови. При этом ошибочные ложноположительные срабатывания (классификация образца из класса „норма“ как образца с гемолизом или

липемией) наблюдались одинаково и в пользу класса „гемолиз“, и в пользу класса „липемия“. Устранение ложноположительных срабатываний может быть настроено с помощью выбора порога вероятности, по которому классификатор относит образец к тому или иному классу.

Для детектирования и классификации образцов в более сложном случае, когда образцы движутся по конвейерной линии, нами была разработана более сложная модель на основе нейронной сети с архитектурой YOLOv5. Было установлено, что такая модель имеет низкую долю ложноположительных срабатываний при поиске образцов на изображении ($mAP50 = 0.984$, что говорит о числе ложноположительных срабатываний на уровне 1.6%). Для обнаруженных на изображении образцов, в случае если для пробирки была видна область, соответствующая сыворотке крови, максимальная чувствительность идентификации гемолиза и липемии составляла 97% при специфичности 74%. При этом наблюдалось, что основной процент ложноположительных срабатываний происходил из-за отнесения образцов нормы к классу „гемолиз“. Предположительно, такая проблема может быть связана с менее однородным уровнем освещенности поля, в котором происходит детектирование пробирок в случае их перемещения по конвейерной линии, влияющего на цветопередачу оттенка сыворотки крови.

Последняя гипотеза, в частности, подтверждается тем, что модель оказывалась чувствительна к уровню освещенности полей зрения камер и поля зрения камеры — в случае изменения указанных параметров, точность обнаружения объектов и их классификации снижались, поэтому итоговые данные собирались в одной измерительной конфигурации. Устойчивость модели к различным уровням освещенности может быть повышена за счет дополнительного дообучения модели на изображениях, полученных в условиях различной освещенности и для различного положения камер, однако это требует пропорционально большего числа изображений в обучающем наборе данных. Несмотря на указанные недостатки, подход на основе компьютерного зрения потенциально применим в преаналитической классификации сыворотки крови, в частности, нами продемонстрирована возможность применения данного метода в идентификации образцов с гемолизом и липемией на конвейерной линии в медицинской лаборатории.

Заключение

Преаналитический отбор образцов сыворотки крови на пригодность для дальнейшего проведения биохимических анализов — критический шаг в правильной постановке диагноза. В настоящей работе исследованы аналитические возможности методов спектроскопии диффузного отражения и методов компьютерного зрения в задаче преаналитической классификации образ-

цов на наличие гемолиза или липемии, являющиеся распространенными причинами ошибочных результатов анализов. Установлено, что спектроскопия диффузного отражения, реализованная в виде двухволоконной схемы, при проведении образца без дополнительного пробоотбора, непосредственно через стенку пробирки, обладает высокими (более 99%) показателями чувствительности и специфичности в идентификации образцов с гемолизом (концентрация более 0.7 g/L) и липемией (более 1.3 g/L). Для идентификации гемолиза и липемии также могут быть использованы методы компьютерного зрения — в частности, анализ цветовых характеристик сыворотки крови, регистрируемых с помощью фотокамеры в стандартных цветовых каналах при фиксированных условиях освещенности обладает высокой селективностью — более $97 \pm 2.5\%$ при специфичности обнаружения гемолиза и липемии $89 \pm 8\%$. При этом использование „продвинутых“ методов компьютерного зрения — сверточных нейронных сетей — позволяет применять методы компьютерного зрения для анализа изображений в режиме реального времени при движении пробирок относительно детектора, при этом, однако, чувствительность модели существенно определяется условиями освещенности, расположением пробирки относительно детектора и качеством получаемых изображений.

Благодарности

Работа выполнена в рамках программы стратегического академического лидерства „Приоритет 2030“ ФГАОУ ВО „Первый МГМУ им. И.М. Сеченова“ Минздрава России (Сеченовский университет) и в рамках работы междисциплинарной научно-образовательной школы Московского университета „Фотонные и квантовые технологии. Цифровая медицина“.

Финансирование работы

Исследование поддержано грантом Правительства Москвы на реализацию научно-практического проекта в медицине № 2212-19/2.

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Список литературы

- [1] J.Z. Ji, Q.H. Meng. *Clinica Chimica Acta*, **412** (17–18), 1550–1553 (2011).
- [2] M.B. Smith, Y.W. Chan, A. Dolci, M.D. Kellogg, C.R. McCudden, M. McLean. *Wayne, PA, USA: Clinical and Laboratory Standards Institute* (2012).
- [3] E.P. Kakorina, A.V. Polikarpov, N.A. Golubev. *Laboratory Service*, **7** (4), 32–39 (2018).
- [4] P.L. Epner, J.E. Gans, M.L. Graber. *BMJ Quality & Safety*, **22** (Suppl 2), ii6–ii10 (2013).
- [5] P. Bonini, M. Plebani, F. Ceriotti, F. Rubboli. *Clinical Chemistry*, **48** (5), 691–698 (2002).
- [6] G. Lippi, G.L. Salvagno, G. Lima-Oliveira, G. Brocco, E. Danese, G.C. Guidi. *Clinica Chimica Acta*, **440**, 164–168 (2015).
- [7] A.M. Simundic, K. Bölenius, J. Cadamuro, S. Church, M.P. Cornes, E.C. van Dongen-Lases, P. Eker, T. Erdeljanovic, K. Grankvist, J.T. Guimaraes, R. Hoke. *Clinical Chemistry and Laboratory Medicine (CCLM)*, **56** (12), 2015–2038 (2018).
- [8] G. Lima-Oliveira, G. Lippi, G.L. Salvagno, M. Montagnana, G. Picheth, G.C. Guidi. *Biochemia Medica*, **22** (2), 180–186 (2012).
- [9] F. Sanchis-Gomar, G. Lippi. *Biochemia Medica*, **24** (1), 68–79 (2014).
- [10] G. Lippi, G.L. Salvagno, E. Danese, G. Lima-Oliveira, G. Brocco, G.C. Guidi. *Clinica Chimica Acta*, **436**, 183–187 (2014).
- [11] A.-M. Simundic, M. Cornes, K. Grankvist, G. Lippi, M. Nybo. *Clinica Chimica Acta*, **432**, 33–37 (2014).
- [12] W. Barcellini. *Transfusion Medicine and Hemotherapy*, **42** (5), 287–293 (2015).
- [13] A. Abdollahi, H. Saffar, H. Saffar. *North Am. J. of Medical Sciences*, **6** (5), 224 (2014).
- [14] G. Lippi, M. Plebani, A.-M. Simundic. *Biochemia Medica*, **20** (2), 126–130 (2010).
- [15] G. Tian, Y. Wu, X. Jin, Z. Zeng, X. Gu, T. Li, X. Chen, G. Li, J. Liu. *J. Plos One*, **17** (1), e0262748 (2022).
- [16] N.J. Heyer, J.H. Derzon, L. Wingses, C. Shaw, D. Mass, S.R. Snyder, P. Epner, J.H. Nichols, J.A. Gayken, D. Ernst, E.B. Liebow. *Clinical Biochemistry*, **45** (13–14), 1012–1032 (2012).
- [17] G. Lippi, N. Blanckaert, P. Bonini, S. Green, S. Kitchen, V. Palicka, A.J. Vassault, M. Plebani. *Clinical Chemistry and Laboratory Medicine*, **46** (6), 764–772 (2008).
- [18] R. Chawla, B. Goswami, D. Tayal, V. Mallika. *Laboratory Medicine*, **41** (2), 89–92 (2010).
- [19] M.R. Glick, K.W. Ryder, S.J. Glick, J.R. Woods. *Clinical Chemistry*, **35** (5), 837–839 (1989).
- [20] G. Lippi, J. Cadamuro, A. von Meyer, A.M. Simundic. *Clinical Chemistry and Laboratory Medicine (CCLM)*, **56** (5), 718–727 (2018).
- [21] A.M. Simundic, N. Nikolac, V. Ivankovic, D. Ferenc-Ruzic, B. Magdic, M. Kvaternik, E. Topic. *Clinical Chemistry and Laboratory Medicine*, **47** (11), 1361–1365 (2009).
- [22] D.A. Noe, V. Weedn, W.R. Bell. *Clinical Chemistry*, **30** (5), 627–630 (1984).
- [23] C. Bürki, M. Volleberg, D. Brunner, M. Schmutz, M. Hersberger. *Clinical Biochemistry*, **100**, 67–70 (2022).
- [24] S. Storti, E. Battipaglia, M.S. Parri, A. Ripoli, S. Lombardi, G. Andreani. *J. Laboratory Medicine*, **43** (2), 67–76 (2019).
- [25] C.-J.L. Farrell, A.C. Carter. *Ann. Clin. Biochem.*, **53** (5), 527–538 (2016).
- [26] Z. Du, J. Liu, H. Zhang, B. Bao, R. Zhao, Y. Jin. *J. Clin. Lab. Anal.*, **33** (4), e22856 (2019).
- [27] Z. Wang, Z. Zhao. In: *MATEC Web Conf. 173*, EDP Sciences (2018).
- [28] H. Wang, H. Huang, X. Wu. *Chemom. Intell. Lab. Syst.*, **231**, 104688 (2022).
- [29] C. Yang, D. Li, D. Sun, S. Zhang, P. Zhang, Y. Xiong, M. Zhao, T. Qi, B. Situ, L. Zheng. *Clin. Chim. Acta*, **531**, 254–260 (2022).

- [30] M. Ashenden, A. Clarke, K. Sharpe, G. d'Onofrio, J. Plowman, C.J. Gore. *Int. J. Lab. Hematol.*, **35** (2), 183–192 (2013).
- [31] Supervise.ly [Электронный ресурс].
URL: <https://supervisely.com/>
- [32] C.R. Harris, K.J. Millman, S.J. Van Der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern. *Nature*, **585**, 357–362 (2020).
DOI: 10.1038/s41586-020-2649-2
- [33] W. McKinney. In: *Proceedings of the 9th Python in Science Conference*, **445** (1), 51–56 (2010).
- [34] J.D. Hunter. *Computing in Science & Engineering*, **9** (03), 90–95 (2007).
- [35] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison. *Advances in Neural Information Processing Systems*, **32** (2019).
- [36] Ultralytics [Электронный ресурс].
DOI: 10.5281/zenodo.7347926
- [37] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas. *J. machine Learning Research*, **12**, 2825–2830 (2011).
- [38] T.Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick. In: *Computer Vision – ECCV 2014: 13th European Conferenc Proceedings, Part V 13* (2014).
- [39] R. Padilla, S.L. Netto, E.A. Da Silva. In: *International conference on systems, signals and image processing (IWSSIP)*, 237–242 (2020).