

Алгоритм определения групп крови по изображениям серологических планшетов

С.А. Корчагин^{1,2}, Е.Е. Зайченкова^{1,3}, Д.А. Шарапов^{1,3}, Е.И. Ершов^{1,3}, Ю.В. Буторин^{1,2}, Ю.Ю. Венгеров^{1,2}

¹ Институт проблем передачи информации им. А.А. Харкевича РАН, 127051, Россия, г. Москва, Большой Каретный переулок, д. 19 стр. 1;

² Московский государственный университет им. М.В. Ломоносова, 119991, Россия, г. Москва, Ленинские горы, д. 1;

³ Московский физико-технический институт (НИУ), 141701, Россия, Московская область, г. Долгопрудный, Институтский пер., д. 9;

⁴ ООО «СИНТЭКО-КОМПЛЕКС», 142530, Россия, Московская область, г. Электрогорск, ул. Будёного, д. 1а

Аннотация

В работе рассматривается алгоритм для автоматизации системы медицинской экспресс-диагностики, предназначенной для определения группы крови путем анализа реакции агглютинации (склеивания эритроцитов). После забора крови и заполнения серологического планшета лаборанту достаточно разместить его в специальном сканере для последующего автоматического определения группы крови. Оцифровка данных и помощь лаборанту в распознавании планшетов позволяет решить сразу две критически важные задачи: хранение результатов анализа и контроль человеческого фактора. Предлагаемый алгоритм распознавания позволяет точно определить границы лунок и оценить степень агглютинации посредством легковесной свёрточной нейронной сети. Был собран уникальный набор данных с независимой оценкой степени агглютинации медицинскими экспертами. Наивысшая достигнутая точность оценки степени агглютинации на собранном наборе данных из 3231 изображения лунок оказалась сопоставима с оценкой среднего эксперта и составила 98,2%.

Ключевые слова: агглютинация, группа крови, классификация, преобразование Хафа, глубокое обучение.

Цитирование: Корчагин, С.А. Алгоритм определения групп крови по изображениям серологических планшетов / С.А. Корчагин, Е.Е. Зайченкова, Д.А. Шарапов, Е.И. Ершов, Ю.В. Буторин, Ю.Ю. Венгеров // Компьютерная оптика. – 2023. – Т. 47, № 6. – С. 958-967. – DOI: 10.18287/2412-6179-CO-1339.

Citation: Korchagin SA, Zaychenkova EE, Sharapov DA, Ershov EI, Butorin UV, Vengerov UU. An algorithm of blood typing using serological plate images. Computer Optics 2023; 47(6): 958-967. DOI: 10.18287/2412-6179-CO-1339.

Введение

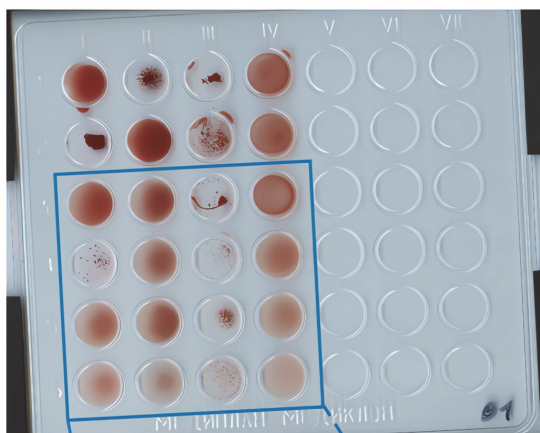
Первые попытки переливания крови начали предприниматься после того, как в 1628 году Карл Гарвей открыл циркуляцию крови [1]. Однако многие из этих экспериментов оказывались неудачными – вскоре после переливания реципиентам становилось значительно хуже. Найти этому медицинское объяснение удалось лишь в 1901 году, когда австрийско-американский иммунолог Карл Ландштейнер открыл группы крови [2]. Он обнаружил, что в плазме крови содержатся антитела – белки, нейтрализующие чужеродные тела при их попадании в организм, а на поверхности эритроцитов расположены антигены – молекулы, вступающие с антителами в реакцию, называемую агглютинацией. Визуально эта реакция проявляется как образование крупных или мелких скоплений эритроцитов. Каждая группа крови задается определенной комбинацией антигенов и является наследственным признаком, неизменным на протяжении жизни.

На март 2023 года, согласно сайту Международного общества переливания крови (ISBT), выявлено 44 системы групп крови, включающих 354 антигена. В одну и ту же систему входят структурно и генетически близкие антигены, связь которых друг с другом хорошо прослеживается. В данной работе рассматривается три системы: ABO (I-IV группы крови), Rh (Резус) и Kell (Келл), наиболее часто требуется определить группу крови именно по ним, потому что входящие в них антигены вызывают наиболее сильные осложнения при переливании несовместимой крови.

Задача определения группы крови является критической составляющей множества медицинских процессов. Если перелить человеку кровь несовместимой группы, содержащиеся в ней антигены могут быть распознаны как чужеродные объекты (если в крови реципиента таких антигенов нет), и тогда в плазме начнут вырабатываться антитела для их нейтрализации, что приведёт к разрушению эритроцитов (гемолизу). Это, в свою очередь, может причинить серьезный вред здоровью реципиента вплоть до летального

исхода. В Российской Федерации неверное определение группы крови влечет уголовную ответственность для врача.

В настоящее время в медицинской диагностике выделяют три основных методики определения группы крови: автоматическая – когда от момента забора крови до получения вердикта используются автоматизированные системы, ручная – исследование полностью осуществляется лаборантом, и полуавтоматическая – лаборант подготавливает образцы крови и тест-систему, а последующий анализ проводится автоматически. Все методы в той или иной степени распространены в клиниках, но, как показывает практика, ручной труд не застрахован от человеческой ошибки, а полностью автоматические системы, хотя и являются наиболее производительными, – чрезмерно дорогие. В данной работе рассматривается полуавтоматический случай, позволяющий сохранить баланс между контролем человеческого фактора и стоимостью используемого оборудования, благодаря чему предлагаемый алгоритм наиболее актуален для небольших лабораторий. Кроме того, он позволяет оцифровывать полученные данные для их дальнейшего безвременного хранения и более удобного анализа. В данной статье для определения группы крови используется методика анализа реакций агглютинации по изображению подготовленного серологического планшета (рис. 1).



Образец	Реагенты				Вердикт
	A	B	D	K	
1	-	+	+	-	B(III) Rh+ K-
2	+	-	+	-	A(II) Rh+ K-
3	-	-	+	+	O(I) Rh+ K+
4	+	-	+	-	A(II) Rh+ K-

Рис. 1. Иллюстрация работы алгоритма

Рассматриваемая процедура проведения анализа заключается в следующем: для каждой пары образец крови–реагент в отдельной лунке серологического планшета стеклянной палочкой перемешиваются несколько капель образца крови и добавленного к нему реагента, затем отводится время на протекание реакции агглютинации (от 30 секунд до 3 минут в зависимости от реагента). После этого планшет сканируется, и к полученному изображению применяется представленный в данной статье алгоритм определения наличия реакции агглютинации. Затем происходит формирование итогового результата исследования в виде групповой принадлежности проанализированных образцов.

В параграфе 1 представлен литературный обзор подходов, используемых для определения групп крови, подробно рассмотрены методы, основанные на проведении реакции агглютинации и определении ее результатов методами цифровой обработки изображений. В параграфе 2 приведено описание и анализ используемого набора данных. Параграф 3 посвящен предлагаемому алгоритму: подпараграф 3.1 – построению множества прямых («решётки»), лежащих в пространстве между лунками, подпараграф 3.2 – нахождению угла наклона планшета на изображении, подпараграф 3.3 – распознаванию точных границ лунок, а подпараграф 3.4 – распознаванию степени агглютинации. В параграфе 4 описан процесс вынесения вердикта по комбинации выходов реакций агглютинации и полученные результаты. В параграфе 5 приведены анализ результатов и обсуждение эффективности предлагаемого решения.

1. Методы определения агглютинации

На сегодняшний день существует множество методов для определения группы крови, но при выборе конкретного приходится выдерживать компромисс между временем проведения, точностью и стоимостью эксперимента. Помимо упомянутой выше реакции агглютинации, в основе теста по определению группы крови может лежать молекулярный импринтинг [3], поверхностный плазмонный резонанс [4], намагничивание эритроцитов [5], твердофазная экстракция [6], проточная цитометрия [7], генотипирование [8] и прочее. Здесь и далее рассматривается подход, основанный на проведении реакции агглютинации, поскольку он является наиболее распространенным в России. В качестве тест-объектов [9] могут использоваться стеклянные пластинки, пробирки, планшеты с микроячейками, гелевые карты [10], пластинки с фильтровальной бумагой [11] и т. д.

Для повышения точности анализа могут использоваться более сложные устройства, усиливающие протекание реакции: например, волновод [12] (2012), ультразвуковой генератор [13] (1999 год), лазер [14] (2010). Существуют также различные системы, позволяющие полностью автоматизировать проведение

теста: Auto-Grouper [15], Olympus PK [16], [17], Ortho Auto value Innova System [18], Immucor Galileo [19], Techno Twin Station [20]. Такие приборы отличаются еще более высокой стоимостью и, хоть и позволяют проводить большое количество тестов одновременно, в случае небольших масштабов (десятков образцов, что свойственно большинству региональных клинико-диагностических центров) каждый отдельный анализ занимает гораздо больше времени, чем требуется при проведении теста вручную и тем более в полуавтоматическом режиме.

Существует множество методов оценки агглютинации, основанных на спектроскопии: агглютинация определяется по изменению или сохранению спектра излучения крови, измеренного спектрофотометром [21] или фотодиодом [22]. Также в основе подхода может лежать подсчет с помощью специального анализатора числа скоплений эритроцитов до и после агглютинации [23]. Тесты на гелевых картах заключаются в оценке степени диффузии эритроцитов в геле после протекания реакции [10]. В данной работе рассмотрен подход, основанный на обработке изображений. Преимуществами данного метода отличаются простота предварительной подготовки и низкая стоимость оборудования (необходима лишь камера или сканер).

Для распознавания агглютинации путем обработки изображений в 2013 году Ана Ферраз предложила использовать для определения наличия агглютинации фиксированный порог по стандартному отклонению значений пикселей изображения реакции в зелёном канале [24]. Подобный подход сильно зависит от набора данных, что приводит к его плохой масштабируемости. Это подтверждается работами [25] и [26], где использовался идентичный подход, но для достижения наилучшего результата были зафиксированы альтернативные глобальные пороги. В 2017 году команда из Бангладеша под руководством профессора Уддина [27] также это подтвердила и взамен предложила использовать альтернативный признак: количество компонент связности карты границ изображения пластинки с реакцией. Однако в данном методе тоже использовался глобальный порог, и, хотя им, как и Ферраз, удалось добиться 100% точности на своем наборе данных, данный метод страдает от тех же проблем. В 2016 году в [28] в качестве признака использовался еще один вариант фиксированного глобального порога – количество особых точек, найденных детектором SURF.

Из более современных решений, использующих глобальные пороги разделения признаков, можно выделить метод [29], разработанный в 2017 году. Он основан на использовании уже двух признаков: числа белых пикселей на бинаризованном изображении выхода реакции и количества компонент связности белых пикселей («объектов») внутри выхода реакции. В 2019 году в [30] использовались сразу три признака: число

белых пикселей, число компонент связности белых пикселей на бинаризованном изображении, а также число компонент связности карты границ Кэнни.

В качестве способов разделения признаков могут также использоваться методы машинного обучения. Так, в [31] применяется нейронная сеть, использующая в качестве признаков плотность и разреженность пикселей на бинарном изображении. В 2013 году был предложен метод [32], в котором использовались изображения прореагировавшей крови под микроскопом и для распознавания агглютинации применялся метод опорных векторов. В качестве признаков были взяты значения тона, насыщенности и светлоты, полученные после перехода в пространство HSL, а также некоторые характеристики текстуры (связанные с контрастностью, энтропией, средним, стандартным отклонением). Метод опорных векторов также применялся в [33] (2016 год), где в качестве признаков использовались особые точки, найденные при помощи детектора SIFT.

Во всех описанных алгоритмах реакция агглютинации проводилась на стеклянных пластинках, из-за чего отдельной задачей являлось распознавание области на изображении, в которой находится прореагировавшая смесь. Когда реагенты раскапываются на серологической планшете, данная задача решается более надежными и простыми методами за счет наличия на планшете лунок, границы которых хорошо различимы на изображении. Поэтому в данной работе рассматривается подход, в котором используется именно планшет. Кроме того, по сравнению с тестами в пробирках тесты на планшетах более точные, а по сравнению с гелевыми картами – более дешевые. Помимо этого, планшеты позволяют анализировать одновременно несколько образцов крови и для каждого образца проводить больше реакций.

2. Набор данных

Набор данных состоит из 86 изображений планшетов модели «Планшет многоцветный для проведения серологических реакций по ТУ 32.50.50-105-51203590-2020 на 42 лунки прозрачный (Медиклон)» (смотри рис. 2), полученных с помощью аппаратно-программного комплекса «Эксперт-Лаб». Сканированные изображения имеют разрешение 236,22 пикселей на сантиметр (7020×5100 пикселей). Лунки на планшетах круглые, расположены в шесть горизонтальных и семь вертикальных рядов. При проведении реакции каждый горизонтальный ряд соответствует отдельному образцу крови, а каждый вертикальный – реагенту, с которым смешивают образец. Максимально на планшете может быть представлено 6 различных образцов крови и 7 различных реагентов: каждая непустая лунка содержит смесь одного образца крови и одного реагента. Суммарно планшеты из используемого набора данных содержат 516 различных образцов крови и 3231 непустую лунку. Данные

о группах крови доноров были получены с их информированного согласия. Процесс сбора данных и их дальнейшей обработки удовлетворяет требованиям Хельсинского соглашения 2013 года.

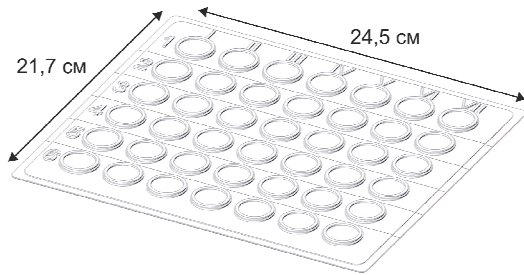


Рис. 2. Схематический вид планшета «Медиклон»

Для каждого образца крови известна истинная группа, взятая из медицинской карты (все использованные образцы принадлежат донорам, а потому их группы крови подтверждены путем многократно проведенных исследований). На основе информации об истинной группе крови определены истинные значения степени агглютинации для каждой реакции.

Кроме того, для каждого планшета получена разметка степени агглютинации в каждой лунке, выполненная пятью независимыми квалифицированными медицинскими экспертами в области серологии. Степень агглютинации определялась как одно из следующих значений: 0 – точно нет агглютинации, 1 – скорее всего нет агглютинации, 2 – затруднительно вынести вердикт, 3 – скорее всего есть агглютинация, 4 – точно есть агглютинация. Для соответствия разметке экспертов истинные результаты реакции принимают значения либо 0, либо 4.

Несмотря на то, что все опрошенные эксперты являются квалифицированными медицинскими работниками, их оценка агглютинации иногда не совпадает с истинными результатами. Такое может произойти по многим причинам: например, эксперты не знают всей информации о проведенном эксперименте (время реакции, отсутствие возможности оценить объемную картину содержимого лунки и т.д.), поэтому была осуществлена проверка адекватности их разметки путем сравнения ответов каждого эксперта с истинными результатами агглютинации и разметкой других экспертов. Для этого для каждого эксперта была вычислена нормированная метрика Вассерштайна для одномерного распределения:

$$\frac{1}{4n} \sum_{i=1}^n |gt_i - markup_i|, \tag{1}$$

где n – число лунок в рассматриваемом наборе данных, gt_i – истинная степень агглютинации в i -й лунке, принимающая значение 0 или 4, $markup_i$ – степень агглютинации в i -й лунке по мнению эксперта, принимающая одно из значений [0, 1, 2, 3, 4].

Смысл данной метрики в том, что если эксперт дает для каждой лунки ответ, противоположный ис-

тинному (например, что агглютинация точно есть, когда её точно нет), то для него метрика равняется 1,0. Если эксперт, напротив, всегда безошибочно и уверенно определяет степень агглютинации, то метрика равняется 0,0. Результаты вычисления этой метрики и точность (ассигу) приведены в табл. 1.

Табл. 1. Степень согласованности оценок степени агглютинации, данных каждым из экспертов, с истинной степенью агглютинации

Эксперт, №	1	2	3	4	5
Введенная метрика	0,075	0,041	0,033	0,017	0,080
Точность, %	96,1	94,6	97,6	100	93,2

Для оценки попарной согласованности экспертов между собой вычислялся коэффициент Каппа Коэна (1,0 означает абсолютное совпадение ответов, 0,0 – их полное рассогласование). Результаты проиллюстрированы на рис. 3.

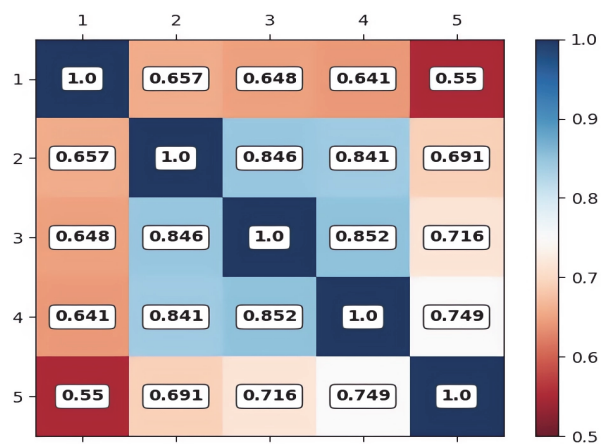


Рис. 3. Степень согласованности оценок степени агглютинации между каждой парой экспертов

По результатам анализа согласованности экспертов между собой и с истинными значениями были выявлены лунки, в которых наблюдалось наибольшее расхождение ответов. На основе анализа этих данных были выявлены наиболее распространённые признаки сложных для определения агглютинации случаев (смотри рис. 4).

Эти случаи можно разделить на 2 основные группы: случаи, вызванные некачественной подготовкой эксперимента лаборантом (несоблюдение пропорций объемов образца и реагента, неверная оценка времени протекания реакции), и случаи, связанные со свойствами реагентов и индивидуальными особенностями образцов. Если корректную подготовку эксперимента можно контролировать, имея опытного лаборанта, то второй тип случаев представляет из себя испытание для оценивающей системы. В него входит мелкодисперсная агглютинация – подвид агглютинации, при котором сгустки эритроцитов имеют небольшой размер. Ее легко перепутать с пересыханием лунки, из-за

которого становится невозможно дать правильный ответ, или «монетными столбиками» – физическим явлением, при котором эритроциты образуют скопления под действием электростатических сил.

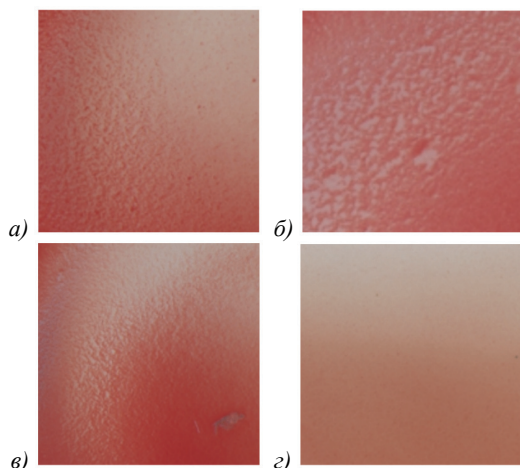


Рис. 4. Наиболее сложные случаи при определении степени агглютинации: (а) мелкодисперсная агглютинация, (б) монетные столбики, (в) пересыхание, (г) неверная пропорция реагентов

Решение по оценке агглютинации разрабатывалось, чтобы определять агглютинацию в любом возможном сложном случае.

3. Структура метода

3.1. Построение «решётки»

В общем случае радиус лунки в пикселях неизвестен и планшет на изображении может быть наклонен, что критично для алгоритма распознавания границ лунок. Поэтому первым важным шагом предлагаемого метода является обнаружение множества прямых («решётки»), лежащих в пространстве между лунками. Шаг «решётки» позволяет оценить радиус лунок, а совокупность наклонов найденных прямых – угол наклона планшета на изображении.

Нахождение прямых «решётки» осуществляется при помощи быстрого преобразования Хафа (БПХ) для двумерного случая [34]. Основная идея этого алгоритма заключается в формировании признакового пространства, в Хаф-образе которого искомые прямые соответствуют локальным максимумам. Хаф-образ – это изображение, в котором значение пикселя с координатами (s, t) равняется сумме значений пикселей по прямой исходного изображения, задаваемой её попиксельным сдвигом s и параметром наклона t . Для формирования признакового пространства к исходному изображению последовательно применяются следующие преобразования:

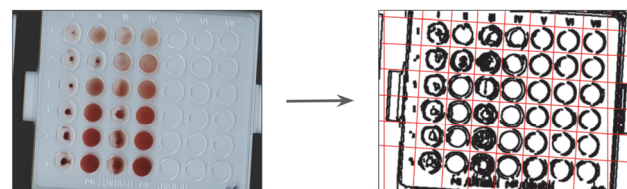
1. Изменяется размер изображения таким образом, чтобы максимальная сторона была равна степени двойки.
2. Строится инвертированная карта границ Кэнни [35].

3. Применяется морфологическая операция эрозии [36].
4. Осуществляется гауссовская свертка [37].
5. Строится дополнение изображения нулями до квадратного.

Дальнейший поиск прямых осуществляется посредством повторения следующих шагов:

1. Находится глобальный максимум на Хаф-образе (точка будет соответствовать одной из искомым прямым);
2. Обновляется признаковое изображение: закрашивается (обнуляется) полоса, в центре которой проходит найденная прямая.

Данные шаги осуществляются отдельно для нахождения преимущественно горизонтальных и преимущественно вертикальных прямых и повторяются до тех пор, пока не найдётся нужное количество прямых на признаковом изображении (или не будет занулено всё изображение) (смотри рис. 5).



Входное изображение, 7020×5100 пикселей

Построение «решётки»

Рис. 5. Иллюстрация работы алгоритма построения «решётки»

3.2. Нахождение угла поворота

Следующим шагом алгоритма является нахождение угла наклона планшета на изображении при помощи построенного на предыдущем шаге набора прямых и последующее выравнивание изображения планшета.

Предлагаемое решение состоит из следующих шагов:

1. Найденные преимущественно вертикальные прямые «решётки» поворачиваются на 90°, сохраняя свой первоначальный сдвиг, таким образом, чтобы их можно было представить на том же Хаф-образе, что и преимущественно горизонтальные прямые.
2. Точки, соответствующие преимущественно горизонтальным и вертикальным прямым, отмечаются на едином Хаф-образе и размываются Гауссовым фильтром.
3. С помощью преобразования БПХ находится прямая, аппроксимирующая построенные в пункте 2 области. Сдвиг этой прямой s задает средний угол наклона «решётки» и, соответственно, примерный угол наклона планшета на изображении (смотри рис. 6).
4. Исходное изображение поворачивается относительно центра на найденный в пункте 3 угол.

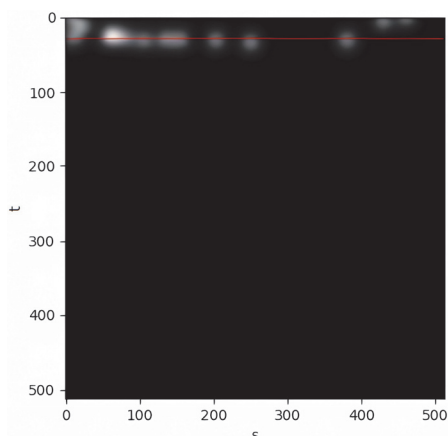


Рис. 6. Пример нахождения аппроксимирующей прямой

На собранном наборе данных угол наклона планшетов был очень маленьким, поэтому для проверки работы данного этапа алгоритма был проведен дополнительный эксперимент. Каждое изображение из набора данных было 10 раз искусственно повернуто на случайный угол от -5° до 5° с устранением артефактов поворота и при помощи описанного алгоритма возвращено в исходное положение. Средняя ошибка определения угла наклона составила меньше $0,5^\circ$.

3.3. Алгоритм распознавания границ лунок

После построения «решётки» и выравнивания планшета находятся границы лунок посредством градиентного преобразования Хафа для окружностей [38] со следующими параметрами:

- максимально возможный радиус искомых лунок max_rad как половина медианного шага найденной «решётки»;
- минимально возможный радиус min_rad как $0,6 \cdot max_rad$;
- наименьшее допустимое расстояние между центрами искомых лунок как $2 \cdot min_rad$.

Выбор данного метода обусловлен тем, что из концентрических окружностей выбирается единственная, причем наибольшего радиуса, что гарантирует распознавание для каждой лунки единственной соответствующей окружности и повышает вероятность распознать лунку полностью. Но даже несмотря на это, в результате работы алгоритма могут быть найдены лишние окружности, не соответствующие лункам, а радиус найденных окружностей может сильно варьироваться от min_rad до max_rad . Поэтому для всех окружностей выполняются следующие шаги:

1. Окружности, не попадающие в один ряд с некоторыми из других распознанных окружностей, удаляются.
2. Горизонтальная координата центра приравнивается к медиане горизонтальных координат центров окружностей в том же вертикальном ряду.
3. Вертикальная координата центра приравнивается к медиане из вертикальных координат центров окружностей в том же горизонтальном ряду.

4. Радиус приравнивается к медианному радиусу всех окружностей (см. рис. 7).

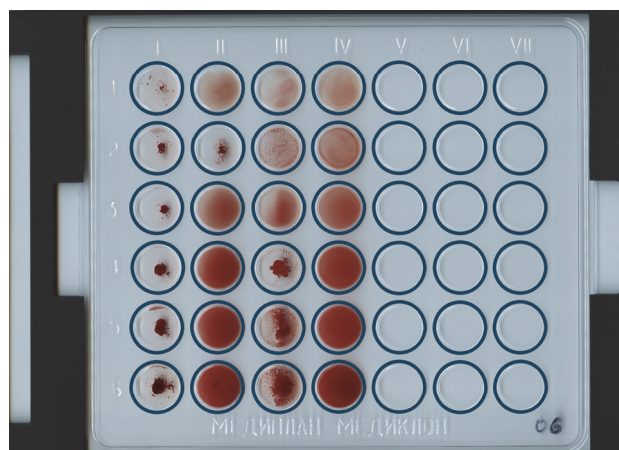


Рис. 7. Пример распознавания границ лунок

В качестве метрики при оценке качества работы алгоритма распознавания лунок используется расстояние Фреше, вычисляемое оптимизированным образом, предложенным в работе [39]. В нашем случае обнаруженная и соответствующая истинной окружности являются почти концентрическими, а для концентрических расстояние Фреше равно модулю разницы их радиусов. Данный алгоритм позволяет находить границы окружностей безошибочно при условии, что отношение расстояния Фреше между найденной и истинной окружностью к радиусу истинной окружности меньше заданного коэффициента 0,2.

3.4. Определение степени агглютинации

В работе предложен нейросетевой алгоритм для определения степени агглютинации и показано, что его точность лучше базового метода, предложенного в работе [27], основанного на вычислении числа компонент связности карты границ Кэнни и дальнейшем сравнении этого числа с обученным глобальным порогом для изображений планшетов из главы 2.

Для формирования набора данных в задаче определения степени агглютинации из изображений планшетов вырезались изображения задействованных лунок. Для каждого из 56 планшетов 4 случайные строки относились к обучающей выборке, 1 строка к валидационной и 1 строка к тестовой выборкам. Разделение таким образом гарантированно сохраняет баланс различных типов реагентов между выборками. Ещё 30 изображений планшетов, собранные позднее, были полностью отнесены к тестовой выборке. Суммарно соотношение обучающей/валидационной/тестовой выборки составило примерно 4:1:5 (1331 лунка в обучающей, 340 в валидационной и 1501 в тестовой выборке). Для базового алгоритма глобальный порог разделения подбирался на объединённых валидационной и обучающей выборках.

Основным методом классификации агглютинации является обученная легковесная нейронная сеть с 5

свёрточными слоями: размер ядра свёртки 3×3 , после свёртки применяется функция активации ReLu, а затем на всех слоях, кроме последнего, – слой подвыборки по максимальному значению (max-pooling) с ядром 2×2 . После этого выделенные особенности преобразуются в вектор, проходящий через два полносвязных слоя: после первого происходит исключение случайных нейронов с вероятностью 0,5 (dropout) и активация ReLu, а после второго – сигмоидная активация, переводящая окончательный ответ нейронной сети в пределы отрезка $[0, 1]$ (смотри рис. 8).

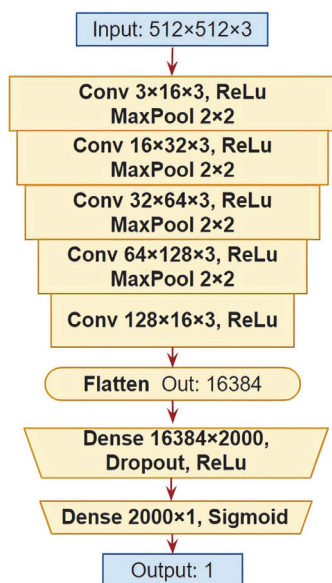


Рис. 8. Архитектура нейронной сети

Важной особенностью предложенной архитектуры является то, что сжатие исходного изображения практически отсутствует (на вход сети подаётся изображение 512×512). Такое решение обусловлено тем, что размер пикселя на изображении лунок должен быть сопоставим с размером сгустков эритроцитов. Это подтверждается проведённой серией экспериментов, в которых варьировался размер входного изображения лунки (128×128 , 256×256 и 512×512 пикселей) с соответствующими изменениями в архитектуре нейронной сети. Для каждого варианта размеров изображения лунки было обучено по 10 нейронных сетей одинаковой архитектуры. В табл. 2 приведена средняя точность (ассигасу) и стандартное отклонение на обучающей, валидационной и тестовой выборках по этим 10 нейронным сетям. Видно, что на тестовой выборке лучшее качество было достигнуто нейронными сетями, обученными на изображениях наибольшего размера. Изображения планшетов были получены при разрешении сканера 236,22 пикселей на сантиметр, из чего следует, что для изображений размером 512×512 пикселей один пиксель соответствует линейному размеру приблизительно 6,5 стандартных эритроцитов.

Матрица ошибок для лучшей обученной нейронной сети приведена в табл. 3. Можно заметить, что

алгоритм ошибается в предсказании отрицательной агглютинации в разы больше, чем при предсказании положительной. Это напрямую связано со сложностью собранного набора данных и наличием большого числа лунок с мелкодисперсной агглютинацией. В дальнейшем во всех экспериментах была использована архитектура, принимающая на вход изображения размера 512×512 пикселей.

Табл. 2. Усредненная точность (%) работы нейронных сетей для различных размеров входных изображений лунок

Размер лунки, пикселей	128x128	256x256	512x512
Обуч. выборка	99,4 ± 0,3	99,3 ± 1,3	97, ± 1,0
Валид. выборка	96,9 ± 0,6	97,2 ± 0,6	97,2 ± 0,4
Тест. выборка	96,4 ± 0,9	97,6 ± 0,4	98,0 ± 0,3

Табл. 3. Матрица числа ошибок сравнения предсказанной и истинной агглютинации

	Истинно положительная	Истинно отрицательная
Предсказанная положительная	1029	6
Предсказанная отрицательная	23	1083

Во всех случаях процесс обучения обладал следующими особенностями:

1. В качестве алгоритма оптимизации нейронной сети использовался Adam [40].
2. Обучение проходило в течение 1500 эпох.
3. В качестве аугментации использовались случайные поворот изображения лунки до 180° , небольшой случайный сдвиг его границ до 15 пикселей, а также небольшой гауссовский шум.
4. Для каждого размера входного изображения лунки обучалась 10 нейронных сетей с одними и теми же параметрами.
5. Оптимальный порог разделения ответа нейронной сети (есть агглютинация или нет) подбирался таким образом, чтобы давать наилучшую точность (ассигасу) на валидационной выборке.

4. Результаты

Групповая принадлежность каждого образца определяется результатом реакций агглютинации, произошедшей после смешивания этого образца с различными реагентами. Выбор реагентов зависит от системы, по которой определяется группа крови. Так, для определения группы по системе АВО образец (эритроциты или цельная кровь) смешивается с реагентами анти-А и анти-В, представляющими собой смеси моноклональных антител к антигенам А и В (см. табл. 4), это так называемая прямая реакция.

В отличие от остальных исследуемых в данной работе систем (Rh и Kell), решение о групповой принадлежности образца по системе АВО принимается на основании комбинации выходов реакции обоих

реагентов. Кроме того, как правило, параллельно выполняется контрольная перекрестная реакция – это реакции сыворотки образца с контрольными эритроцитами, антигенный состав (групповая принадлежность) которых известен. Результат агглютинации перекрестной реакции должен быть противоположен результату прямой. В случаях, когда результат перекрестной реакции противоречит результату прямой, вердикт о групповой принадлежности образца по системе АВО выносится по результатам прямой реакции, но исследование должно быть проведено повторно с использованием других реагентов или других методов.

Табл. 4. Соответствие результатов агглютинации и групп крови по системе АВО

	O(I)	A(II)	B(III)	AB(IV)
Эритроциты образца + анти-А	–	+	–	+
Эритроциты образца + анти-В	–	–	+	+

Для определения резус-принадлежности образца по системе Rh (резус) проверяется присутствие на его эритроцитах антигена D, в случае обнаружения которого образец считается резус-положительным (Rh+). Также часто проводится исследование на выявление минорных антигенов системы Rh: C, E, c, e и Cw, которые менее выражено вступают в реакцию, но все же могут вызвать в организме иммунологический конфликт. В системе Kell наиболее сильные осложнения может вызвать антиген K (Келл). При проведении реакции, аналогично системе АВО, образец крови смешивается с сывороткой, содержащей антитела к соответствующему антигену.

В табл. 5 приведена точность для лучшей из 10 обученных нейронных сетей, подобранной таким образом, чтобы максимизировать точность на валидационной выборке.

Табл. 5. Точность (%) определения агглютинации и групп крови по системам в сравнении с точностью работы среднего эксперта

	ABO + контроль	Rh	Kell	Аггл.	Средний эксперт
Базовый	65,8	94,5	93,1	93,1	96,3
Нейро-сетевой	98,3	97,6	97,4	98,2	

5. Обсуждение

Важно понимать, что при ручном анализе выходов реакций на серологическом планшете лаборант выносит вердикт о группе крови, опираясь на большее количество факторов, нежели просто двумерное изображение реакций. У него есть возможность оценить трехмерную картину реакции, посмотреть на содержимое лунки в движении (потряхивая и наклоняя планшет), а также наблюдать за реакцией на протя-

жении всего времени ее протекания. Таким образом, разработанная система находится в невыгодном положении относительно ручного труда, однако она может использоваться для ускорения обработки реакций и в качестве вспомогательного инструмента для контроля работы лаборанта при вынесении окончательного вердикта.

Анализируя полученные результаты работы системы (смотри табл. 5), можно сделать вывод, что разработанный нейросетевой подход работает в разы лучше базового метода, а его качество сопоставимо с качеством среднего эксперта. Вердикт о группе крови по системе АВО без учета контрольных реагентов выносится безошибочно. Для того, чтобы полноценно сравнить точность работы предлагаемого алгоритма с точностью работы медицинских экспертов, качество алгоритма было дополнительно подсчитано на той же подвыборке, которую оценивали эксперты (смотри параграф 2). Вычисленная таким образом точность работы алгоритма составила 96,4%, что практически совпадает с точностью работы среднего эксперта.

Сравнение с другими работами в данной области осложнено отсутствием единого набора данных (benchmark) для проведения экспериментов. В каждой из статей [24–33] был собран свой, как правило, небольшой (до 100 изображений реакций) набор данных. Из перечисленных работ из параграфа 1 наиболее репрезентативными являются статьи [29] и [30] с тестовыми выборками из 80 и 52 изображений реакций, на которых была достигнута точность (accuracy) 98,0% и 84,6% соответственно, что не идет в сравнение с анализируемым в данной статье набором данных из 3231 изображения с достигнутой точностью на тестовой подвыборке в 98,2%.

Заключение

В данной работе представлено решение задачи определения группы крови человека путем анализа изображения отсканированного серологического планшета с реакциями агглютинации. Был собран уникальный набор данных для определения групп крови по трём наиболее распространённым системам: АВО (I-IV группы), Rh (Резус) и Kell (Келл). На всем наборе данных были размечены лунки, и информация об истинной группе крови образцов была взята из медицинских карт доноров. Кроме того, на части набора данных степень агглютинации была оценена 5-ю медицинскими экспертами. Метод, разработанный на основе обученной легковесной нейронной сети, позволил достичь на собранном наборе данных точности определения агглютинации в 98,2%, что сопоставимо с точностью работы среднего медицинского эксперта.

Направления дальнейшего развития данной работы включают в себя:

1. Повышение точности распознавания агглютинации за счёт использования информации о распространённости антигенов и частоте комбина-

ций выходов реакции, а также информации о сложности лунок из разметки агглютинации медицинскими экспертами.

2. Расширение метода для других типов планшетов.
3. Калибровку нейронной сети с целью преобразовать её ответы в уверенность в выдаваемом результате.
4. Исследование качества работы нейронных сетей других архитектур.
5. Оценку сложности данных, полученную на основе согласованности экспертов между собой и с истинными результатами.
6. Распознавание особых случаев, таких как мелкодисперсная агглютинация (вид агглютинации, когда сгустки эритроцитов небольшого размера) и «монетные столбики» (электростатическое явление, при котором эритроциты образуют скопления, а потому его легко перепутать с агглютинацией).

References

- [1] Harvey W. On the motion of the heart and blood in animals: A new edition of William Harvey's *Exercitatio anatomica de motu cordis et sanguinis in animalibus*. Benton: Resource Publications; 2016. ISBN: 1-4982-3508-5.
- [1] Landsteiner K. Über Agglutinationserscheinungen normalen menschlichen Blutes. 1901 [Agglutination phenomena of normal human blood]. *Wiener Klinische Wochenschrift*; 2001; 768-769.
- [2] Piletsky SS, Rabinowicz S, Yang Z, Zagar C, Piletska EV, Guerreiro A, Piletsky SA. Development of molecularly imprinted polymers specific for blood antigens for application in antibody-free blood typing. *ChemComm* 2017; 11: 1793-1796. DOI: 10.1039/c6cc08716g.
- [3] Hounkhamhang N, Vongsakulyanon A, Peungthum P, Sudprasert K, Kitpoka P, Kunakorn M, Srikihirin T. ABO blood-typing using an antibody array technique based on surface plasmon resonance imaging. *Sensors* 2013; 13(9): 11913-11922. DOI: 10.3390/s130911913.
- [4] Bouix O, Ferrera V, Delamaire M, Redersdorff JC, Roubinet F. Erythrocyte-magnetized technology: an original and innovative method for blood group serology. *Transfusion* 2008; 48(9): 1878-1885. DOI: 10.1111/j.1537-2995.2008.01790.x.
- [5] Ching E. Solid phase red cell adherence assay: A tubeless method for pretransfusion testing and other applications in transfusion science. *Transfus Apher Sci* 2012; 46(3): 287-291. DOI: 10.1016/j.transci.2012.03.018.
- [6] Krishnan NS, Fleetwood P, Higgins RM, Hathaway M, Zehnder D, Mitchell D, Briggs D. Application of flow cytometry to monitor antibody levels in ABO incompatible kidney transplantation. *Transplantation* 2008; 86(3): 474-477. DOI: 10.1097/01.tp.0000332337.86453.42.
- [7] Gorakshakar A, Gogri H, Ghosh K. Evolution of technology for molecular genotyping in blood group systems. *Indian J Med Res* 2017; 146(3): 305. DOI: 10.4103/ijmr.IJMR_914_16.
- [8] Malomgré W, Neumeister B. Recent and future trends in blood group typing. *Anal Bioanal Chem* 2009; 393(5): 1443-1451. DOI: 10.1007/s00216-008-2411-3.
- [9] Lapierre Y, Rigal D, Adam J, Josef D, Meyer F, Greber S, Drot C. The gel test: a new way to detect red cell antigen-antibody reactions. *Transfusion* 1990; 30(2): 109-113. DOI: 10.1046/j.1537-2995.1990.30290162894.x.
- [10] Noiphung J, Talalak K, Hongwarittorn I, Pupinyo N, Thirabowonkitphithan P, Laiwattanapaisal W. A novel paper-based assay for the simultaneous determination of Rh typing and forward and reverse ABO blood groups. *Biosens Bioelectron* 2015; 67: 485-489. DOI: 10.1016/j.bios.2014.09.011.
- [11] Uno S, Tanaka T, Ashiba H, Fujimaki M, Tanaka M, Hatta Y, Makishima M. Sensitive typing of reverse ABO blood groups with a waveguide-mode sensor. *J Biosci Bioeng* 2018; 126(1): 131-137. DOI: 10.1016/j.jbiosc.2018.01.019.
- [12] Heucke UW, Cobet U. Blood group typing by ultrasound backscattering—Quantitative measurements of agglutinates and their shear-dependent behavior. *Instrum Sci Technol* 2000; 28(4): 311-321. DOI: 10.1081/ci-100100980.
- [13] Doubrovski VA, Dolmashkin AA. Human blood group typing based on digital photographs of RBC agglutination process. *Opt Spectrosc* 2010; 109(2): 263-267. DOI: 10.1134/s0030400x10080187.
- [14] Sturgeon P. Automation: its introduction to the field of blood group serology. *Immuno-hematology Journal of Blood Group Serology and Education* 2001; 17(4): 100-105. DOI: 10.21307/immuno-hematology-2019-561.
- [15] Lifshitz MS, De Cresce RP. The Olympus PK 7100 automated pretransfusion blood testing system. *Lab Med* 1987; 18(3): 182-183. DOI: 10.1093/labmed/18.3.182.
- [16] Novaretti MCZ, Navarro SP, Dorliac-Liacer PE, Chamone DDAF. K phenotyping using a PK-7200 automated analyzer. *Immuno-hematology* 1998; 14(1): 22-25. DOI: 10.21307/immuno-hematology-2019-649.
- [17] Dada A, Beck D, Schmitz G. Automation and data processing in blood banking using the Ortho AutoVue® Innova System. *Transfus Med Hemother* 2007; 34(5): 341-346. DOI: 10.1159/000106558.
- [18] Wittmann G, Frank J, Schramm W, Spannagl M. Automation and data processing with the Immucor Galileo® system in a university blood bank. *Transfus Med Hemother* 2007; 34(5): 347-352. DOI: 10.1159/000107936.
- [19] Shin SY, Kwon KC, Koo SH, Park JW, Ko CS, Song JH, Sung JY. Evaluation of two automated instruments for pretransfusion testing: AutoVue Innova and Techno TwinStation. *The Korean Journal of Laboratory Medicine* 2008; 28(3): 214-420. DOI: 10.3343/kjlm.2008.28.3.214.
- [20] Narayanan S, Galloway L, Nonoyama A, Leparc GF, Garcia-Rubio LH, Potter RL. UV-visible spectrophotometric approach to blood typing II: Phenotyping of subtype A2 and weak D and whole blood analysis. *Transfusion* 2002; 42(5): 619-626. DOI: 10.1046/j.1537-2995.2002.00090.x.
- [21] Ramasubramanian MK, Alexander SP. An integrated fiberoptic-microfluidic device for agglutination detection and blood typing. *Biomed Microdevices* 2009; 11(1): 217-229. DOI: 10.1007/s10544-008-9227-y.
- [22] Sheng N, Liu L, Liu H. Quantitative determination of agglutination based on the automatic hematology analyzer and the clinical significance of the erythrocyte-specific antibody. *Clin Chim Acta* 2020; 510: 21-25. DOI: 10.1016/j.cca.2020.06.042.
- [23] Ferraz A. Automatic system for determination of blood types using image processing techniques. 2013 IEEE 3rd Portuguese Meeting in Bioengineering (ENBENG) 2013: 1-6. DOI: 10.1109/enbeng.2013.6518441.
- [24] HasanTalukder M, Reza M, Begum B, Islam R, Hasan M. Improvement of accuracy of human blood groups determination using image processing techniques. *Int J Adv Res Comput Commun Eng* 2015; 4(10): 411-412. DOI: 10.17148/ijarcc.2015.41090.

- [25] Dong Y, Fu W, Zhou Z, Chen N, Liu M, Chen S. ABO blood group detection based on image processing technology. 2017 2nd Int Conf on Image, Vision and Computing (ICIVC) 2017: 655-659. DOI: 10.1109/icivc.2017.7984637.
- [26] Rahman S, Rahman A, Khan FA, Shahjahan SB, Nahar K. Blood group detection using image processing techniques. Doctoral dissertation. Bangladesh: BRAC University; 2017.
- [27] Sahastrabudde AP, Aji SD. Blood group detection and RBC, WBC counting: an image processing approach. Int J Comput Sci Eng 2016; 5(10): 18635-18639. DOI: 10.18535/IJECS/V5I10.49.
- [28] Yamin A, Imran F, Akbar U, Tanvir SH. Image processing based detection & classification of blood group using color images. 2017 Int Conf on Communication, Computing and Digital Systems (C-CODE) 2017: 293-298. DOI: 10.1109/c-code.2017.7918945.
- [29] Atici H, Koçer HE, Kader S. Determination of blood group by image processing using digital images. Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Dergisi 2020; 7(2): 649-659. DOI: 10.35193/bseufbd.646847.
- [30] Anuradha T. Human blood type classification by ROI based pixel segmentation and neural network. Des Eng 2021; 23(8): 9828-9836. DOI: 10.26438/ijcse/v7i7.230234.
- [31] Fathima SMN. Classification of blood types by microscope color images. Int J Mach Learn Comput 2013; 3(4): 376. DOI: 10.7763/ijmlc.2013.v3.342.
- [32] Rathod A, Pathan A. Determination and classification of human blood types using SIFT transform and SVM classifier. Int J Adv Res Electr Electron Instrum Eng 2016; 5: 8467-8473. DOI: 10.15662/IJAREEIE.2016.0511031.
- [33] Ershov EI, Terekhin AP, Nikolaev DP. Generalization of the fast hough transform for three-dimensional images. J Commun Technol Electron 2018; 63(6): 626-636. DOI: 10.1134/s1064226918060074.
- [34] Canny J. A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 1986; 6: 679-698. DOI: 10.1016/b978-0-08-051581-6.50024-6.
- [35] Fisher R, Perkins S, Walker A, Wolfart E. Hypermedia image processing reference. Hoboken: John Wiley & Sons Ltd; 1996: 118-130. ISBN: 0-4719-6243-0.
- [36] Nixon M, Aguado A. Feature extraction and image processing for computer vision. Amsterdam: Elsevier Ltd; 2019. ISBN: 0-1281-4976-0.
- [37] Bradski G, Kaehler A. Learning OpenCV: Computer vision with the OpenCV library. Sebastopol: O'Reilly Media Inc; 2008. ISBN: 0-5965-1613-4.
- [38] Figueira JP. Fast discrete Fréchet distance. 2020. Source: <<https://towardsdatascience.com/fast-discrete-fr%C3%A9chet-distance-d6b422a8fb77>>.
- [39] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv Preprint. 2014. Source: <<https://arxiv.org/abs/1412.6980>>.

Сведения об авторах

Корчагин Сергей Андреевич, 1997 года рождения, учится в Московском государственном университете имени М.В. Ломоносова на факультете вычислительной математики и кибернетики. Работает младшим научным сотрудником в секторе 11.1 «Репродукция и синтез цвета» Института проблем передачи информации имени А.А. Харкевича. Область научных интересов: обработка изображений, машинное обучение, цифровая фотография, спектральный анализ цвета. E-mail: korchagin.sergey.97@gmail.com.

Зайченкова Екатерина Евгеньевна, 2001 года рождения, учится в Московском физико-техническом институте (НИУ) по специальности «Прикладные математика и физика». Работает стажёром-исследователем в секторе 11.1 «Репродукция и синтез цвета» Института проблем передачи информации имени А.А. Харкевича. Область научных интересов: обработка изображений, машинное обучение, цифровая фотография. E-mail: zaychenkova.ee@phystech.edu.

Шарапов Денис Александрович, 2002 года рождения. Учится в Московском физико-техническом институте (НИУ) по специальности «Прикладные математика и физика» на кафедре Института проблем передачи информации в секторе 11.1 «Репродукции и синтеза цвета». Область научных интересов: цифровая фотография, компьютерное зрение, машинное обучение. E-mail: sharapov.d@phystech.edu.

Ершов Егор Иванович, 1990 года рождения, в 2019 получил степень кандидата физико-математических наук по специальности 05.13.18 в Институте проблем передачи информации. Руководит сектором «Репродукции и синтеза цвета» Института проблем передачи информации. E-mail: e.i.ershov@gmail.com.

Буторин Юрий Владимирович, 1976 года рождения. В 1999 году окончил Российский химико-технологический университет им. Д. И. Менделеева. Работает руководителем отдела разработок в ООО «СИНТЭКО-КОМПЛЕКС». Область профессиональных интересов: обработка графических изображений, программирование, автоматизация производственных процессов, организация внутрипроизводственного контроля качества. E-mail: byv@synteco.ru.

Венгеров Юрий Юрьевич, 1991 года рождения, в 2012 году окончил бакалавриат факультета экономики НИУ ВШЭ, в 2017 году окончил магистратуру МГУ по направлению «Инновационный менеджмент», с 2016 работает генеральным директором производителя оборудования для лабораторной диагностики ООО «СИНТЭКО-КОМПЛЕКС». Область научных интересов: клиническая лабораторная диагностика, приборостроение, иммунология, инновационный менеджмент. E-mail: wvengerov@gmail.com.

ГРНТИ: 28.23.15

Поступила в редакцию 10 мая 2023 г. Окончательный вариант – 20 июня 2023 г.

An algorithm of blood typing using serological plate images

S.A. Korchagin^{1,2}, E.E. Zaychenkova^{1,3}, D.A. Sharapov^{1,3}, E.I. Ershov^{1,3}, Y.V. Butorin^{1,2}, Y.Y. Vengerov^{1,2}

¹The Institute for Information Transmission Problems,

127051, Moscow, Russia, Bolshoy Karetny per. 19, build.1;

²Lomonosov Moscow State University, 119991, Moscow, Russia, Leninskie Gory 1;

³The Moscow Institute of Physics and Technology,

141701, Russia, Moscow region, Dolgoprudny, Institutskiy per. 9;

⁴LLC «SYNTECO», 142530, Russia, Moscow region, Elektrogorsk, Budionova 1a

Abstract

This paper describes an in vitro medical express diagnostic system designed to determine the blood group by analyzing the agglutination reaction (gluing of erythrocytes). The medical staff only needs to take a blood sample, put it on a serological plate, placing it in a special scanner for the blood group to be automatically determined. Data digitizing and machine-assisted plate identification allows two critical tasks to be addressed at once: storing the analysis results and controlling the human factor. The proposed recognition algorithm allows the alveolus boundaries to be accurately determined and the agglutination degree to be evaluated using a lightweight convolutional neural network. A unique dataset was collected with the independent assessment of agglutination degree conducted by medical experts. The agglutination estimation accuracy on the collected dataset of 3231 alveole was comparable to the accuracy of an average medical expert and equal to 0.98.

Keywords: agglutination, blood typing, classification, Hough transform, deep learning.

Citation: Korchagin SA, Zaychenkova EE, Sharapov DA, Ershov EI, Butorin UV, Vengerov UU. An algorithm of blood typing using serological plate images. *Computer Optics* 2023; 47(6): 958-967. DOI: 10.18287/2412-6179-CO-1339.

Authors' information

Sergey Andreevich Korchagin, (b. 1997), studies at the Lomonosov Moscow State University of Computational Mathematics and Cybernetics faculty. Works as junior researcher in Sector 11.1 "Reproduction and synthesis of color" of the Institute for Information Transmission Problems named after A.A. Kharkevich. Research interests: image processing, machine learning, digital photography, color spectral analysis. E-mail: korchagin.sergey.97@gmail.com.

Ekaterina Evgenievna Zaychenkova, (b. 2001), studies at the Moscow Institute of Physics and Technology (NRU) with a degree in Applied Mathematics and Physics. Works as a trainee researcher in Sector 11.1 "Reproduction and synthesis of color" of the Institute for Information Transmission Problems named after A.A. Kharkevich. Research interests: image processing, machine learning, digital photography. E-mail: zaychenkova.ee@phystech.edu.

Denis Alexandrovich Sharapov, (b. 2002), studies at the Moscow Institute of Physics and Technology (NRU) with a degree in Applied Mathematics and Physics at the Department of the Institute for Information Transmission Problems in sector 11.1 "Color reproduction and synthesis". Research interests: digital photography, computer vision, machine learning. E-mail: sharapov.d@phystech.edu.

Egor Ivanovich Ershov, (b. 1990), in 2019 received a PhD in Physics and Mathematics, specialty 05.13.18 at the Institute for Information Transmission Problems. Head of the "Reproduction and color synthesis" sector of the Institute for Information Transmission Problems named after A.A. Kharkevich. E-mail: e.i.ershov@gmail.com.

Yury Vladimirovich Butorin, (b. 1976). Graduated from the Russian Chemical-Technological University named after D. I. Mendeleev. Works as the head of the development department at SYNTECO-COMPLEX LLC. Area of professional interests: processing of graphic images, programming, automation of production processes, organization of in-house quality control. E-mail: byv@synteco.ru.

Yury Yuryevich Vengerov, (b. 1991), graduated from the HSE faculty of Economics in 2012, graduated from the Moscow State University in 2017 with a master's degree in Innovation Management. Works as the General Director of Sinteko-Complex LLC. Research interests: clinical laboratory diagnostics, instrumentation, immunology, innovative management. E-mail: wvengerov@gmail.com.

Received May 10, 2023. The final version – June 20, 2023.
