

© Коллектив авторов, 2021

П.А. ВИШНЯКОВА¹, Е.А. КАПРУЛЕВИЧ², А.О. КИРИЛЛОВА¹,
В.В. АНАНЬЕВ², А.Ю. НАУМОВ², Т.Х. ФАТХУДИНОВ³**ПРИМЕНЕНИЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ
В ПАТОМОРФОЛОГИИ И ВСПОМОГАТЕЛЬНЫХ РЕПРОДУКТИВНЫХ
ТЕХНОЛОГИЯХ**¹ФГБУ «Национальный медицинский исследовательский центр акушерства, гинекологии и перинатологии
имени академика В.И. Кулакова» Минздрава России, Москва, Россия²ФГБУН «Институт системного программирования им. В.П. Иванникова» Российской академии наук, Москва, Россия³ФГАОУ ВО «Российский университет дружбы народов», Министерства науки и высшего образования
Российской Федерации, Москва, Россия

Модели машинного обучения повсеместно применяются для анализа изображений, сигналов и видео. На первый взгляд, это хорошо разработанный процесс, сводящийся к этапам сбора данных, разметке, обучению модели и в итоге — ее применению в той или иной области (распознавание автомобильных номеров, лиц в смартфонах и т.д.). Однако в области медицины все гораздо сложнее: применение моделей искусственного интеллекта — серьезный вызов. Методы машинного обучения становятся все более и более используемыми в морфологических науках и биомедицинских исследованиях. Внедрение искусственного интеллекта для анализа изображений позволяет снизить нагрузку на оператора (патолога, гистолога), исключить фактор субъективной оценки и снизить вероятность ошибки. В данном обзоре приводится краткий экскурс в историю возникновения методов машинного обучения, рассматриваются примеры их использования в двух областях, где они получили наибольшее распространение — патоморфологии и вспомогательных репродуктивных технологиях, а также указываются ограничения и сложности, с которыми сталкиваются разработчики при обучении нейронных сетей.

Заключение: Авторами также предлагаются решения для преодоления трудностей связанных со сбором, совместной разметкой данных и обучением моделей: создание инфраструктуры высокого качества, привлечение высококвалифицированных специалистов, размечающих данные, передовой научный подход к технологиям искусственного интеллекта, в качестве основы для масштабируемого хранения и анализа биомедицинских данных предлагается использовать облачные платформы.

Ключевые слова: нейронная сеть, машинное обучение, тайм-лапс микроскопия.

Вклад авторов: Вишнякова П.А., Капрулевич Е.А., Кириллова А.О., Ананьев В.В., Наумов А.Ю., Фатхудинов Т.Х. — анализ литературы, суммирование данных, написание статьи.

Конфликт интересов: Авторы статьи заявляют об отсутствии конфликта интересов.

Финансирование: Работа выполнена в рамках госзадания №121040600436-7. Исследование выполнено при финансовой поддержке РФФИ и Правительства Москвы в рамках научного проекта № 21-315-70048.

Для цитирования: Вишнякова П.А., Капрулевич Е.А., Кириллова А.О., Ананьев В.В.,
Наумов А.Ю., Фатхудинов Т.Х. Применение алгоритмов машинного обучения
в патоморфологии и вспомогательных репродуктивных технологиях.
Акушерство и гинекология. 2021; 10: 38-46
<https://dx.doi.org/10.18565/aig.2021.10.38-46>

©A group of authors, 2021

P.A. VISHNYAKOVA¹, E.A. KAPRULEVICH², A.O. KIRILLOVA¹, V.V. ANANIEV²,
A.YU. NAUMOV², T.KH. FATKHU DINOV³**APPLICATION OF MACHINE LEARNING ALGORITHMS
IN MORPHOPATHOLOGY AND IN ASSISTED REPRODUCTIVE TECHNOLOGIES**¹Academician V.I. Kulakov National Medical Research Center of Obstetrics, Gynecology, and Perinatology,
Ministry of Health of Russia, Moscow, Russia²V.P. Ivannikov Institute for System Programming, Russian Academy of Sciences, Moscow, Russia³Peoples' Friendship University of Russia (RUDN University),

Ministry of Education and Higher Science of the Russian Federation, Moscow, Russia

Machine learning models are used everywhere to analyze images, signals, and videos. At first glance, this is a well-designed process that involves the stages of data collection, mark-up, and training a model, and, as a result, its application in a particular field (recognition of vehicle plate numbers, smartphone faces, etc.). However, everything is much more complicated in the field of medicine: the use of artificial intelligence models is a serious challenge. Machine learning methods are becoming more and more used in morphological sciences and biomedical studies.

The introduction of artificial intelligence for image analysis can lower the burden on an operator (a pathologist, a histologist), eliminate the factor of subjective assessment, and reduce the likelihood of an error.

This review provides a brief excursion into the history of machine learning methods, considers the examples of their use in two areas where they are most widespread: morphopathology and assisted reproductive technologies, and also indicates the limitations and difficulties that developers face when training neural networks.

Conclusion: *The authors also propose solutions to overcome the difficulties associated with the collection and joint marking of data, and model training: creation of a high-quality infrastructure, attraction of highly qualified specialists who mark data, an advanced scientific approach to artificial intelligence technologies; cloud platforms are offered to be used as a basis for the scalable storage and analysis of biomedical data.*

Keywords: *neural network, machine learning, time-lapse microscopy.*

Authors' contributions: Vishnyakova P.A., Kaprulevich E.A., Kirillova A.O., Ananiev V.V., Naumov A.Yu., Fatkhudinov T.Kh. – literature analysis; data summation; writing the article.

Conflicts of interest: The authors declare that there are no conflicts of interest.

Funding: The investigation was conducted within the framework of State Assignment No. 121040600436-7.

This was funded by the Russian Foundation for Basic Research and the Government of Moscow within the framework of Research Project No. 21-315-70048.

For citation: Vishnyakova P.A., Kaprulevich E.A., Kirillova A.O., Ananiev V.V., Naumov A.Yu., Fatkhudinov T.Kh. Application of machine learning algorithms in morphopathology and in assisted reproductive technologies. Akusherstvo i Ginekologiya/Obstetrics and Gynecology. 2021; 10: 38-46 (in Russian) <https://dx.doi.org/10.18565/aig.2021.10.38-46>

Историческая справка

Методы глубокого машинного обучения разработаны еще в 1960-е гг., но из-за низкой производительности компьютеров они не нашли массового применения и вновь проявили себя уже только в 2012 г., когда сверточная нейронная сеть (НС) выиграла у классических конкурсов алгоритмов распознавания изображений [1]. К сожалению, до сих пор данные технологии недостаточно широко применяются для анализа изображений при микроскопической визуализации клеток, тканей и органов в биологии и медицине. Это прежде всего обусловлено техническими сложностями получения большого количества воспроизводимых по качеству микроскопических изображений анализируемых объектов.

Одной из областей, где машинное обучение получило наибольшее распространение, является патоморфология. Сканирование и оцифровка препаратов позволили перейти от техники микроскопирования к визуализации всего препарата целиком в сверхвысоком разрешении, размечать и использовать эти изображения для решения задач классификации, семантической сегментации и детекции объектов с помощью глубоких НС, что дало начало целому направлению – digital pathology. В PubMed при поисковом запросе «histopathology neural network» за последние 5 лет обнаруживаются тысячи статей и с каждым годом их количество растет (https://pubmed.ncbi.nlm.nih.gov/?term=histopathology+neural+network&filter=datesearch.y_5). Это свидетельствует о том, что технические возможности, наконец, достигли необходимого уровня, и данная тематика становится актуальной, но все еще малоисследованной. При этом микроскопическая морфология содержит тысячи различных объектов, характеристики которых можно получать и анализировать с помощью современных компьютеров и программного обеспечения. Морфологические

науки достигли некоторого предела своего развития, а с помощью машинного обучения (и глубоких НС, в частности) морфометрия может приобрести интенсивное развитие благодаря внедрению цифровизации и алгоритмизации исследований в цитологии, гистологии, эмбриологии и патологической анатомии.

Машинное обучение в патоморфологии

Исходным материалом для анализа и машинного обучения в морфологических науках являются микроскопические изображения субклеточных структур, клеток, тканей, органов и даже целых организмов, которые можно получать с помощью электронной, световой, фазово-контрастной, флуоресцентной и других видов микроскопии. При этом необходимо соблюдать два основных условия: 1) изображения должны быть качественно воспроизводимыми и 2) изображений должно быть много (от сотен тысяч, и чем изображений больше, тем точнее анализ). Современные сканеры гистологических препаратов, тайм-лапс микроскопы и пополняющиеся открытые базы данных позволяют получать огромное количество качественных изображений самых разных биологических объектов для обучения НС, но эффективные инструменты анализа таких данных развиваются с существенным отставанием. Использование сверточных НС позволяет разработать подобные инструменты в самых различных областях.

При использовании НС для семантической сегментации гистологических изображений задачи формулируются таким образом, чтобы находить конкретные структурные элементы или морфологические паттерны в изучаемом объекте. Для обучения НС необходимо определить целевые структуры или паттерны структур, несущих или не несущих (по принципу +/-, есть/нет) какой-либо

признак (или признаки), количественная оценка совокупного присутствия которого сигнализирует о нарушении структуры клетки/ткани/органа/организма. К сожалению, такой подход обучения НС по заданным паттернам не позволяет искусственному интеллекту самостоятельно найти какие-либо новые, неизвестные человеку различия между патологическими процессами.

подавляющее большинство публикаций по глубокому обучению в медицине посвящено использованию НС в онкоморфологии – наиболее актуальному и востребованному направлению патологической анатомии. Задача машинного обучения в патологической анатомии сводится к тому, чтобы отличить опухолевую ткань от нормальной и в данной области уже достигнуты некоторые успехи. J. Ker et al. [1] продемонстрировали, что НС без ошибок различает нормальную ткань мозга от глиобластомы, а при разделении нормальной ткани молочной железы от рака *in situ* допускает менее 10% ошибок. Другие авторы [2] добились уровня ошибки менее 3% при разделении доброкачественных и злокачественных опухолей молочной железы. Yan et al. получили схожие результаты: ошибки при дифференцировке злокачественных инвазивных и *in situ* опухолей от нормальных тканей составили 5% и 3%, соответственно, в то время как при определении доброкачественных опухолей их уровень был несколько выше – до 13% [3]. Hekler et al. [4] при дифференциальном разделении невуса и меланомы продемонстрировали точность, составившую 68% и оказавшуюся статистически выше, чем при анализе, проведенном на том же материале патологоанатомом (59%). Таким образом, можно заключить, что с относительно несложными задачами НС справляется не хуже квалифицированного специалиста [5]. Более сложной задачей является сопоставление схожих патологических процессов или объектов, локализующихся в морфологически близких структурах ткани (строме, паренхиме). При сравнении глиомы высокой и низкой степени злокачественности (*high grade* и *low grade*) была получена точность 97,5% [1]. Korbar et al. [6] с помощью НС добились точности в 93% при дифференцировке доброкачественных опухолей толстой кишки и нормальной ткани.

Реже НС используют в диагностике неопухолевых заболеваний. Wei et al. [7] использовали НС для выявления целиакии в биоптатах двенадцатиперстной кишки. С точностью 87% удалось отличать нормальную слизистую оболочку, неспецифическое воспаление и целиакию. Martin et al. [8] показали, что НС выявляет гастрит почти с 100% точностью, и с 90% – реактивную гастропатию. Также НС используются для оценки качества репродуктивных тканей и органов в программах сохранения и восстановления репродуктивной функции. С помощью сверточной НС удалось произвести высокоточную оценку плотности фолликулов в овариальной ткани [9]. Кроме того анализ гистологических изображений с помощью искусственного интеллекта позволяет оценивать патоморфоз эндометрия [10]. Перечисленные работы были проведены на неболь-

ших выборках и безусловно требуют дальнейшей валидации. Также важно отметить, что для отдельного ряда клинических случаев требуется разработка отдельного программного продукта.

Для повышения качества прогностической способности используется технология дообучения предобученной НС. Так, можно получить хорошие результаты даже с относительно небольшим объемом выборки. Для исследователей доступен ряд глубоких НС. Крупные компании в области информационных технологий уже выпустили и выложили в свободный доступ нейросети (ResNet, Inception V3, AlexNet и т.д.), которые предобучены на размеченных объектах реального мира из набора изображений ImageNet [11]. Во многих задачах применение дообучения предобученных НС позволяет существенно повысить точность по сравнению с обучением НС с нуля на имеющемся морфологическом материале.

Большой проблемой обучения НС при анализе микроскопических изображений является подбор, разметка и стандартизация этих изображений, определение зоны интереса, которая на данном этапе не может быть решена без участия человека. НС обучается на изображениях одинакового или разного размера; сканированные гистологические изображения нарезают на квадраты или прямоугольники, при этом патологоанатом выделяет зону интереса. Количественная представленность анализируемого признака в разных изображениях (участках ткани) может существенно различаться, поэтому должно быть проанализировано большое количество изображений, чтобы максимально уменьшить вероятность ошибки. При анализе изображений также применяется аугментация – методика создания дополнительных обучающих материалов из имеющихся данных, позволяющая увеличить количество изображений, тем самым повысить качество обучения НС. Например, используется вращение изображений вокруг центральной оси, искажение изображения для имитации наклона в горизонтальной плоскости, отражение в горизонтальной или вертикальной плоскости, масштабирование и т.д. [2]. Аугментация позволяет увеличить исходный объем данных в разы, но ключевым является не финальное число изображений, а однородность исходных микропрепаратов и количество исходных клинических или экспериментальных объектов.

Благодаря применению современных подходов к обучению НС, получается достаточно хорошо различать как опухолевую, так и неопухолевую патологию, однако доверить НС диагностике сложных случаев пока не получается. Мнение, полученное с помощью НС, специалист может принять во внимание, но решение он примет с учетом комплекса факторов. Получается, что НС обеспечивает поддержку диагностического решения (*diagnostic decision support*): НС может провести морфологический скрининг и все подозрительные случаи передать патологоанатому, чтобы врач определил окончательный морфологический диагноз. Еще одним потенциальным направле-

нием развития является морфометрия, которая может быть выполнена ИС как в клинической диагностике, так и в рамках экспериментальных исследований.

Машинное обучение в программах вспомогательных репродуктивных технологий

Одним из результатов применения машинного обучения являются клинические прогностические модели. Такие модели – это результат применения аналитических методов, обнаруживающих прогностическое влияние различных входных параметров на результаты лечения с определенным уровнем достоверности. В рамках программ вспомогательных репродуктивных технологий (ВРТ) методы машинного обучения применяются для оценки клинических характеристик пациентов, параметров цикла, а также эмбриологических данных с целью предикции исхода ВРТ. Первое подобное исследование с применением искусственных ИС было проведено Kaufmann et al. [12] в 1997 г. На основании 455 циклов ВРТ авторы создали клиническую прогностическую модель, построенную с использованием таких переменных, как возраст пациентки, количество полученных ооцитов, количество перенесенных эмбрионов и факта криоконсервации эмбриона. Точность полученной модели составила 59%. По прошествии 20 лет был накоплен большой массив данных об успехах программ ВРТ в крупных клинических центрах. Техническое развитие позволило реализовывать математические алгоритмы для создания десятков прогностических моделей с различной точностью и различными параметрами входных данных. Существует ряд решений для построения предиктивных моделей в области ВРТ на основе различных моделей машинного обучения, включая: support vector machines (SVM), neural network (NN), k-nearest neighbors (kNN), naive bayes (NB), random forest (RF) и decision tree (DT). Недостатком полученных моделей является их слабая валидация в клинике [13]. Различия в системах сбора данных о пациентах, субъективность вносимой информации или ее недостаточность не позволяет сделать ту или иную разработанную модель универсальной.

Наиболее эффективным подходом ВРТ является проведение стимуляции суперовуляции яичников с получением достаточного количества зрелых ооцитов. Часто это приводит к формированию больше одного эмбриона, пригодного для переноса в полость матки. Таким образом, очень важным становится вопрос селекции эмбрионов для приоритетного переноса и достижения беременности в кратчайшие сроки [14]. Морфологическая оценка была основным критерием отбора эмбрионов на протяжении десятилетий [15]. Оценка в первый день (16–20 часов после оплодотворения) включает оценку пронуклеусов и наличие второго полярного тела как индикаторов правильного оплодотворения. Оценка на третий день – это оценка дробления эмбриона, согласно критериям Европейского

общества репродукции человека и эмбриологии (ESHRE): количество и размер бластомеров, наличие мультинуклеации, фрагментации или вакуолей. Оценка эмбрионов на стадии бластоцист проходит на пятый и шестой день культивирования по системе Гарднера. На основании перечисленных оценок проводится отбор эмбриона с наилучшими морфологическими параметрами для переноса в полость матки.

Тем не менее, несмотря на очевидные имеющиеся корреляции, морфология эмбриона не всегда является надежным и абсолютным показателем его имплантационного потенциала: оптимальная по морфометрии бластоциста не всегда приводит к беременности, тогда как морфологически субоптимальный эмбрион может привести к рождению здорового ребенка.

С развитием технологий преимплантационного генетического скрининга (ПГС) стало ясно, что нельзя всецело полагаться на морфологический анализ при выборе эуплоидного эмбриона. Так, было показано, что нет корреляции между хромосомным статусом эмбриона и морфологической оценкой эмбриона на стадии дробления [16]. Для эмбрионов на стадии бластоцисты есть более высокая вероятность эуплоидии среди морфологически хороших бластоцист, но все же влияние анеуплоидии на качество эмбриона трудноуловимо [16]. Даже эмбрионы с хорошей морфологией могут нести генетические аномалии, и, наоборот, среди эмбрионов низкого качества по морфологии, есть эуплоидные эмбрионы. Преимплантационное генетическое тестирование может повысить эффективность ЭКО для определенной группы пациентов, тем не менее данный подход имеет ряд недостатков. Во-первых, это инвазивный метод, который теоретически может понизить жизнеспособность эмбриона. Во-вторых, ПГС требует дополнительные финансовые затраты со стороны пациенток. В-третьих, данный метод предполагает сегментацию лечебного цикла с криоконсервацией всех эмбрионов и переноса размороженного эмбриона в последующих циклах. Другим перспективным подходом к селекции эмбрионов является оценка их метаболического статуса по анализу среды для культивирования. Данный метод является неинвазивным и основывается на анализе концентраций веществ в среде для культивирования эмбрионов. При данном подходе оценивается, как потребление веществ, так и секреция их в культуральную среду. Было показано, что профили концентраций аминокислот, пирувата и глюкозы коррелируют с имплантационным потенциалом эмбрионов [17–20]. Для оценки метаболического статуса эмбриона также можно применять спектроскопию в ближнем инфракрасном диапазоне. Данный метод основывается на сравнении спектров поглощения среды для культивирования эмбрионов и контрольной среды с последующим вычислением «балла жизнеспособности эмбриона» [21]. Так было продемонстрировано, что балл жизнеспособности коррелирует с имплантационным потенциалом эмбрионов на стадии бластоцисты [22].

Тем не менее, несмотря на успехи в данной области, оценка метаболического статуса эмбриона в настоящее время не вошла в рутинную клиническую практику. Данный метод требует дальнейшей доработки и валидации. Также он имеет ряд ограничений, так как требует крайне стабильных условий культивирования, высокотехнологичного оборудования и разработку специальных алгоритмов анализа данных.

Рутинная процедура оценки эмбрионов в клиниках ЭКО основана на наблюдении с помощью световой микроскопии в установленное время. В двухтысячных годах наряду с классическими показателями морфологии blastocysts, внимание клиницистов привлекли морфокинетические параметры, получаемые с помощью тайм-лапс микроскопии, что дало дополнительные параметры для оценки качества эмбрионов [23]. Тайм-лапс микроскопия позволяет получить важную информацию о развитии эмбрионов, которую зачастую нельзя получить при краткосрочной оценке эмбрионов раз в несколько дней. Система тайм-лапс объединяет три основных элемента: инкубатор, оптический микроскоп и программное обеспечение. Эти элементы обеспечивают постоянное наблюдение за развитием эмбриона при сохранении оптимальных условий культивирования. Маркеры, выявленные тайм-лапс микроскопией, обладают клинической значимостью при их отборе: например, время между оплодотворением и первым дроблением является важным параметром, который коррелирует с определенными прогностическими показателями жизнеспособности эмбрионов [24].

Тайм-лапс данные обладают рядом преимуществ по сравнению с рутинным морфологическим анализом. Они позволяют определить не только точное время до первого дробления и общее число клеточных делений, но и добиться более тщательного морфологического мониторинга, связанного с развитием эмбриона: тайм-лапс обнаруживает такие события, как нерегулярные деления, начало уплотнения или появление blastocell, детекцию мейотического и митотического делений. В исследовании Lemmen et al. [25] было оценено 102 эмбриона в течение 20–24 часов, анализировалось время и координация дробления во время раннего развития зиготы. Авторы отметили, что успешно имплантированные эмбрионы дробились быстрее и более синхронно после первого расщепления и что время исчезновения пронуклеуса на стадии 4-клеток у них было меньше. Значимое преимущество тайм-лапс микроскопии при отборе эмбрионов показано в проспективном когортном исследовании Adamson et al. [26]: из 319 пациентов клиническая беременность при отборе эмбриона с помощью оценки морфокинетических параметров наступила в 46% случаев, в то время как традиционная оценка по стандартной морфологической классификации привела к 32% беременностей.

Метаанализ, проведенный Pribenszky et al. [27], указывает на перспективность и клиническую значимость использования систем визуализации с тайм-лапс технологией в программах ВРТ. Авторы проанализировали имеющиеся данные об исходах

беременностей, где применялся выбор эмбриона по морфокинетическим параметрам, и сделали вывод, что использование этой технологии приводит к снижению процента ранней потери беременности и увеличению продолжительности наступившей беременности. На основании данных нескольких рандомизированных испытаний был получен вывод, что смена рутинной практики стандартного наблюдения в фиксированные моменты времени на непрерывное наблюдение с оценкой морфокинетических параметров обоснована, и данное направление будет активно развиваться.

С момента начала использования тайм-лапс микроскопии в клинической практике было предложено несколько критериев оценки эмбрионов, которые будут подвергнуты трансферу, чтобы достичь наибольшей частоты живорождения. Предлагались следующие критерии: время первого дробления ($14,3 \pm 6,0$ мин); время между первым и вторым митотическим делением или до стадии 2–3 blastomeres ($11,1 \pm 2,2$ ч); и время между вторым и третьим дроблением или до стадии 3–4 blastomeres ($1,0 \pm 1,6$ ч) [28]. Meseguer et al. [29] предложили критерии прогноза успешной имплантации: t_5 время между ICSI и стадии 5 blastomeres ($48,8–56,6$ ч); $t_3–t_2$, время длительности стадии двух blastomeres ($\leq 11,9$ ч); переход к стадии 4 blastomeres ($\leq 0,76$ ч). На основании предложенных критериев Meseguer et al. [29] предложили 10 категорий (A+/A–, B+/B–, C+/C–, D+/D–, E, F) с разной частотой успешной имплантации. Дальнейшие исследования в целом подтвердили соответствие такой классификации эмбрионов и частоты успешной имплантации [24].

Несмотря на увеличивающееся количество исследований применения тайм-лапс микроскопии, ряд авторов критически относятся к данной технологии, аргументируя это тем, что результаты оцениваются по удачной имплантации, а не по частоте живорождений [30]. В более современных системах оценки качества эмбрионов с помощью тайм-лапс микроскопии это учитывается. Так сравнительный анализ TLIA (time-lapse imaging algorithm), разработанный Fishel et al. [31], показал большую селективную силу алгоритма по сравнению с обычными общепринятыми критериями оценки качества blastocyst, которые основаны на морфологических критериях, и использующих в качестве конечной точки живорождение. Несмотря на то, что критерии оценки эмбрионов Гарднера также оказались эффективными для определения потенциала живорождения, они оказались слабее, чем TLIA [32]. Система TLIA базируется на присвоении определенного ранга эмбриону от A до D, которые основываются на таких показателях, как время, прошедшее от оплодотворения до начала дробления, время формирования blastocyst и другие. Для переноса выбирается тот эмбрион, который получил наибольший ранг [31].

В работе [33] также с привлечением тайм-лапс микроскопии показана ценность таких показателей как время формирования morula (time of morulation (tM)) и оценка качества трофобластической оболочки,

которые, по результатам исследования оказались единственными признаками для прогнозирования частоты живорождения, хотя как отмечают сами авторы, время начала формирования бластоцисты также является значимым критерием. Было показано, что если время формирования морулы было меньше 80 ч, а качество трофэктодермы высокое, то частота живорождения составляла порядка 55,2%, а если более 80 ч – 25,5%.

Таким образом, оценка качества эмбрионов на основе тайм-лапс микроскопии является эффективным инструментом [34]. Однако для оптимизации данного метода и внедрения в широкую клиническую практику требуется разработка более специфичных критериев и автоматизированных алгоритмов, позволяющих распознавать эмбрионы, перенос которых закончится живорождением с наибольшей вероятностью, а также дальнейшая валидация таких критериев и протоколов.

Проблемы применения алгоритмов искусственного интеллекта в биомедицине

Проведенный анализ литературы демонстрирует высокую актуальность темы. За последние несколько лет разработаны и широко используются методы оцифровки микроскопических изображений клеток и тканей, но отсутствуют эффективные инструменты анализа и интерпретации получаемых массивов данных и почти отсутствуют валидированные и внедренные в клиническую практику программные продукты.

Причин отсутствия готовых решений с использованием искусственного интеллекта несколько: низкий коэффициент согласия экспертов, повышенные требования к качеству моделей машинного обучения, открытый вопрос ответственности за принятие машинной врачебных решений, недостаточный объем данных в рамках отдельного медицинского учреждения, отсутствие междисциплинарных команд, отсутствие нормативно-правовой базы.

Одним из показательных примеров нетривиальности применения искусственного интеллекта в медицине является отсутствие результатов соизмеримых с затратами у подразделения IBM Watson Health, фирмы разработавшей суперкомпьютер, оснащенный системой искусственного интеллекта (<https://spectrum.ieee.org/biomedical/diagnostics/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care>).

Watson Health и ряд других не таких известных компаний предприняли попытки применить искусственный интеллект к медицине путем переноса методов машинного обучения с изображений бытовых предметов на биомедицинский домен. Неуспех подобных инициатив говорит о том, что в области присутствует скрытая сложность. Одна из ключевых проблем применения методов искусственного интеллекта к медицинскому домену – вероятностная природа глубоких НС. Какой бы качественной ни была НС всегда есть вероятность, что алгоритм не определит больного человека и в этот момент возникает как вопрос ответственности, так и ряд других

вопросов, которые не позволяют широко внедрить искусственный интеллект во все сферы медицины.

Однако существует область систем поддержки принятия врачебных решений, аудита врачебных решений и скрининговых исследований, «домедицина», в которых кажется возможным постепенное внедрение высококачественных моделей искусственного интеллекта. Качество моделей тем не менее остается важным вопросом, так как в случае скрининга большого количества пациентов большое количество ложноположительных срабатываний модели приведет к перегрузке врачей.

Для качественного обучения моделей прежде всего нужен большой объем качественных обучающих данных. Зачастую одно подразделение или даже медицинский центр целиком не может обеспечить достаточное количество данных и/или их качественную разметку врачами.

На успех применения моделей машинного обучения влияет не только качество собранных данных, но и качество разметки. Например, в классическом data science для обучения НС распознаванию объектов реального мира по изображениям котиков, собак и самолетов привлекаются люди без специализированного образования, и стоимость такой разметки сравнительно невысока. В свою очередь, в медицинском домене стоимость качественной разметки врачами-экспертами крайне высока. К тому же, даже у экспертов зачастую наблюдается невысокая степень согласия [35]. Вопросам согласия врачей посвящаются целые исследования [36]. При этом экономия на разметке данных экспертами низкой квалификации обычно не приводит к получению модели достаточного качества.

Дополнительные сложности при обучении моделей искусственного интеллекта могут возникнуть при использовании приборов разных производителей, ручных методов работы с образцами, например, при ручной окраске иммуногистохимических препаратов. Также сложности возникают при попытке обработки данных с неструктурированными диагнозами на естественном языке и зачастую с различными вариантами сокращений.

Инфраструктура и передовые методы для анализа медицинских данных

Для того, чтобы обучать модели искусственного интеллекта, применимые на практике, требуется:

- **Высокое качество инфраструктуры.** Одной из основных тенденций в области развития IT-инфраструктуры является использование облачных технологий. Облачные платформы позволяют разделить зоны ответственности на обслуживание низкоуровневой базовой платформы и на сопровождение конечных сервисов с использованием технологий получения виртуализированных ресурсов по запросу.
- **Передовой научный подход к сбору и хранению данных.** Необходимо учитывать, что при росте требований к качеству обученной модели на основе глубоких НС объем обучающих данных растет.

Также требуется заложить в системе соответствующий уровень безопасности хранения большого объема персональных данных пациентов.

- **Передовой научный подход к технологиям искусственного интеллекта** позволяет эффективно обрабатывать медицинские данные и извлекать из них смысл. Один из современных многообещающих подходов — это обучение с частичным привлечением учителя: модель, предобученная на большом количестве размеченных данных, дообучается на небольшом количестве качественно размеченных данных [37].
- **Высокая квалификация специалистов, размечающих данные** (интеграция врачей в процесс). На основе некачественно размеченных данных зачастую невозможно обучить модель, которая обеспечит приемлемое соотношение сигнал/шум. Качество разметки данных определяется, помимо прочего, степенью согласия экспертов.

Реализация всех этих процессов представляет собой сложную задачу, причем даже для такой хорошо изученной области, как патоморфологические исследования. Помимо всех вышеупомянутых пунктов для эффективной работы желательны также наличие единой специализированной системы разметки данных, так как отображение гистологических изображений в обычных системах для разметки изображений затруднительно из-за их размера. Отдельная проблема — запись заключений на естественном языке. Помимо различных сокращений состав заключения может существенно отличаться из-за различия врачебных школ и терминологии как в мире, так и внутри России в зависимости от региона. Для обучения моделей, способных анализировать гистологические изображения, в научных статьях используются сотни, а иногда и тысячи размеченных изображений.

Заключение

Обучение моделей искусственного интеллекта для создания качественных программных продуктов, применимых на практике, требует принципиально новых подходов. Для создания и внедрения такого рода продуктов, обеспечивающих здоровье и долголетие людей посредством развития предиктивной, превентивной и персонализированной медицины, требуются как использование современной инфраструктуры хранения и анализа данных (облачные технологии), так и передовые методы, которые позволяют минимизировать количество разметки (дообучение нейронной сети, адаптация к домену, контрастное обучение) (рисунок).



Литература/References

1. Ker J., Bai Y., Lee H.Y., Rao J., Wang L. Automated brain histology classification using machine learning. *J. Clin. Neurosci.* 2019; 66: 239-45. <https://dx.doi.org/10.1016/j.jocn.2019.05.019>.
2. Alom M.Z., Yakopcic C., Nasrin M.S., Taha T.M., Asari V.K. Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network. *J. Digit. Imaging.* 2019; 32(4): 605-17. <https://dx.doi.org/10.1007/s10278-019-00182-7>.
3. Yan R., Ren F., Wang Z., Wang L., Zhang T., Liu Y. et al. Breast cancer histopathological image classification using a hybrid deep neural network. *Methods.* 2019; 173: 52-60. <https://dx.doi.org/10.1016/j.ymeth.2019.06.014>.
4. Hekler A., Utikal J.S., Enk A.H., Solass W., Schmitt M., Klode J. et al. Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. *Eur. J. Cancer.* 2019; 118: 91-6. <https://dx.doi.org/10.1016/j.ejca.2019.06.012>.
5. Hannun A.Y., Rajpurkar P., Haghpanahi M., Tison G.H., Bourn C., Turakhia M.P., Ng A.Y. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat. Med.* 2019; 25: 65-9. <https://dx.doi.org/10.1038/s41591-018-0268-3>.
6. Korbar B., Olofson A., Mirafior A., Nicka C., Suriawinata M., Torresani L. et al. Deep learning for classification of colorectal polyps on whole-slide images. *J. Pathol. Inform.* 2017; 8: 30. https://dx.doi.org/10.4103/jpi.jpi_34_17.
7. Wei J.W., Jackson C.R., Ren B., Suriawinata A.A., Hassanpour S. Automated detection of celiac disease on duodenal biopsy slides: A deep learning approach. *J. Pathol. Inform.* 2019; 10: 7. https://dx.doi.org/10.4103/jpi.jpi_87_18.
8. Martin D.R., Hanson J.A., Gullapalli R.R., Schultz F.A., Sethi A., Clark D.P. A deep learning convolutional neural network can recognize common patterns of injury in gastric pathology. *Arch. Pathol. Lab. Med.* 2020; 144(3): 370-8. <https://dx.doi.org/10.5858/arpa.2019-0004-OA>.
9. İnik Ö., Ceyhan A., Balçioğlu E., Ülker E. A new method for automatic counting of ovarian follicles on whole slide histological images based on convolutional neural network. *Comput. Biol. Med.* 2019; 112: 103350. <https://dx.doi.org/10.1016/j.combiomed.2019.103350>.
10. Sun H., Zeng X., Xu T., Peng G., Ma Y. Computer-aided diagnosis in histopathological images of the endometrium using a convolutional neural network and attention mechanisms. *IEEE J. Biomed. Health Inform.* 2020; 24(6): 1664-76. <https://dx.doi.org/10.1109/JBHI.2019.2944977>.
11. Deng J., Dong W., Socher R., Li L.-J., Li K., Fei-Fei L. ImageNet: A large-scale hierarchical image database. In: 2009 IEEE Conference on computer vision and pattern recognition. Miami, FL, USA 20-25 June 2009: 248-55. <https://dx.doi.org/10.1109/cvpr.2009.5206848>.
12. Kaufmann S.J., Eastaugh J.L., Snowden S., Smye S.W., Sharma V. The application of neural networks in predicting the outcome of in-vitro fertilization. *Hum. Reprod.* 1997; 12(7): 1454-7. <https://dx.doi.org/10.1093/humrep/12.7.1454>.

13. Raef B., Ferdousi R. A review of machine learning approaches in assisted reproductive technologies. *Acta Inform. Med.* 2019; 27(3): 205-11. <https://dx.doi.org/10.5455/aim.2019.27.205-211>.
14. Balaban B., Brison D., Calderón G., Catt J., Conaghan J., Cowan L. et al. Istanbul consensus workshop on embryo assessment: Proceedings of an expert meeting, *Reprod. Biomed. Online.* 2011; 22(6): 632-46. <https://dx.doi.org/10.1016/j.rbmo.2011.02.001>.
15. Cummins J.M., Breen T.M., Harrison K.L., Shaw J.M., Wilson L.M., Hennessey J.F. A formula for scoring human embryo growth rates in vitro fertilization: its value in predicting pregnancy and in comparison with visual estimates of embryo quality. *J. In Vitro Fert. Embryo Transf.* 1986; 3(5): 284-95. <https://dx.doi.org/10.1007/bf01133388>.
16. Fragouli E., Alfarawati S., Spath K., Wells D. Morphological and cytogenetic assessment of cleavage and blastocyst stage embryos. *Mol. Hum. Reprod.* 2014; 20(2): 117-26. <https://dx.doi.org/10.1093/MOLEHR/GAT073>.
17. Gardner D.K., Lane M., Stevens J., Schoolcraft W.B. Noninvasive assessment of human embryo nutrient consumption as a measure of developmental potential. *Fertil. Steril.* 2001; 76(6): 1175-80. [https://dx.doi.org/10.1016/S0015-0282\(01\)02888-6](https://dx.doi.org/10.1016/S0015-0282(01)02888-6).
18. Leese H.J. Metabolism of the preimplantation embryo: 40 years on. *Reproduction.* 2012; 143(4): 417-27. <https://dx.doi.org/10.1530/REP-11-0484>.
19. Сысоева А.П., Макарова Н.П., Калинина Е.А., Скибина Ю.С., Занишевская А.А., Янчук Н.О., Грязнов А.Ю. Повышение эффективности вспомогательных репродуктивных технологий с помощью искусственного интеллекта и машинного обучения на эмбриологическом этапе. *Акушерство и гинекология.* 2020; 7: 28-36. [Sysoeva A.P., Makarova N.P., Kalinina E.A., Skibina Yu.S., Zanishevskaya A.A., Yanchuk N.O., Gryaznov A.Yu. Enhancing the efficiency of assisted reproductive technologies using artificial intelligence and machine learning at the embryological stage. *Obstetrics and Gynecology.* 2020; 7: 28-36. (in Russian)]. <https://dx.doi.org/10.18565/aig.2020.7.28-36>.
20. Валиахметова Э.З., Кулакова Е.В., Скибина Ю.С., Грязнов А.Ю., Сысоева А.П., Макарова Н.П., Калинина Е.А. Неинвазивное тестирование преимплантационных эмбрионов человека in vitro как способ прогнозирования исходов программ экстракорпорального оплодотворения. *Акушерство и гинекология.* 2021; 5: 5-16. [Valiakhmetova E.Z., Kulakova E.V., Skibina Yu.S., Gryaznov A.Yu., Sysoeva A.P., Makarova N.P., Kalinina E.A. Non-invasive testing of human preimplantation embryos in vitro as a way to predict the outcomes of in vitro fertilization programs. *Obstetrics and Gynecology.* 2021; 5: 5-16. (in Russian)]. <https://dx.doi.org/10.18565/aig.2021.5.5-16>.
21. Montag M., Toth B., Strowitzki T. New approaches to embryo selection, *Reprod. Biomed. Online.* 2013; 27(5): 539-46. <https://dx.doi.org/10.1016/j.rbmo.2013.05.013>.
22. Ahlström A., Wikland M., Rogberg L., Barnett J.S., Tucker M., Hardarson T. Cross-validation and predictive value of near-infrared spectroscopy algorithms for day-5 blastocyst transfer. *Reprod. Biomed. Online.* 2011; 22(5): 477-84. <https://dx.doi.org/10.1016/j.rbmo.2011.01.009>.
23. Johnson M.H., Day M.L. Egg timers: how is developmental time measured in the early vertebrate embryo? *Bioessays.* 2000; 22: 57-63. [https://dx.doi.org/10.1002/\(SICI\)1521-1878\(200001\)22:1<57::AID-BIES10>3.0.CO;2-L](https://dx.doi.org/10.1002/(SICI)1521-1878(200001)22:1<57::AID-BIES10>3.0.CO;2-L).
24. Castelló D., Motato Y., Basile N., Remohí J., Espejo-Catena M., Meseguer M. How much have we learned from time-lapse in clinical IVF? *Mol. Hum. Reprod.* 2016; 22(10): 719-27. <https://dx.doi.org/10.1093/MOLEHR/GAW056>.
25. Lemmen J.G., Agerholm I., Ziebe S. Kinetic markers of human embryo quality using time-lapse recordings of IVF/ICSI-fertilized oocytes. *Reprod. Biomed. Online.* 2008; 17(3): 385-91. [https://dx.doi.org/10.1016/S1472-6483\(10\)60222-2](https://dx.doi.org/10.1016/S1472-6483(10)60222-2).
26. Adamson G.D., Abusief M.E., Palao L., Witmer J., Palao L.M., Gvakharia M. Improved implantation rates of day 3 embryo transfers with the use of an automated time-lapse-enabled test to aid in embryo selection. *Fertil. Steril.* 2016; 105(2): 369-75.e6. <https://dx.doi.org/10.1016/j.fertnstert.2015.10.030>.
27. Pribenszky C., Nilselid A.M., Montag M. Time-lapse culture with morphokinetic embryo selection improves pregnancy and live birth chances and reduces early pregnancy loss: a meta-analysis. *Reprod. Biomed. Online.* 2017; 35(5): 511-20. <https://dx.doi.org/10.1016/j.rbmo.2017.06.022>.
28. Wong C.C., Loewke K.E., Bossert N.L., Behr B., De Jonge C.J., Baer T.M., Pera R.A.R. Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. *Nat. Biotechnol.* 2010; 28(10): 1115-21. <https://dx.doi.org/10.1038/nbt.1686>.
29. Meseguer M., Herrero J., Tejera A., Hilligsoe K.M., Ramsing N.B., Remohí J. The use of morphokinetics as a predictor of embryo implantation. *Hum. Reprod.* 2011; 26(10): 2658-71. <https://dx.doi.org/10.1093/humrep/der256>.
30. Polanski L.T., Coelho Neto M.A., Nastri C.O., Navarro P.A., Ferriani R.A., Raine-Fenning N., Martins W.P. Time-lapse embryo imaging for improving reproductive outcomes: systematic review and meta-analysis. *Ultrasound Obstet. Gynecol.* 2014; 44(4): 394-401. <https://dx.doi.org/10.1002/uog.13428>.
31. Fishel S., Campbell A., Foad F., Davies L., Best L., Davis N. et al. Evolution of embryo selection for IVF from subjective morphology assessment to objective time-lapse algorithms improves chance of live birth. *Reprod. Biomed. Online.* 2020; 40: 61-70. <https://dx.doi.org/10.1016/j.rbmo.2019.10.005>.
32. Gardner D.K., Lane M., Stevens J., Schlenker T., Schoolcraft W.B. Blastocyst score affects implantation and pregnancy outcome: Towards a single blastocyst transfer. *Fertil. Steril.* 2000; 73(6): 1155-8. [https://dx.doi.org/10.1016/S0015-0282\(00\)00518-5](https://dx.doi.org/10.1016/S0015-0282(00)00518-5).
33. Rienzi L., Cimadomo D., Delgado A., Minasi M.G., Fobozzi G., del Gallego R. et al. Time of morulation and trophectoderm quality are predictors of a live birth after euploid blastocyst transfer: a multicenter study. *Fertil. Steril.* 2019; 112(6): 1080-93.e1. <https://dx.doi.org/10.1016/j.fertnstert.2019.07.1322>.
34. Романов А.Ю., Ковальская Е.В., Макарова Н.П., Сыркашева А.Г., Долгушина Н.В. Использование цейтраферной съемки для оценки качества эмбрионов человека в программах экстракорпорального оплодотворения. *Цитология.* 2017; 59(7): 462-6. [Romanov A.Yu., Kovalskaya E.V., Makarova N.P., Syrkasheva A.G., Dolgushina N.V. Using time-lapse photography to assess the quality of human embryos in in vitro fertilization programs. *Cytology.* 2017; 59(7): 462-6. (in Russian)].
35. Benchoufi M., Matzner-Lober E., Molinari N., Jannot A.S., Soyser P. Interobserver agreement issues in radiology. *Diagn. Interv. Imaging.* 2020; 101(10): 639-41. <https://dx.doi.org/10.1016/j.diii.2020.09.001>.
36. Kang J.H., Choi S.H., Lee J.S., Park S.H., Kim K.W., Kim S.Y. et al. Interreader agreement of liver imaging reporting and data system on MRI: a systematic review and meta-analysis. *J. Magn. Reson. Imaging.* 2020; 52(3): 795-804. <https://dx.doi.org/10.1002/jmri.27065>.
37. Lu M.Y., Chen R.J., Wang J., Dillon D., Mahmood F. Semi-supervised histology classification using deep multiple instance learning and contrastive predictive coding. Available at: <http://arxiv.org/abs/1910.10825> Accessed June 10, 2021.

Поступила 11.06.2021

Принята в печать 03.08.2021

Received 11.06.2021

Accepted 03.08.2021

Сведения об авторах:

Вишнякова Полина Александровна, к.б.н., с.н.с. лаборатории регенеративной медицины, Национальный медицинский исследовательский центр акушерства, гинекологии и перинатологии им. академика В.И. Кулакова Минздрава России, p_vishnyakova@orapina4.ru, 117997, Россия, Москва, ул. Академика Опарина, д. 4.

Карпелевич Евгений Андреевич, н.с. отдела «Информационные системы», Институт системного программирования им. В.П. Иванникова РАН, karpulevich@ispras.ru, 109004, Россия, Москва, ул. Александра Солженицына, д. 25.

Кириллова Анастасия Олеговна, к.б.н., с.н.с. I-го гинекологического отделения, Национальный медицинский исследовательский центр акушерства, гинекологии и перинатологии им. академика В.И. Кулакова Минздрава России, stasia.kozyreva@gmail.com, 117997, Россия, Москва, ул. Академика Опарина, д. 4.

Ананьев Владислав Валерьевич, программист отдела «Информационные системы», Институт системного программирования им. В.П. Иванникова РАН, novisp53@ispras.ru, 109004, Россия, Москва, ул. Александра Солженицына, д. 25.

Наумов Антон Юрьевич, стажер-исследователь отдела «Информационные системы», Институт системного программирования им. В.П. Иванникова РАН, anton-naumov@yandex.ru, 109004, Россия, Москва, ул. Александра Солженицына, д. 25.

Фатхудинов Тимур Хайсамудинович, д.м.н., заместитель директора, Научно-исследовательский институт морфологии человека РАН; заведующий кафедрой гистологии, цитологии и эмбриологии, заместитель директора по научной работе Медицинского института, РУДН, tfat@yandex.ru, 117997, Россия, Москва, ул. Миклухо-Маклая, д. 8.

Authors' information:

Polina A. Vishnyakova, PhD, Senior Researcher, Laboratory of Regenerative Medicine, Academician V.I. Kulakov National Medical Research Center for Obstetrics, Gynecology and Perinatology, Ministry of Health of the Russian Federation, p_vishnyakova@oparina4.ru, 117997, Russia, Moscow, Ac. Oparina str., 4.

Evgeniy A. Karpulevich, Researcher, Department of Information Systems, V.P. Ivannikov Institute for System Programming, Russian Academy of Sciences, karpulevich@ispras.ru, 109004, Russia, Moscow, Alexander Solzhenitsyn str., 25.

Anastasia O. Kirillova, PhD, Senior Researcher of the 1st Gynecological Department, Academician V.I. Kulakov National Medical Research Center for Obstetrics, Gynecology and Perinatology, Ministry of Health of the Russian Federation, stasia.kozyreva@gmail.com, 117997, Russia, Moscow, Ac. Oparina str., 4.

Vladislav V. Ananiev, programmer, Department of Information Systems, V.P. Ivannikov Institute for System Programming, Russian Academy of Sciences, novisp53@ispras.ru, 109004, Russia, Moscow, Alexander Solzhenitsyn str., 25.

Anton Yu. Naumov, Research Assistant, Department of Information Systems, V.P. Ivannikov Institute for System Programming, Russian Academy of Sciences, anton-naumov@yandex.ru, 109004, Russia, Moscow, Alexander Solzhenitsyn str., 25.

Timur Kh. Fatkhudinov, Dr. Med. Sci., Deputy Director, Research Institute of Human Morphology of the Russian Academy of Sciences; Head of the Department of Histology, Cytology and Embryology, Deputy Director for Research of the Medical Institute, RUDN University, tfat@yandex.ru, 117997, Russia, Moscow, Miklukho-Maklaya str., 8.