

*В.Р. Хонов, А.П. Кукареко*

**ОБРАЗОВАТЕЛЬНО-ИССЛЕДОВАТЕЛЬСКАЯ ИНТЕГРИРОВАННАЯ ПЛАФОРМА ПОДДЕРЖКИ ПРИНЯТИЯ ВРАЧЕБНЫХ РЕШЕНИЙ НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ГИСТОЛОГИЧЕСКИХ ДАННЫХ. СОЗДАНИЕ НЕЙРОСЕТЕВОЙ МОДЕЛИ ДЛЯ ОЦЕНКИ МОРФОЛОГИЧЕСКИХ ИЗМЕНЕНИЙ ЛЕГКИХ ПРИ КОРОНАВИРУСНОЙ ИНФЕКЦИИ COVID-19**

*Научные руководители: канд. мед. наук, доц. В.С. Гайдук,  
канд. мед. наук, доц. Т.А. Бич*

*Кафедра гистологии, цитологии и эмбриологии,  
кафедра патологической анатомии*

*Белорусский государственный медицинский университет, г. Минск  
Белорусский государственный университет информатики и  
радиоэлектроники, г. Минск*

*V.R. Khonov, A.P. Kukareko*

**EDUCATIONAL AND RESEARCH INTEGRATED PLANFORM FOR SUPPORTING MEDICAL DECISION-MAKING BASED ON INTELLIGENT ANALYSIS OF HISTOLOGICAL DATA. CREATING A NEURAL NETWORK MODEL FOR ASSESSING MORPHOLOGICAL CHANGES IN THE LUNGS DURING COVID-19 CORONAVIRUS INFECTION**

*Tutors: PhD, associate professor V. S. Gayduk,  
PhD, associate professor T. A. Bich*

*Department of Histology, Cytology and Embryology,  
Department of Pathological Anatomy*

*Belarusian State Medical University, Minsk  
Belarusian State University of Informatics and Radioelectronics, Minsk*

**Резюме.** Разработана модель с использованием компьютерного анализа изображений и подходы машинного обучения для описания патоморфологических изменений на оцифрованных аутопсийных изображениях тканей легкого у пациентов с COVID-19.

**Ключевые слова:** COVID-19, цифровая патология, машинное обучение.

**Resume.** In our model, we used computer image analysis and machine learning approaches to describe pathomorphological changes in digitized autopsy images of lung tissue in patients with COVID-19.

**Keywords:** COVID-19, computational pathology, machine learning.

**Актуальность.** Последние достижения в области компьютерного зрения, появление возможностей быстрой оцифровки гистологических препаратов повысили интерес к приложениям искусственного интеллекта (ИИ) для патологии. Современные алгоритмы способны, к примеру, обнаруживать рак, опухоли, помогать в постановке диагноза пациенту [1]. Многие алгоритмы полагаются только на полнослайдовое изображение (WSI) и метаданные, некоторые также требуют аннотаций клеток, тканей и других объектов на WSI. Эта дополнительная информация помогает связать гистологические изображения с лежащей в их основе

биологией патологических изменений и может улучшить интерпретируемость и обобщаемость полученных алгоритмов [2]. Легкие как орган дыхательной системы играют важную роль в поддержании гомеостаза организма, обеспечивая газообмен, поддерживая кислотно-щелочной баланс крови. Нарушение этих ключевых функций вследствие патологических состояний может приводить к необратимым для организма последствиям. Так, наиболее частой причиной смерти пациентов с инфекцией COVID-19 был респираторный дистресс-синдром взрослого типа (РДСВ), морфологически проявляющийся диффузным альвеолярным повреждением (ДАП). Создание алгоритма оценки морфологических изменений легких при инфекции COVID-19 с помощью подходов компьютерного анализа позволяет внедрить элементы цифровой патологии в научно-исследовательскую работу, а также образовательный процесс.

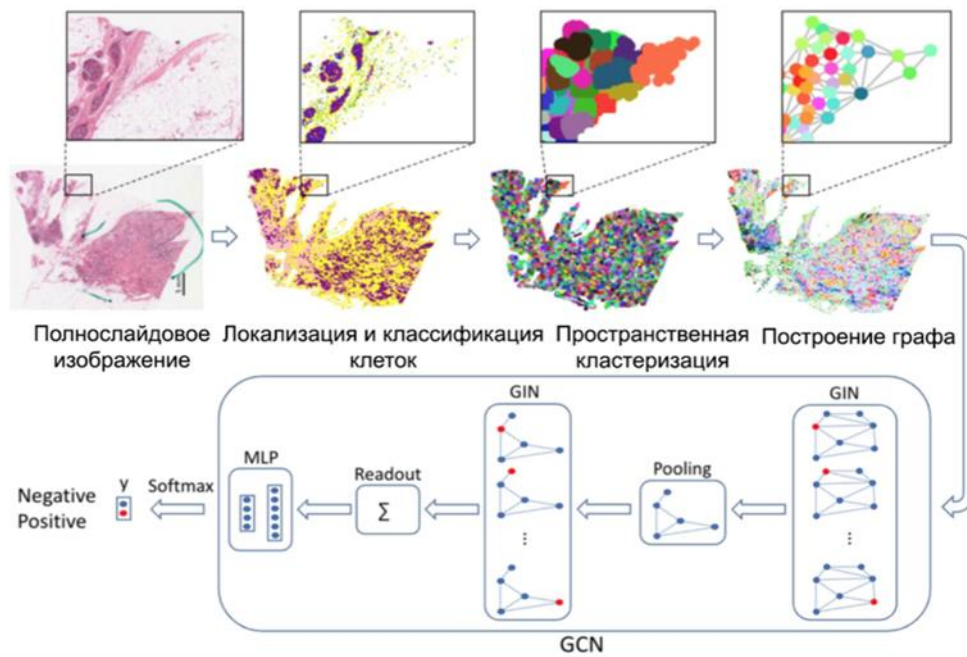
**Цель:** разработать цифровую образовательно-исследовательскую платформу для анализа гистологических данных, создать на ее базе модель нейронной сети, способную графически определять патоморфологические изменения, выявлять количественные и пространственные закономерности иммунного ответа в лёгких при коронавирусной инфекции COVID-19 для дальнейшей интеграции программы в научно-образовательный и практический процессы.

**Задачи:**

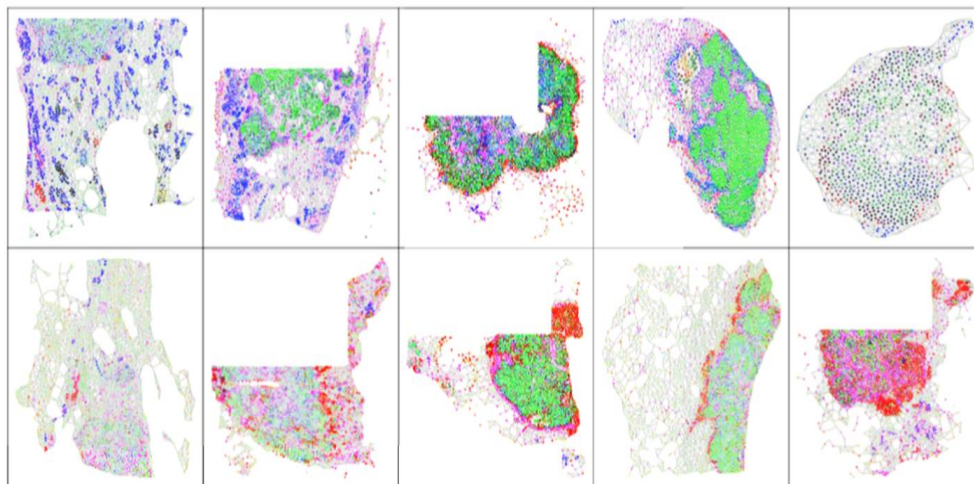
1. Сформировать целевую выборку (датасеты) для обучения модели на множественных примерах патологий.
2. Выбрать оптимальную архитектуру для выполнения задачи и обучить модель-классификатор.
3. Исследовать точность полученной модели.
4. Разработать программный комплекс, интегрирующий ИИ и систему графической аннотации.

**Материалы и методы.** Набор данных состоял из полнослайдовых изображений (WSI) гистологических препаратов легких от 10 пациентов, умерших от COVID-19. Препараты окрашены гематоксилином и эозином. Данные были предоставлены открытой международной научной базой Национального института здоровья США (NIH) и Objective Whole Slide Image Server.

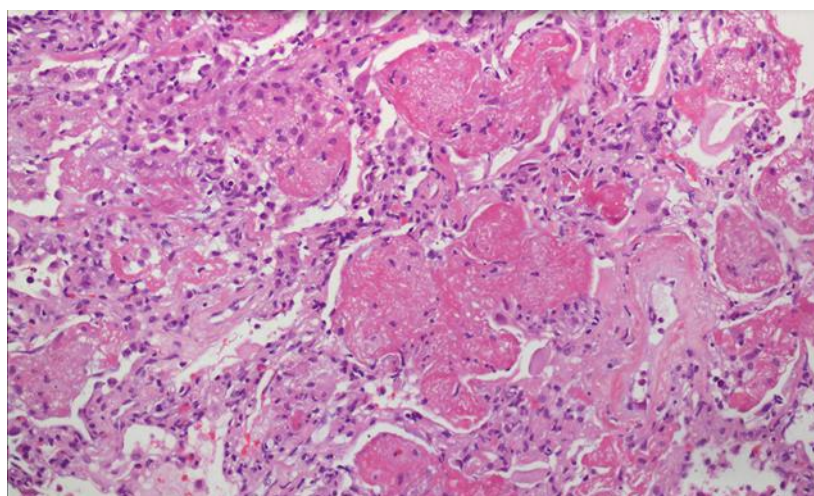
Из данного датасета случайным образом были выделены 200 фрагментов размером  $1024 \times 882$  пикселей. Все изображения были нормализованы в цвете и увеличении. С помощью инструмента аннотирования LabelImg была произведена графическая разметка объектов. Далее, обработанные данные были использованы в качестве датасета для обучения многоуровневой архитектуры свёрточных нейронных сетей (линейный выпрямитель с «утечкой») [3].



**Рис. 1** – Этапы классификации



**Рис. 2** – Визуализация кластеров



**Рис. 3** – Пример полнослайдового изображения из датасета: острое повреждение легких с внутриальвеолярными гиалиновыми мембранами, фибриновыми пробками и интерстициальным воспалением



**Результаты и их обсуждение.** Чтобы решить эту проблему, была разработана система интегрированная платформа поддержки принятия врачебных решений (ИПППВР), позволяющая создавать эффективные и высококачественные аннотации в любом масштабе, производить обмен данными, использовать всю информацию для обучения как моделей глубокого обучения, так и пользователей системы (студентов, исследователей). ИПППВР состоит из трех программных компонентов — модели сегментации, модели классификатора и модели активного ученика, — которые совместно поддерживают пользователя-аннотатора платформы. Модель сегментации определяет расположение всех клеток в небольшой области. Аннотатор, обычно квалифицированный патологоанатом, начинает маркировать клетки. Эти метки используются для обучения модели классификатора, которая идентифицирует и помечает области, которые может впоследствии скорректировать аннотатор. По мере того, как классификатор учится на этих исправлениях, его прогнозы постепенно улучшаются и требуют меньшего количества исправлений. Когда интересующие патологии достаточно размечены (20–30 аннотаций на класс), модель активного ученика определяет следующую наиболее информативную область для аннотации, и процесс повторяется. Так, очевидны преимущества данного подхода. Во-первых, модель классификатора ИПППВР снижает рабочую нагрузку, заменяя трудоемкую задачу первичных аннотаций более простой задачей исправления аннотаций. Во-вторых, модель активного ученика ИПППВР повышает эффективность работы с данными, указывая аннотатору на более информативные области на WSI. Чтобы продемонстрировать эти преимущества, были произведены измерения рабочей нагрузки (доля предложений ИИ, требующих исправления) и эффективность обработки данных (точность проверки по сравнению с размером выборки) для семи патологов, использующих ИПППВР для маркировки морфологических изменений легких при коронавирусной инфекции COVID-19. Использование ИПППВР привело к значительному снижению рабочей нагрузки для всех задач, от 96% для разметки патологических клеток (самая простая задача) до 83% для эозинофилов. Получается, что патологоанатом, вооруженный ИПППВР, создаст примерно в 10 раз больше аннотаций за тот же промежуток времени. В дополнение к этим объективным улучшениям расширение ИИ также позволяет патологоанатомам отслеживать улучшения классификатора и выявлять ошибки на протяжении всего процесса разметки. Это может способствовать более тесному сотрудничеству с инженерами-программистами и способствовать пониманию недостатков модели и распределению ресурсов для маркировки [4]. Подобные системы широко представлены как коммерческие решения и используются в таких областях, как распознавание объектов. В итоге нашей работы была создана нейронная сеть с возможностью определения следующих патоморфологических изменений: тромбы в просветах сосудов различных видов (смешанный, организующийся, фибриновый), серозный и фибринозно-гнойный экссудат, утолщение стенок альвеол, геморрагии. В области иммунного ответа нейронная сеть способна различать клетки иммунной и "неиммунной" природы [5]. Общая точность классификации составила 83,1% и 79,7% по данным обучения и независимого тестирования датасетов соответственно. Идентифицируя клетки и классифицируя типы клеток, нейросеть может преобразовать изображение патологии в «пространственную карту» клеток [6]. Из

этой пространственной карты мы можем извлечь признаки, характеризующие микроокружение исследуемых объектов. На основе этих особенностей разработана прогностическая модель классификатора.

#### **Выводы:**

1. Созданная нейросетевая модель позволяет классифицировать клетки иммунного ответа и определять ограниченный спектр морфологических изменений в легких.
2. Разработанная образовательно-исследовательская программная платформа позволила значительно оптимизировать рабочие процессы в практике патологоанатомов, в частности аннотацию и разметку, платформа представляет новые возможности обучения как студентов, так и нейросетевых моделей в направлении патологической анатомии.

#### **Литература**

1. WHO, "Coronavirus disease (COVID-19) Situation Report-111," 2020.
2. Dosovitskiy A, Fischer P, Springenberg JT, Riedmiller M, Brox T. Discriminative unsupervised feature learning with exemplar convolutional neural networks. *IEEE Trans Pattern Anal Mach Intell* 2016;38(9):1734-47.
3. Snead DR, Tsang YW, Meskiri A, et al. Validation of digital pathology imaging for primary histopathological diagnosis. *Histopathology* 2016; 68(7):1063–72.
4. Bankhead P, Loughrey MB, Fernandez JA, et al. QuPath: open source software for digital pathology image analysis. *Sci Rep* 2017; 7(1):16878.
5. Wang S, Yang DM, Rong R, Zhan X, Xiao G. Pathology image analysis using segmentation deep learning algorithms. *Am. J. Pathol.* 2019
6. Nienhold R, Ciani Y, Koelzer VH, et al. Two distinct immunopathological profiles in autopsy lungs of COVID-19. *Nat Commun* 2020; 11:5086.