

<https://doi.org/10.60043/2949-5938-2025-1-31-40>



# Полуавтоматический метод аннотирования фазово-контрастных изображений живых культур клеток для сегментации ядер на основе машинного обучения

М.В. Балясин<sup>1,2,\*</sup>, А.Г. Демченко<sup>2</sup>, А.В. Люндуп<sup>1,2</sup>

<sup>1</sup> ФГАОУ ВО «Российский университет дружбы народов имени Патриса Лумумбы», Россия, 117198, Москва, ул. Миклухо-Маклая, д. 6

<sup>2</sup> ФГБНУ «Медико-генетический научный центр имени академика Н.П. Бочкова», Россия, 115522, Москва, ул. Москворечье, д. 1

\* Адрес для корреспонденции: [balyasin\\_mv@rudn.ru](mailto:balyasin_mv@rudn.ru)

## Аннотация

В текущей работе разработан комплексный подход для определения ядер живых клеток на изображениях без флуоресцентных меток. Поскольку в клеточной биологии актуальными являются подсчет клеток, оценка динамики роста клеток и конfluентности, то существует целесообразность в автоматизации получения этих данных. Для автоматизации применяют алгоритмы на основе машинного обучения, которые необходимо обучать на изображениях конкретных культур клеток. Обучение алгоритмов является трудоемким процессом и требует длительной ручной разметки. Также доступные методы анализа на основе машинного обучения обладают низкой точностью определения живых клеток без флуоресцентной окраски. **Цель исследования:** упростить создание набора данных аннотированных клеток с последующим обучением алгоритмов на изображениях живых культур клеток. **Материалы и методы.** Методика включала использование сверточных нейронных сетей на основе алгоритма для сегментации ядер клеток на флуоресцентных и гистологических изображениях StarDist. Для создания аннотированных фазово-контрастных изображений культуры клеток образцы окрашивали ядерным флуоресцентным красителем DAPI с последующей отбраковкой некачественных изображений при помощи классификации в программе Cellprofiler Analyst. Обучение модели на основе StarDist проводили на 1130 изображениях автоматически аннотированных ядер на фазово-контрастных изображениях культуры эпителиальных клеток респираторного тракта человека, полученных на объективе 10×, размером 1600×1200 пикселей и глубиной цвета 16 бит. **Результаты исследования.** Полученная модель показала хорошую точность ( $F_1 = 0,765$ ) сегментации ядер на валидационном наборе данных. Модель применили для определения времени удвоения популяции культуры эпителиальных клеток. **Заключение:** разработанный подход позволил создать аннотации и обучить модель машинного обучения для получения данных без применения флуоресцентных меток «label-free» на живых культурах клеток.

**Ключевые слова:** машинное обучение, сверточные нейронные сети, сегментация ядер, фазово-контрастные изображения, label-free

**Конфликт интересов:** авторы заявляют об отсутствии конфликта интересов.

**Финансирование:** работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации для федерального государственного

бюджетного научного учреждения «Медико-генетический научный центр имени академика Н.П. Бочкова».

**Для цитирования:** Балясин М.В., Демченко А.Г., Люндуп А.В. Полуавтоматический метод аннотирования фазово-контрастных изображений живых культур клеток для сегментации ядер на основе машинного обучения. *Регенерация органов и тканей*. 2025;3(1): 31–40. <https://doi.org/10.60043/2949-5938-2025-1-31-40>

Получена 27.02.2025

Обработана 18.03.2025

Принята 25.03.2025

## Semi-automatic method of annotating phase-contrast images of live cell cultures for nuclei segmentation based on machine learning

Maxim V. Balyasin<sup>1,2,\*</sup>, Anna G. Demchenko<sup>2</sup>, Alexey V. Lyundup<sup>1,2</sup>

<sup>1</sup> Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University), Russia, 117198, Moscow, Miklukho-Maklaya str., 6

<sup>2</sup> Research Centre for Medical Genetics, Russia, 115522, Moscow, Moskvorechye str., 1

\* Corresponding author: balyasin\_mv@rudn.ru

### Abstract

This study developed a comprehensive approach for identifying live cell nuclei in images without fluorescent labels. Since cell biology involves counting cells and assessing cell growth dynamics and confluence, it is expedient to automate the collection of this data. Machine learning algorithms are used for automation, which must be trained on images of specific cell cultures. Training algorithms is a labor-intensive process and requires lengthy manual annotation. Also, available machine learning-based analysis methods have low accuracy in identifying living cells without fluorescent staining. **Aim of the study.** To simplify the creation of a dataset of annotated cells with subsequent training of algorithms on images of living cell cultures. **Materials and methods.** The methodology involved the use of convolutional neural networks based on an algorithm for segmenting cell nuclei in fluorescent and histological images using StarDist. To create annotated phase-contrast images of cell cultures, samples were stained with the nuclear fluorescent dye DAPI, followed by the rejection of poor-quality images using classification in the CellProfiler Analyst program. The StarDist-based model was trained on 1,130 images of automatically annotated nuclei in phase-contrast images of human respiratory tract epithelial cell cultures, obtained with a 10x lens, 1,600x1,200 pixels in size, and 16-bit color depth. **Results.** The resulting model showed good accuracy ( $F_1 = 0.765$ ) in segmenting nuclei on the validation dataset. The model was used to determine the population doubling time of the epithelial cell culture population. **Conclusion.** The developed approach made it possible to create annotations and train a machine learning model to obtain data without the use of fluorescent labels (“label-free”) on live cell cultures.

**Keywords:** machine learning, convolutional neural networks, nuclear segmentation, phase-contrast images, label-free

**Conflicts of Interest:** The authors declare no conflicts of interest.

**Funding:** This work was carried out within the framework of the state assignment of the Ministry of Science and Higher Education of the Russian Federation for the Research Centre for Medical Genetics named after Academician N.P. Bochkov.

**For citation:** Balyasin M.V., Demchenko A.G., Lyundup A.V. Semi-automatic method of annotating phase-contrast images of live cell cultures for nuclei segmentation based on machine learning. *Tissue and organ regeneration*. 2025;3(1):31–40. <https://doi.org/10.60043/2949-5938-2025-1-31-40>

Received 27.02.2025

Received 18.03.2025

Accepted 25.03.2025

#### Список сокращений:

**CNN** — сверточные нейронные сети (convolutional neural networks)

**ML** — машинное обучение (machine learning)

**PDT** — время удвоения популяции (population doubling time)

## Введение

Сегментация клеток является важной задачей в биомедицинских исследованиях. Классические методы сегментации и морфометрического анализа изображений не отличаются высокой точностью, а также обладают низкой способностью к обобщению [1]. На смену им пришли алгоритмы машинного обучения, но даже самые актуальные алгоритмы требуют разметки исходных данных для последующего обучения [2]. Аннотация объектов интереса на микроскопических изображениях требует значительных затрат времени и труда на ручную разметку данных [3]. Задача сегментации медико-биологических изображений клеток и структур ткани решается многими коллективами, некоторые сосредоточены на решении узких задач, например сегментации окрашенных ядер культур клеток и гистологических срезов, другие — на разработке генерализованных алгоритмов [4–6]. Разнообразие задач исследователя не позволяет на текущий момент использовать предобученные модели без их точной настройки [7]. Другая проблема — это необходимость применения селективных красителей для визуализации структур клетки, без которых наиболее трудно выполнить сегментацию отдельных клеток в живой культуре [8]. Фазово-контрастная микроскопия представляет собой оптимальный метод для наблюдения живых неокрашенных клеток, но автоматическая сегментация отдельных клеток и их структур на таких изображениях остается сложной задачей из-за низкого контраста и артефактов оптической системы [9, 10]. Также есть исследование по применению сегментации ядер для подсчета клеток на изображениях светлого поля [11, 12]. Однако

все это требует трудоемкого процесса подготовки набора данных для обучения моделей [13]. В данном исследовании представлен подход, который может помочь исследователям облегчить задачу подготовки набора данных для последующего обучения необходимой ML модели на примере задачи сегментации ядер культуры клеток без применения меток.

## Материалы и методы

### Подготовка образцов и набора данных

Для обучения модели использовали ФК-изображения культуры эпителиальных клеток респираторного тракта человека, полученные с помощью микроскопии в режиме реального времени на приборе Celena X (Logos biosystems, Корея) на объективе с увеличением 10×. Клетки культивировали в соответствии с протоколом, опубликованном ранее [14]. Для формирования обучающего набора данных клетки пассировали в различной концентрации начиная с 50 000 клеток/см<sup>2</sup> в 24-луночный планшет с шагом разведения 1:1,2 в двух технических повторях. Клетки культивировали в течение 12 часов, после чего фиксировали при помощи 4% раствора параформальдегида в фосфатно-солевом буфере в течение 15 минут, трижды проводили отмывку раствором DPBS («ПанЭко», Россия) и окрашивали раствором DAPI (Abcam, США) в концентрации 1 мкг/мл. Затем получали флуоресцентные изображения DAPI с соответствующим полем зрения на ФК.

### Обучение модели

Полученный набор данных с 24-луночного планшета подготовили, проведя двухэтапный полуавтоматический отбор изображений с помощью

программ CellProfiler версии 4.2.6 (Broad Institute of MIT and Harvard, США) и CellProfiler Analyst (Broad Institute of MIT and Harvard, США) версии 3.0.4 [15, 16]. На первом этапе удаляли фотографии, где ядра клеток в канале DAPI находились не в фокусе. На втором этапе на полученных изображениях проводили сегментацию ядер при помощи предобученной модели StarDist [17]. На полученных сегментированных изображениях ядер клеток с помощью CellProfiler Analyst удаляли артефакты, клетки на краю лунки и за пузырями воздуха. Таким образом всего было получено 1130 изображений и 22 636 сегментированных объектов, из которых 960 изображений были использованы для обучения модели и 170 — для валидации и проверки ее точности. Для обучения и сегментации изображений был взят алгоритм StarDist на основе CNN архитектуры U-net с последующей сегментацией объектов звездчатой формы. Архитектура U-Net была выбрана, поскольку она является специализированной и широко признанной моделью для задач семантической сегментации, демонстрирующей более высокую точность по сравнению с моделями детекции объектов (такими, как Mask R-CNN или YOLO), которые часто показывают худшие результаты в подобных задачах. Для этого использовали предварительно обученную модель на наборе данных DSB2018 для сегментации флуоресцентно окрашенных ядер клеток. Обучение длилось 450 эпох, с размером пакета 4, размером изображения для обучения 256×256 пикселей, скоростью обучения 3E-4. Время обучения составило 8–9 часов. В ходе обучения использовали аугментацию данных: отражение и поворот в случайном направлении и на случайный угол, изменение размера и яркости изображения. Для оценки хода обучения и точности модели использовали loss функцию и метрику IoU (коэффициент Жаккара). Итоговую точность модели оценивали по таким показателям, как precision, recall, accuracy, F1 score (коэффициент Дайса), mean true score, mean matched score, panoptic quality. Для обучения и применения модели использовали ноутбук с характеристиками: ЦП AMD Ryzen 7 7735HS 3.2 ГГц 8 ядер (16 логических), оперативная память SODIMM DDR5 16Gb 4800 МГц, видеокарта NVIDIA GeForce RTX4060 laptop.

Обученная модель, набор изображений (2600 изображений), исходный код для обучения и применения модели доступны в открытом облачном хранилище Яндекс-диск <https://disk.yandex.ru/d/eff1D0qDw3gACg>. Также в облачном хранилище находится подробная документация по установ-

ке необходимого программного обеспечения, примеры использования модели и инструкции по адаптации подхода для других типов изображений. Для воспроизведения результатов предоставлены файлы с параметрами обучения и предварительно обработанные данные.

### Проверка модели на фазово-контрастных изображениях культуры клеток в реальном времени

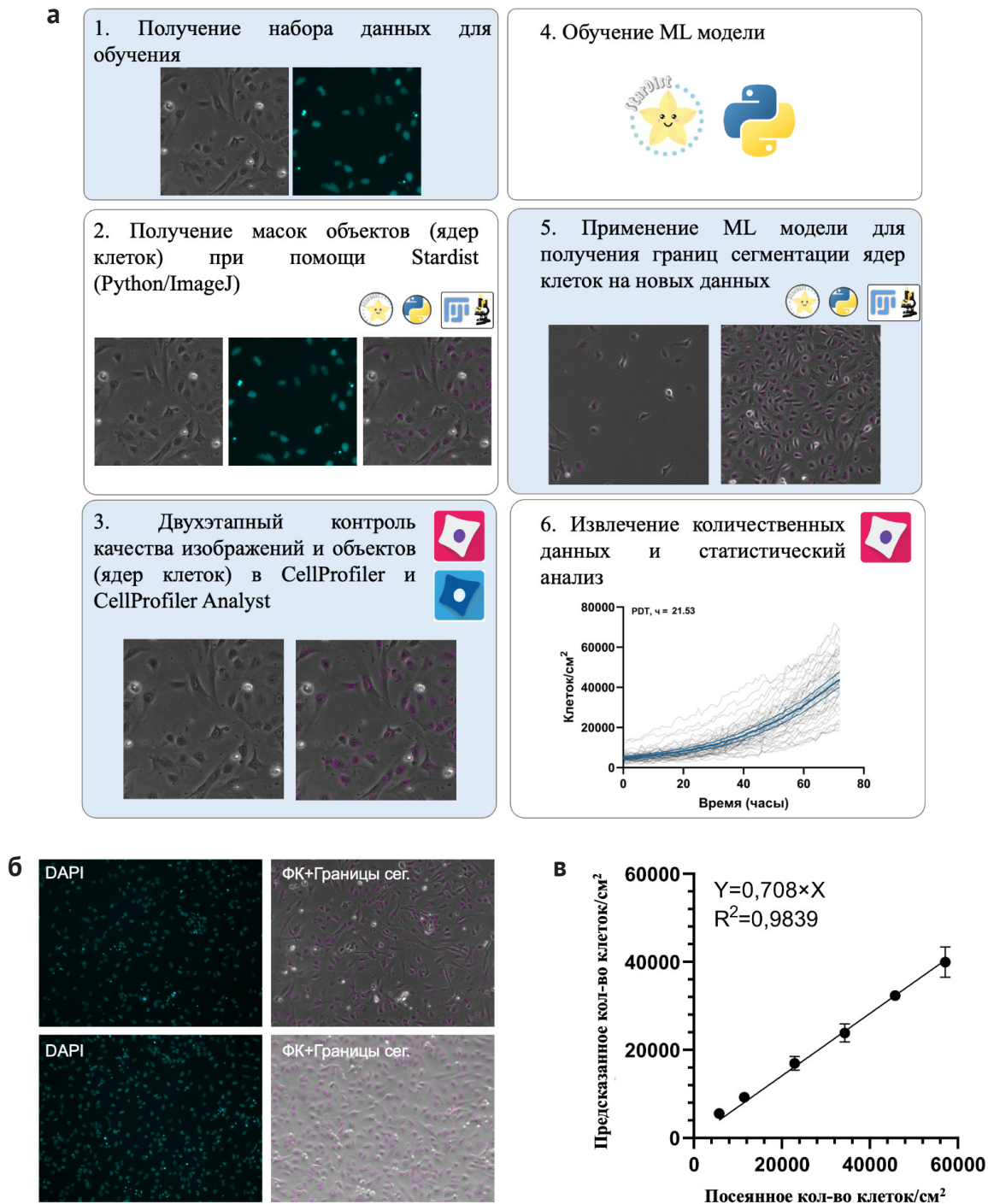
Для оценки применимости модели проводили пассирование клеток на 24-луночный планшет с концентрацией клеток  $50 \times 10^3$  клеток/см<sup>2</sup> в 3-х технических повторах. Культуры инкубировали в течение 72 часов при +37 °С, высокой влажности и атмосфере 5% CO<sub>2</sub>. Затем каждый час проводили получение изображений с 16 полей зрения на лунку планшета. На полученном наборе изображений во времени проводили сегментацию ядер клеток, оценивали плотность клеток и PDT.

### Статистический анализ

Для статистического анализа применяли программное обеспечение GraphPad Prism версии 10.1 (США). Методы статистической обработки данных включали среднее значение и 95% доверительный интервал (ДИ). Для оценки соответствия плотности клеток на изображениях обучающей выборки и количества посеянных клеток рассчитывали коэффициент корреляции Пирсона плотности клеток на изображениях и в зависимости от концентрации посаженных клеток. Также оценивали скорость роста живой культуры клеток методом оценки PDT при помощи расчета угла наклона натурального логарифма плотности клеток от времени роста. Статистическую значимость принимали при  $p$  значение  $< 0,05$ .

### Результаты

В ходе работы был разработан протокол подготовки аннотаций изображений для обучения и применения модели машинного обучения, который включает в себя этапы: получение пар флуоресцентных и ФК изображений, контроль качества изображений и объектов, обучение модели, применение модели и анализ данных (рис. 1а). В результате двухэтапного контроля качества изображений и объектов в программах CellProfiler и CellProfiler Analyst был получен набор данных с 1130 пар изображений, на которых была проведена сегментация ядер (рис. 1б). Для обучения модели допускались пары



**Рис. 1.** Набор данных для обучения ML модели с последующей сегментацией ядер на изображениях культуры эпителиальных клеток на ФК, окрашенных DAPI: а – схема протокола подготовки данных, обучения и анализа изображений; б – исходные флуоресцентные и ФК-изображения культур клеток с нанесенными границами сегментированных ядер клеток (GT); в – результаты автоматической сегментации ядер на основе изображений DAPI при помощи предварительно обученной модели CNN в соответствии с плотностью посева клеток; объектив 10×; данные представлены как среднее значение ± 95% ДИ; коэффициент корреляции Пирсона  $R^2$

**Fig. 1.** Training dataset for an ML model with subsequent nucleus segmentation in phase-contrast (PC) images of an epithelial cell culture stained with DAPI: а – schematic of the data preparation, training, and image analysis workflow; б – raw fluorescence and PC images of cell cultures with overlaid boundaries of segmented cell nuclei (ground truth, GT); в – results of automated nucleus segmentation from DAPI images using a pre-trained CNN model according to cell seeding density; 10× objective; data are presented as mean ± 95% CI; Pearson correlation coefficient ( $R^2$ )

изображений при наличии небольшого отклонения от фокуса на ФК или полученное около края лунки планшета при условии, что флуоресцентное изображение ядра клетки находится в фокусе и хорошо визуализируется (рис. 1б).

**Таблица 1.** Оценка производительности модели по основным метрикам

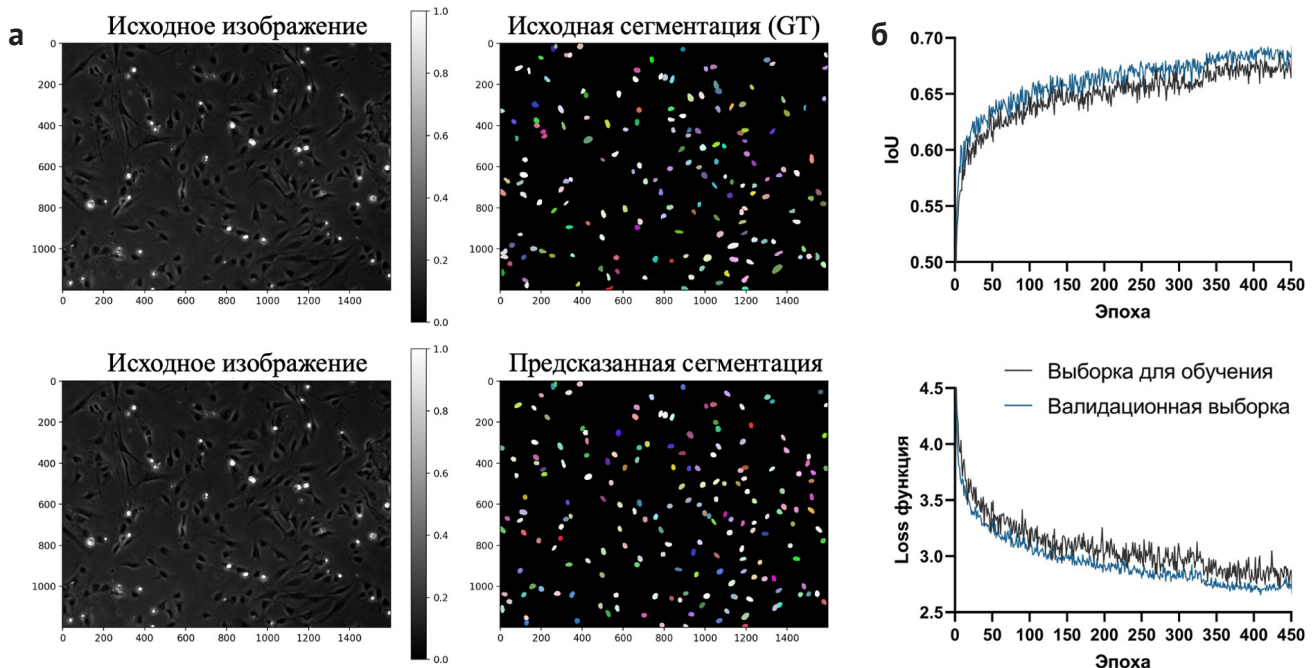
**Table 1.** Model performance evaluation using the main metrics

Параметр	Оценка
$N$ true	22 636
$N$ predicted	22 342
$N$ true positive	17 196
$N$ false positive	5146
$N$ false negative	5440
Precision	0,77
Recall	0,76
Accuracy	0,619
$F_1$ score	0,765
Mean true score	0,577
Mean matched score	0,759
Panoptic quality	0,58

Коэффициент корреляции между количеством посеянных клеток и оценкой плотности клеток на изображениях методом сегментации флуоресцентных ядер алгоритмом StarDist через 24 часа после посева составил  $R^2 = 0,9839$ .

Разные условия, в которых были получены изображения, такие как высокая и низкая плотность клеток, искажение фазового контраста на границе лунки, а также аугментация изображений, позволили добавить вариабельность для обучающей выборки, что привело к увеличению стабильности результирующей модели и возможности к обобщению на новых наборах данных. Обучение модели заняло ~8,5 часа (рис. 2б). Метрика IoU для обучающей выборки составила 0,68, для валидационной выборки — 0,689, что соответствует хорошему качеству результирующих масок (рис. 2а).

Более высокое значение IoU для валидационной выборки связано с применением аугментации изображений на обучающей выборке, которая не используется при оценке валидационной выборки. Параметры точности модели представлены в таблице 1. Так, обучение производилось на 22 636 флуоресцентно окрашенных ядрах.



**Рис. 2.** Прогресс обучения и результаты модели CNN на тестовом наборе изображений: а – границы сегментированных ядер клеток, сгенерированные при помощи CNN на тестовом наборе изображений; б – параметры обучения модели в зависимости от эпохи, IoU и loss функция для обучающей и валидационной выборки

**Fig. 2.** Training progress and CNN model performance on the test image set: a – boundaries of segmented cell nuclei generated by the CNN on the test image set; б – training parameters as a function of epoch: IoU and loss for the training and validation sets

Метрика точности precision составила 0,77, accuracy = 0,619, а  $F_1 = 0,765$ . Положительная прогностическая ценность для определения ядер составила 76,97%, чувствительность — 76,99%. Полученные показатели соответствуют приемлемой предиктивной модели.

Для оценки качества сегментации использовали следующие метрики:

$N$  true — общее количество истинных объектов (ground truth) в наборе данных;

$N$  predicted — общее количество предсказанных объектов моделью;  $N$  true positive — количество истинно-положительных предсказаний;

$N$  false positive — количество ложноположительных предсказаний;

$N$  false negative — количество ложноотрицательных предсказаний.

Precision вычисляется как доля правильно предсказанных положительных результатов из всех предсказаний модели ( $N$  true positive / ( $N$  true positive +  $N$  false positive)). Высокое значение означает низкий уровень ложноположительных результатов.

Recall вычисляется как доля правильно предсказанных положительных результатов из всех истинных положительных результатов ( $N$  true positive / ( $N$  true positive +  $N$  false negative)). Высокое значение означает низкий уровень ложноотрицательных результатов.

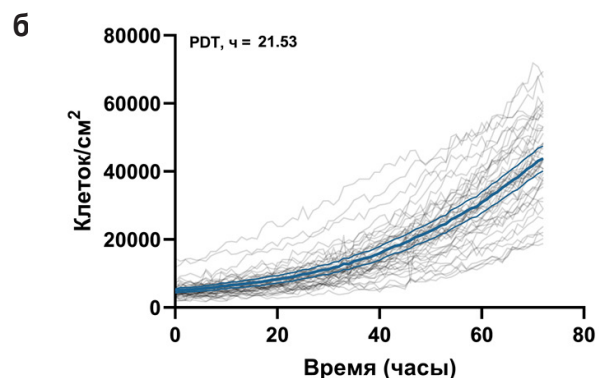
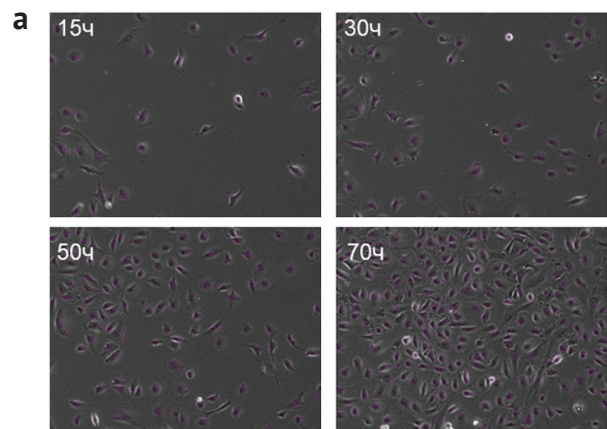
Accuracy вычисляется как общая точность предсказаний модели с учетом всех типов ошибок ( $N$  true positive / ( $N$  true positive +  $N$  false positive +  $N$  false negative)). В контексте сегментации это часто называют «средней точностью» (average precision).

$F_1$  score — вычисляется как гармоническое среднее Precision и Recall ( $2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$ ). Высокое значение указывает на хорошую производительность модели.

Mean true score — среднее значение метрики IoU (intersection over union) для всех истинных объектов, которые были сопоставлены с предсказанными объектами. Нормализуется по общему количеству истинных объектов.

Mean matched score — среднее значение метрики IoU для всех истинно-положительных сопоставлений. Это среднее значение IoU для тех объектов, которые были успешно сопоставлены.

Panoptic quality — метрика, объединяющая сегментацию и классификацию объектов. Вычисляется как  $(\text{sum\_matched\_score}) / (N \text{ true positive} + N \text{ false positive} / 2 + N \text{ false negative} / 2)$ , где  $\text{sum\_matched\_score}$  равен сумме IoU по всем  $N$  true positive. Она учитывает как качество сегментации (через IoU), так и качество обнаружения (через  $N$  true positive,  $N$  false positive,  $N$  false negative).



**Рис. 3.** Оценка на новом наборе данных скорости роста культуры эпителиальных клеток на ФК изображений без меток: а — пример роста культуры клеток через 15 часов, 30, 50 и 70 часов после начала эксперимента с нанесенными границами ядер при помощи CNN модели 1; б — PDT культуры клеток на основе результатов сегментации ядер клеток; объектив 10×, регрессионная линия — среднее значение  $\pm$  95% ДИ

**Fig. 3.** Evaluation of the growth rate of an epithelial cell culture on a new dataset of label-free PC images: а — example of cell culture growth at 15, 30, 50, and 70 h after the start of the experiment with nuclear boundaries overlaid using CNN model 1; б — population doubling time (PDT) of the cell culture based on nucleus segmentation results; 10× objective; regression line shown as mean  $\pm$  95% CI

Полученную модель применяли для оценки скорости роста живой культуры эпителиальных клеток респираторного эпителия в режиме реального времени. Модель эффективно сегментировала ядра клеток на всех стадиях роста культуры клеток начиная с низкой плотности на 0 час, заканчивая 70 часами роста (рис. 3а). В результате оценки количества ядер во времени на живой культуре клеток PDT составила 21,53 часа (рис. 3б), а конечная плотность — 43 786 клеток/см<sup>2</sup> (40 130–47 441, 95% ДИ).

Таким образом, разработан подход для получения аннотированных изображений, который показывает высокую точность в сегментации ядер клеток на ФК изображениях и который можно применять для решения задач количественного анализа живых культур клеток без применения флуоресцентных меток.

### Заключение

Разработанный полуавтоматический метод аннотирования для обучения модели машинного обучения сегментации ядер на ФК-изображениях показал высокую эффективность и точность. Несмотря на то что метрика точности IoU на ФК-изображениях в сравнении с флуоресцентно окрашенными ядрами обладает более низкой точностью, сегментация ядер в нашей работе показала точность предсказанных масок выше, чем при сегментации всей клетки на фазовом контрасте [17]. Метод определения ядер клеток по ФК-изображениям

является перспективным для введения неинвазивных контролей качества в режиме реального времени в области клеточных технологий, для скрининга и тестирования веществ, а также в исследовательских целях.

Одним из ограничений для исследователей в применении машинного обучения на текущий момент может стать необходимость использования кода и его написание. В предложенном подходе выбраны такие алгоритмы, которые имеют наиболее полную документацию по установке, применению, а также имеют примеры, которые можно повторить для воспроизведения методик. Другим ограничением текущей работы является то, что все изображения для анализа получены на одном микроскопе. В дальнейшем мы планируем провести тестирование на изображениях с других микроскопов, а также расширить алгоритм сегментации, добавив возможность отслеживания траектории отдельной клетки или ядра [18–20].

Стоит отметить, что, несмотря на наличие в открытом доступе большого количества проектов по сегментации микроскопических изображений, нет хорошо генерализованных моделей, поэтому исследователям необходимо до обучать доступные модели на своих данных. Предложенный нами метод позволяет облегчить исследователям применение методов машинного обучения и компьютерного зрения, а также адаптировать их под свои задачи путем валидации модели на изображениях.

### Литература

1. Ayanzadeh A., Yağar H.O., Özuysal Ö.Y., Okvur D.P., Töreyn B.U., Ünay D., Önal S. Cell segmentation of 2D phase-contrast microscopy images with deep learning method // 2019 Medical Technologies Congress (TIPTEKNO). IEEE, 2019. P. 1–4.
2. Kirillov A., Mintun E., Ravi N., Mao H., Rolland C., Gustafson L., et al. Segment anything // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023. P. 4015–4026.
3. Kakumani A.K., Padma Sree L. A Deep Learning Approach for Segmenting Time-Lapse Phase Contrast Images of NIH 3T3 Fibroblast Cells // New Trends in Computational Vision and Bio-inspired Computing: Selected works presented at the ICCVBIC 2018, Coimbatore, India. Cham: Springer International Publishing, 2020. P. 855–862.
4. Shamshad F., Khan S., Zamir S.W., Khan M.H., Hayat M., Khan F.S., Fu H. Transformers in medical imaging: A survey // Medical Image Analysis. 2023. Vol. 88. P. 102802.
5. Greenwald N.F., Miller G., Moen E., Kong A., Kagel A., Dougherty T., et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning // Nature Biotechnology. 2022. Vol. 40, № 4. P. 555–565.

6. Stringer C., Wang T., Michaelos M., Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation // *Nature Methods*. 2021. Vol. 18, № 1. P. 100–106.
7. Lee M.Y., Bedia J.S., Bhate S.S., Barlow G.L., Phillips D., Fantl W.J., et al. CellSeg: a robust, pre-trained nucleus segmentation and pixel quantification software for highly multiplexed fluorescence images // *BMC Bioinformatics*. 2022. Vol. 23, № 1. P. 46.
8. Holme B., Bjørnerud B., Pedersen N.M., de la Ballina L.R., Wesche J., Haugsten E.M. Automated tracking of cell migration in phase contrast images with CellTraxx // *Scientific Reports*. 2023. Vol. 13, № 1. P. 22982.
9. Hörst F., Rempe M., Heine L., Seibold C., Keyl J., Baldini G., et al. Cellvit: Vision transformers for precise cell segmentation and classification // *Medical Image Analysis*. 2024. Vol. 94. P. 103143.
10. Berg S., Kutra D., Kroeger T., Straehle C.N., Kausler B.X., Haubold C., et al. Ilastik: interactive machine learning for (bio) image analysis // *Nature Methods*. 2019. Vol. 16, № 12. P. 1226–1232.
11. Ali R., Gooding M., Szilágyi T., Vojnovic B., Christlieb M., Brady M. Automatic segmentation of adherent biological cell boundaries and nuclei from brightfield microscopy images // *Machine Vision and Applications*. 2012. Vol. 23, № 4. P. 607–621.
12. Cross-Zamirski J.O., Mouchet E., Williams G., Schönlieb C.B., Turkki R., Wang Y. Label-free prediction of cell painting from brightfield images // *Scientific Reports*. 2022. Vol. 12, № 1. P. 10001.
13. Pachitariu M., Stringer C. Cellpose 2.0: how to train your own model // *Nature Methods*. 2022. Vol. 19, № 12. P. 1634–1641.
14. Demchenko A., Belova L., Balyasin M., Kochergin-Nikitsky K., Kondrateva E., Voronina E., et al. Airway basal cells from human-induced pluripotent stem cells: a new frontier in cystic fibrosis research // *Frontiers in Cell and Developmental Biology*. 2024. Vol. 12. P. 1336392.
15. Stirling D.R., Swain-Bowden M.J., Lucas A.M., Carpenter A.E., Cimini B.A., Goodman A. CellProfiler 4: improvements in speed, utility and usability // *BMC Bioinformatics*. 2021. Vol. 22, № 1. P. 433.
16. Stirling D.R., Carpenter A.E., Cimini B.A. CellProfiler Analyst 3.0: accessible data exploration and machine learning for image analysis // *Bioinformatics*. 2021. Vol. 37, № 21. P. 3992–3994.
17. Schmidt U., Weigert M., Broaddus C., Myers G. Cell detection with star-convex polygons // *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer International Publishing, 2018. P. 265–273.
18. Tsai H.F., Gajda J., Sloan T.F., Rares A., Shen A.Q. Usiigaci: Instance-aware cell tracking in stain-free phase contrast microscopy enabled by machine learning // *SoftwareX*. 2019. Vol. 9. P. 230–237.
19. Fazeli E., Roy N.H., Follain G., Laine R.F., von Chamier L., Hänninen P.E., et al. Automated cell tracking using StarDist and TrackMate // *F1000Research*. 2020. Vol. 9. P. 1279.
20. Moen E., Borba E., Miller G., Schwartz M., Bannon D., Koe N., et al. Accurate cell tracking and lineage construction in live-cell imaging experiments with deep learning // *bioRxiv*. 2019. 803205.

### Об авторах

**Балысин Максим Витальевич** — младший научный сотрудник Научно-образовательного ресурсного центра «Клеточные технологии» ФГАОУ ВО «Российский университет дружбы народов имени Патриса Лумумбы»; научный сотрудник Лаборатории мутагенеза ФГБНУ «Медико-генетический научный центр имени академика Н.П. Бочкова».

**Демченко Анна Григорьевна** — к.б.н., старший научный сотрудник Лаборатории редактирования генома ФГБНУ «Медико-генетический научный центр имени академика Н.П. Бочкова».

**Льондуп Алексей Валерьевич** — к.м.н., директор Научно-образовательного ресурсного центра «Клеточные технологии» ФГАОУ ВО «Российский университет дружбы народов имени Патриса Лумумбы»; ведущий научный сотрудник Лаборатории

Балясин М.В., Демченко А.Г., Люндуп А.В.

Полуавтоматический метод аннотирования изображений культур

мутация ФГБНУ «Медико-генетический научный центр имени академика Н.П. Бочкова». SPIN: 4954-3004

### Authors

**Maxim V. Balyasin** — Junior Researcher, Scientific and Educational Resource Center “Cell Technologies”, RUDN University; Researcher, Laboratory of Mutagenesis, Research Centre for Medical Genetics.

**Anna G. Demchenko** — Cand. Sci. (Biology), Senior Researcher, Laboratory of Genome Editing, Research Centre for Medical Genetics.

**Alexey V. Lyundup** — Cand. Sci. (Medicine), Director, Scientific and Educational Resource Center “Cell Technologies”, RUDN University; Leading Researcher, Laboratory of Mutagenesis, Research Centre for Medical Genetics.

---

### Вклад авторов

**М.В. Балясин** — концепция или дизайн работы, сбор данных, анализ и интерпретация данных.

**А.Г. Демченко** — сбор данных, составление статьи, итоговая переработка статьи.

**А.В. Люндуп** — окончательное утверждение версии для публикации.

### Author contribution statement

**Maxim V. Balyasin** — conception or design of the work; data acquisition; data analysis and interpretation.

**Anna G. Demchenko** — data acquisition; drafting the manuscript; critical revision of the manuscript.

**Alexey V. Lyundup** — final approval of the version to be published.