

Методы трансферного обучения для сегментации клеток глиомы С6

Д.А. Илюхин¹, В.О. Ячная^{1,2}, Р.О. Малашин^{1,2}, М.К. Ермаченкова¹, А.В. Волков^{3,4},
С.Г. Пашкевич⁴, А.А. Денисов^{3,4}

¹ Институт физиологии имени И.П. Павлова РАН, 199034, Россия, г. Санкт-Петербург, наб. Макарова, д.6;

² Государственный университет аэрокосмического приборостроения,
190000, Россия, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. а;

³ Белорусский государственный университет, 220030, Беларусь, г. Минск, пр. Независимости, д. 4;

⁴ Институт физиологии НАН Беларуси, 220072, Беларусь, г. Минск, ул. Академическая, д. 28

Аннотация

В настоящей работе представлен алгоритм бинарной сегментации клеток культуры глиомы С6 с применением методов глубокого обучения для упрощения и ускорения анализа роста данной культуры. Для разработки и тестирования алгоритма впервые собрана база, содержащая 30 микроскопических фазово-контрастных снимков клеток глиомы С6. Благодаря собранной базе данных, впервые исследуются свойства методов трансферного обучения при работе со снимками клеточных культур в высоком разрешении. Для этого используются дополнительные базы данных микроскопических и неспециализированных изображений. Эксперименты показали, что предобучение нейронной сети с использованием микроскопических снимков обеспечивает значительное преимущество при распознавании клеток глиомы в низком разрешении, однако при увеличении разрешения и возникающей из-за этого рассогласованности размеров объектов лучшие результаты могут быть получены с использованием энкодеров, обученных на повседневных снимках. В работе объясняются причины этого и предлагаются способы повышения обобщающих способностей «специализированных» весов к работе в высоком разрешении за счет использования техник аугментации, обеспечивающих согласованность размеров объектов. Также демонстрируется возможность получать разные обобщающие свойства итогового алгоритма за счет правильного подбора данных для предварительного обучения, что может быть полезно при необходимости обнаруживать или исключать из рассмотрения плотные кластеры клеток.

Ключевые слова: бинарная сегментация, микроскопические снимки, компьютерное зрение, U-Net, нейронные сети, глиома С6.

Цитирование: Илюхин, Д.А. Методы трансферного обучения для сегментации клеток глиомы С6 / Д.А. Илюхин, В.О. Ячная, Р.О. Малашин, М.К. Ермаченкова, А.В. Волков, С.Г. Пашкевич, А.А. Денисов // Компьютерная оптика. – 2025. – Т. 49, № 5. – С. 794-804. – DOI: 10.18287/2412-6179-CO-1609.

Citation: Ilyukhin DA, Yachnaya VO, Malashin RO, Ermachenkova MK, Volkov AV, Pashkevich SG, Denisov AA. Transfer learning methods for glioma C6 cell segmentation. Computer Optics 2025; 49(5): 794-804. DOI: 10.18287/2412-6179-CO-1609.

Введение

Данная работа посвящена разработке модели искусственной нейронной сети для бинарной сегментации клеток глиомы С6. Для разработки и тестирования алгоритма нами впервые собрана и размечена база данных (БД) микроскопических фазово-контрастных снимков клеток глиомы С6. Глиома С6 – линия глиальных клеток, выделенная из мозга крысы с глиомой. Глиомы – опухоли, развивающиеся из глиальных клеток, они являются наиболее распространенными опухолями центральной нервной системы (например, мультиформная глиобластома человека). В связи со схожестью глиомы С6 и глиобластомы, первую используют для изучения роста опухоли и нахождения эффективного метода лечения [1].

Создание системы анализа изображений биологических нейронных сетей может существенно повлиять на выявление нарушений процессов развития структур

нервной системы в онтогенезе, а также стать инструментом новых подходов к диагностике и прогнозу развития патологий центральной нервной системы.

Одним из неинвазивных подходов является использование световой микроскопии для анализа формы и поведения нейронных клеток. Такой метод микроскопической визуализации, как фазовый контраст, может быть использован для захвата внешнего вида клеток без их фиксации и окрашивания. В этом случае взаимодействие на уровне клеток может быть проанализировано в режиме реального времени с помощью методов компьютерного зрения. Используемые в данной работе методы сбора данных позволяют не вмешиваться в процессы роста и развития культур клеток и экономить ресурсы на диагностические маркеры. Как будет показано, собираемая БД снимков клеток глиомы С6 не имеет аналогов, поэтому при разработке модели искусственной нейронной сети нами исследуются методы

трансферного обучения с использованием данных, имеющих сходные визуальные параметры.

В данной статье, во-первых, описывается процесс формирования БД фазово-контрастных микроскопических снимков клеток глиомы С6. Во-вторых, разрабатывается бинарный сегментатор на основе искусственной нейронной сети. Мы впервые исследуем свойства методов трансферного обучения при работе со снимками клеточных культур в высоком разрешении. В частности, выявляем значительную зависимость эффективности переноса признаков от согласованности размеров объектов во время обучения и дообучения, а также предлагаем и экспериментально подтверждаем эффективность метода обеспечения необходимой согласованности при обучении.

1. Смежные работы

Несмотря на широкую популярность исследований, связанных с анализом медицинских снимков, в частности с сегментацией микроскопических снимков клеток, до сих пор не проводилось разработок систем анализа фазово-контрастных микроскопических снимков клеток глиомы С6.

В связи с этим в данном параграфе рассмотрим исследования, посвящённые сегментации клеток на микроскопических снимках, имеющих схожие визуальные параметры. Наиболее распространённые БД микроскопических снимков обычно содержат в себе флуоресцентные снимки (например, Data Science Bowl 2018 [3] для сегментации ядер клеток, CellPose [4] и её расширение OmniPose [5] для сегментации клеток). Как уже отмечалось, фиксация клеток и получение флуоресцентных снимков – это более ресурсоёмкая задача, которая к тому же не позволяет проводить исследования, связанные с динамикой жизни клеток (например, отслеживание их роста или перемещения).

Для данного исследования для этапа предварительного обучения системы нами была выбрана представленная в 2021 году БД LIVECell [6]. Эта база содержит более 5000 фазово-контрастных микроскопических снимков 8 типов клеток и предназначена для сегментации по образцам (всего более 1,6 млн индивидуальных клеток). Представленные в базе типы клеток охватывают различные виды морфологий: от небольших круглых (BV-2) до крупных плоских клеток (SK-OV-3), а также, что наиболее ценно для текущего исследования, нейроноподобные (SH-SY5Y). Значительное количество снимков в базе содержит примеры контактов и перекрытий нескольких клеток, что усложняет процесс сегментации (особенно в случае сегментации по образцам), однако такие ситуации более репрезентативны для реальных задач в биомедицинских исследованиях.

БД LIVECell широко применяется в различных исследованиях и разработках [7 – 18]. Например, она эффективно используется в качестве базы для предварительного обучения системы в работах [7 – 9]. В [7] описывается процесс обучения модели CellPose [4],

основанной на архитектуре U-Net [2] с остаточными блоками. В [8] сначала проводится предобучение модели Mask R-CNN, затем адаптируется алгоритм [19] повышения контраста изображения (для более точного выделения слабоконтрастных границ клеток) и далее разрабатывается модуль поточечного внимания (PAPF – Point-wise attention mechanism) для сети с агрегацией путей (PANet – Path Aggregation Network). В [9] авторы комбинируют архитектуры Mask R-CNN и Cascade R-CNN. В данных работах в целом отмечается зависимость точности сегментатора от количества обучающих данных, в частности в [9] для обучения используются и изображения, предусмотренные авторами БД как валидационные.

Кроме того, LIVECell используется и для тестирования итоговых моделей [10 – 17]. Например, в [10] используется архитектура EfficientDet [20]; в [11] разрабатывается архитектура Single-Cell Transformer Segmenter (SCTS) на основе Swin Transformer [21]; в [13] – архитектура Dense Cell Segmentation Network (DCSN); в [14] – архитектура End-to-end Cell Segmentation (ECS-Net), которая детектирует, сегментирует и подсчитывает количество объектов. В серии работ [15 – 17] используется БД EVICAN [22] и позже БД LIVECell. В [15] (система DeepCeNS) и [17] (система DeepMuCS) представляется следующая архитектура сегментатора: извлечение признаков с помощью Feature Pyramid Network (FPN) [23] и ResNeSt-200 [24], определение региона с помощью Regional Proposal Network (RPN), определение класса, ограничивающей рамки и бинарной маски для каждого региона с помощью Cascade Mask R-CNN [25]. Авторами также ставится задача сегментации ко-культуры, когда на изображении представлены клетки нескольких культур [16]. В работе [12] разрабатывается конвейер трансферного обучения сегментаторов (а именно, предназначенных для сегментации по образцам изображений клеток) для тех случаев, когда доступно большое количество данных, мало из которых при этом аннотированы.

Выбранная нами для текущей работы архитектура U-Net [2] хорошо себя зарекомендовала в задачах сегментации медицинских снимков (например, в [4, 26 – 33]). Кроме того, следует отметить, что стандартная архитектура U-Net имеет множество эффективно применяемых модификаций, среди которых использование в качестве энкодера альтернативных архитектур: ResNet [34] (например, в [31, 35]), VGG [36] (например, в [31, 35, 37]) и Трансформеров (например, в [32, 33]).

2. Данные и методы

Данные. Поскольку данная работа посвящена сегментации клеток глиомы С6 на фазово-контрастных снимках – задаче, ранее не решавшейся, то для предварительного обучения разрабатываемой системы нами была выбрана БД LIVECell, имеющая сходные визуальные параметры. Более подробно эта БД была описана в предыдущем параграфе.

Сбор собственных данных проводился в Центре мозга государственного научного учреждения «Институт физиологии Национальной академии наук Беларуси». Культивирование клеток и последующая съёмка материала состояли из следующих этапов:

1. Перевиваемые глиальные клетки C6 (ATCC CCL-107) крысы *Rattus norvegicus*, ATCC CCL-107 культивировали (концентрация 2×10^5 клеток/мл) на флаконах в среде F-12K (модификация Кайна среды Хэма F-12 ATCC 30-2004) с добавлением 10%-й эмбриональной телячьей сыворотки (Cargilcom, NCS-1A) и 10^{-4} г/мл раствора сульфата гентамицина. Флаконы размещали в CO₂-инкубаторе при 5 % CO₂ и температуре 37°C.
2. Изображения получали с помощью микроскопа BestScope BS-2092 в фазово-контрастном режиме с установленными объективами увеличением 10× и 20× через 24 и 48 часов после высевания суспензии клеток. Через 24 часа клетки глиомы закрепляются к поверхности чашки Петри, приобретают специфическую форму и приступают к активному делению.
3. Цифровая съёмка культур производилась цифровой камерой TourCam UCMOS05100KPA.

Разметка данных производилась вручную в системе CVAT (Computer Vision Annotation Tool) [38]. Маски сохранялись в формате аннотации COCO [39], на основе которой формировались бинарные изображения. Некоторые иные подробности процесса ручной разметки рассказываются в [40]. Всего собрано и размечено 30 изображений размером 2592×1952. Примеры изображений из LIVECell и из собранной базы представлены на рис. 1. Собранный БД далее в работе обозначается как “Glioma C6”.

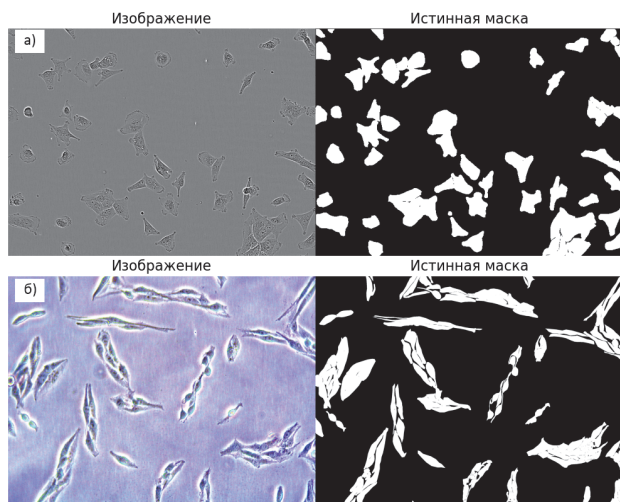


Рис. 1. Примеры изображений и их масок из БД: (а) LIVECell (класс A172 после 8 часов культивирования), (б) Glioma C6. Контраст увеличен

В Приложении А приведено сравнение характеристик БД Glioma C6 и рассмотренных смежных БД. В частности, отмечаются высокое разрешение собранных данных по сравнению с аналогами и различия в цветовом режиме изображений.

Архитектура. Для данного проекта была выбрана модифицированная архитектура U-Net [2, 37], позволяющая использовать различные предварительно обученные на ImageNet модели-классификаторы в качестве энкодера, некоторые из них были рассмотрены ранее.

Декодерная часть модифицированной архитектуры U-Net универсальна: для увеличения изображений используется билинейная интерполяция, а в качестве активационной функции применяется ReLU.

На рис. 2 приведена общая структура модифицированной архитектуры U-Net [37], использованной в настоящей работе.

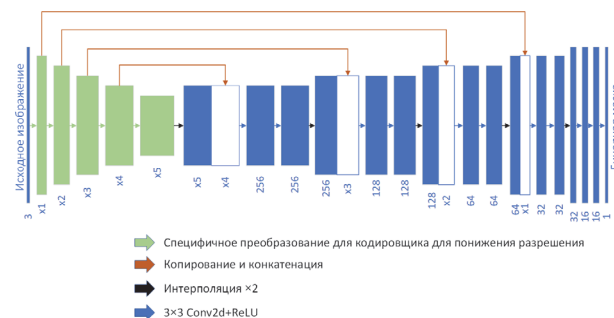


Рис. 2. Общая структура использованной архитектуры, каждый зелёный блок представляет собой многоканальную карту признаков, полученную путем выполнения комбинированных преобразований, которые зависят от выбранной архитектуры энкодера. Каждый синий блок представляет собой многоканальную карту признаков, полученную путем выполнения предыдущих преобразований. Высота блока показывает относительный пространственный размер карты, в то время как его ширина пропорциональна количеству каналов (число явно указано под соответствующим блоком, если данное значение не зависит от архитектуры энкодера, в ином случае в виде обозначения используется значение “x”)

Эксперименты были направлены на создание алгоритмов сегментации, обеспечивающих максимальную точность сегментации набора данных Glioma C6, но при этом рассматривалась возможность работы в более низком разрешении, что может быть полезно для работы во встроенных системах. Для обучения использовались видеокарты NVIDIA GeForce GTX 3090 и NVIDIA A6000 Ada.

Параметры обучения. Для обучения моделей использовался оптимизатор Adam с фиксированным шагом обучения 0,0001. Выбор такого шага обучения был определен по результатам предварительных экспериментов, где тестировались различные стратегии автоматического изменения шага обучения при выходе на плато, и фиксированное значение показало лучшие результаты, что может быть обусловлено способностью Adam к адаптивному изменению скоростей обучения для каждого параметра.

Нейронная сеть всегда получает на вход трехканальное изображение. Для базы данных LIVECell данные преобразуются путем дублирования значений яркости полутонового изображения по всем трем каналам. Кроме

того, было выявлено, что увеличение контраста изображений значительно улучшает обучение моделей. Поэтому во всех экспериментах перед подачей в модель предварительно повышался контраст изображений.

Для оценки качества сегментации использовалась стандартная для этой задачи метрика DiceScore. Обучение всех моделей на БД LIVECell производилось в течение 60 эпох, а на БД Glioma C6 — 300 эпох. Количество эпох было выбрано на основе наблюдений за изменением функции потерь: оно позволяет достичь стабильной сходимости без признаков недообучения. Для контроля переобучения используется метод ранней остановки через мониторинг функции потерь на валидационной выборке.

Размер пакета данных для всех экспериментов выбирался максимально возможным в зависимости от объема памяти видеокарты.

Разбиение данных на тренировочную, валидационную и тестовую выборки для каждой БД определялось следующим образом:

1. Для LIVECell использовались разбиения данных, предложенные авторами базы данных [41–43].
2. Для Glioma C6 выборки формировались случайным образом в соотношении 15, 5 и 10 соответственно.

При обучении моделей на БД LIVECell всегда использовались веса энкодеров, предварительно обученных на БД ImageNet, поскольку, как было показано в [37], это позволяет значительно повысить точность сегментации без использования предварительного обучения на схожей базе данных. В экспериментах с дообучением на БД Glioma C6 исследовалось влияние различных способов инициализации (включая случайно инициализированные, веса энкодера ImageNet и веса, полученные после дообучения на LIVECell).

Сравнение результатов и метрики. В ходе экспериментов было выявлено, что различие в разрешениях изображений для обучения и тестирования значительно влияет на оцениваемые метрики, вследствие чего в данной работе используются модифицированные метрики DiceScore. Стандартный DiceScore рассчитывается следующим образом (1):

$$\text{DiceScore}(A, B) = 2 \times \frac{|A \cap B|}{|A| + |B|}, \quad (1)$$

где A — это предсказанная маска сегментации, а B — это истинная маска сегментации.

Поскольку для корректного вычисления метрики DiceScore маски A и B должны иметь одинаковое разрешение, а обученные модели чувствительны к размеру входного изображения, была введена метрика DS^c (2), которая использует функцию билинейной интерполяции ς для приведения масок к размеру R_2 перед вычислением метрики:

$$\begin{aligned} DS^c(x, y, R_1, R_2) &= \text{DiceScore}(\varsigma(z, R_2), \varsigma(y, R_2)), \\ z &= f(\varsigma(x, R_1); \Theta), \end{aligned} \quad (2)$$

где x — исходное изображение, y — истинная маска сегментации для изображения x , R_1 — разрешение изображения для подачи в нейронную сеть, z — выход функции f , реализующей прямой проход модели с весами Θ . Разрешения R_1 и R_2 представляют из себя кортежи («Новая Ширина», «Новая Высота»), которые используются для масштабирования передаваемого изображения. Согласно формуле (2), в некоторых экспериментах требовалось применение билинейной интерполяции к эталонной маске. В этом случае для того, чтобы гарантировать, что объект на эталонной маске будет выделен целиком до 1, округляются все элементы со значением больше 0.

Для повышения достоверности результатов при дообучении моделей на БД Glioma C6 использовалась кросс-валидация на основе повторного случайного разбиения (repeated random sub-sampling). Данные были случайным образом разделены на тренировочную (15 изображений), валидационную (5 изображений) и тестовую (10 изображений) выборки. Этот процесс повторялся 10 раз с использованием различных начальных значений генератора случайных чисел. После этого для каждой из данных выборок обучалась отдельная модель, при этом все модели имели одинаковую архитектуру и параметры.

Усредненное значение метрики DS^c при использовании кросс-валидации определяет метрику DiceScore^{AV} (3):

$$\text{DiceScore}^{AV}(R_1, R_2) = \frac{1}{KN} \sum_{j=0}^{K,N} DS^c(x_i, y_i, R_1, R_2), \quad (3)$$

где N — количество элементов в тестовой выборке Glioma C6, $K = 10$ — количество повторений для кросс-валидации, x_i и y_i — i -й пример выборки.

В отдельных случаях используется метрика потерь DiceLoss^{AV} (4), выражаемая через DiceScore^{AV} :

$$\text{DiceLoss}^{AV}(R_1, R_2) = 1 - \text{DiceScore}^{AV}(R_1, R_2). \quad (4)$$

В данной работе исследуется функционирование моделей с тремя типами разрешений R_i :

1. Исходное разрешение БД LIVECell — 704×512 пикселей.
2. Максимально доступное разрешение для обучения U-Net с энкодером MiT-B3 — 1472×1120 пикселей.
3. Исходное разрешение БД Glioma C6 — 2592×1952 пикселей.

Для удобства далее они будут обозначаться как 704p, 1472p и 2592p соответственно.

3. Результаты экспериментов

3.1. Выбор архитектуры энкодера и оценка размера обучающей выборки

Для выбора нескольких энкодеров для дальнейших экспериментов была оценена точность распознавания LIVECell, также были исследованы способности сегментаторов обобщаться на целевой БД Glioma C6 в

режиме zero-shot (без обучающих примеров). Были проанализированы следующие энкодеры: DenseNet169 [44], DPN92 [45], EfficientNet-b5 [46], InceptionV4 [47], MiT-B3 [48], MobileNet v2 [49], ResNet50 [34], ResNeXt50 32x4d [50], SE-ResNet50 [51], SE-ResNeXt50 32x4d [51], GERNetL [52], RegNetx 064 [53], Res2Net50 26w 4s [54], ResNeSt50d [24], VGG19 [36], xception [55].

На основе полученных результатов предпочтение было отдано архитектурам с наивысшими показателями метрики DiceScore: VGG19 и MiT-B3, последняя из которых единственная показала некоторую работоспособность в режиме zero-shot.

Далее при обозначении модели U-Net через дефис будет указываться архитектура используемого энкодера.

3.2. Трансферное обучение для сегментации изображений клеток в высоком разрешении

На рис. 3 приведен график зависимости ошибки $DiceLoss^{AV}(704p, 704p)$ от размера обучающей выборки. Стоит отметить, что разница в результатах при использовании 10 и 15 изображений менее значительна по сравнению с использованием только 5 изображений. Таким образом, выбранный размер выборки можно считать оптимальным для проведения исследований. Для улучшения обобщающих способностей системы ниже более подробно рассматривается подход на основе трансферного обучения.

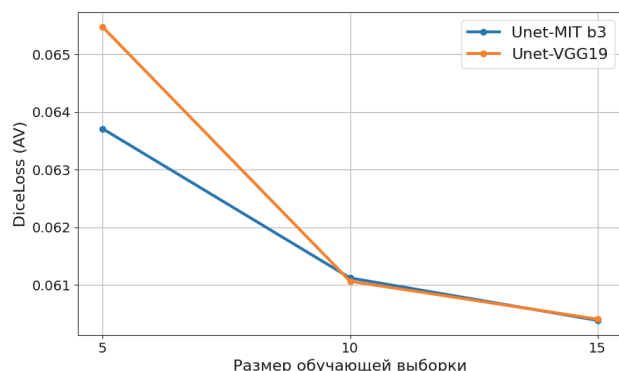


Рис. 3. График зависимости ошибки $DiceLoss^{AV}(704p, 704p)$ от размера тренировочной выборки

3.3. Трансферное обучение

Ввиду ограниченного размера исходного набора данных основной целью исследования было повышение обобщающей способности моделей с использованием предварительной подстройки весов на наборе данных LIVECell.

В табл. 1 приведены результаты вычисления метрики $DiceScore^{AV}$ для всех использованных весов моделей.

На рис. 4 приведен график зависимости значения ошибки $DiceLoss^{AV}(704p, 704p)$ от эпохи обучения для разных типов весовых коэффициентов.

Как можно наблюдать по рис. 4, использование весовых коэффициентов после предобучения моделей на

наборе данных LIVECell позволяет значительно ускорить дообучение моделей на Glioma C6. Этот результат отличается от результата, приведенного в [37], где использование предварительного обучения позволяет лишь очень незначительно повысить точность по сравнению с использованием энкодеров-классификаторов, обученных на ImageNet. Это можно связать с большой визуальной схожестью LIVECell и Glioma C6. Кроме того, в данном случае контролируется согласованность размеров объектов, что позволяет с большей эффективностью использовать веса, полученные при обучении решению родственных задач.

Табл. 1. Результаты вычисления метрики $DiceScore^{AV}(704p, 704p)$ для разных инициализирующих весов после дообучения на Glioma C6 в разрешении 704p

Архитектура	Случайные	ImageNet (энкодер)	LIVECell (вся модель)
UNet-MiT-B3	0,868	0,930	0,940
UNet-VGG19	0,913	0,932	0,940

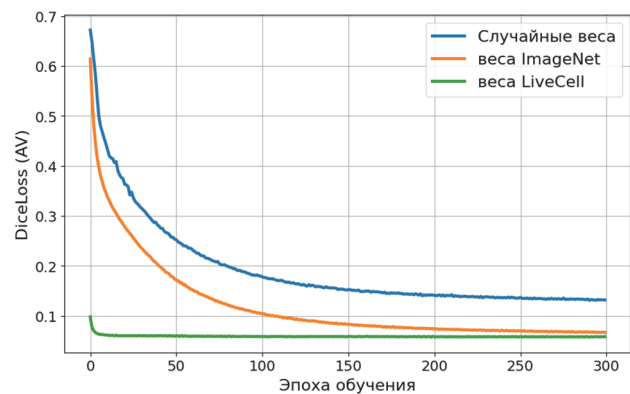


Рис. 4. График зависимости значения ошибки $DiceLoss^{AV}(704p, 704p)$ от эпохи обучения для разных типов инициализации весовых коэффициентов

Также можно заметить, что начальная ошибка при инициализации весов на основе LIVECell значительно ниже, чем при использовании других методов инициализации. Это свидетельствует о том, что данный набор данных позволяет формировать признаки, хорошо подходящие для переноса на исследуемую задачу, однако в первую очередь наблюдаемый эффект, видимо, связан с тем, что после дообучения на LIVECell в сети инициализируются все веса, а не только веса энкодера. При этом можно отметить большую зависимость архитектуры на основе UNet-MiT-B3 от предварительного обучения энкодера (см. табл. 1) по сравнению с UNet-VGG19.

При тестировании моделей на полном разрешении Glioma C6 их эффективность ниже, как это видно из табл. 2.

Также можно отметить, что во всех случаях подача на вход модели разрешения, на котором обучение сети не проводилось, значительно ухудшает результаты. Это свидетельствует об ограниченных способностях сети к обобщению на разные масштабы образов.

Мы объясняем это тем, что предварительное обучение на микроскопических снимках приводит к формированию признаков, пригодных для эффективного распознавания объектов лишь определенного масштаба.

Табл. 2. Результаты тестирования моделей на разрешении 2592p после их дообучения на Glioma C6 на разрешении 704p. В названиях архитектур в скобках указаны базы данных, использованные для предварительного обучения

Архитектура	DiceScore ^{AV} (704p, 2592p)	DiceScore ^{AV} (2592p, 2592p)
UNet-MiT-B3 (ImageNet)	0,876	0,696
UNet-VGG19 (ImageNet)	0,878	0,742
UNet-MiT-B3 (LIVECell)	0,886	0,817
UNet-VGG19 (LIVECell)	0,887	0,632

Далее мы исходим из того, что отношение размера объектов к размеру изображения является одинаковым для БД LIVECell и Glioma C6 (сходство можно оценить, например, по маскам объектов на рис. 1). Сопоставление размеров объектов напрямую не представляется возможным, поскольку собранная база данных предназначена для бинарной сегментации, а не для сегментации по экземплярам.

В связи с этим далее приводится оценка метрики DiceScore^{AV}(x, 2592p), где x – разрешение, которое использовалось для обучения на Glioma C6. Из табл. 3 видно, что если для подстройки на Glioma C6 использовать изображения с более высоким разрешением, то

лучшим вариантом становится обучение, начинающееся с весов ImageNet, а не LIVECell.

В табл. 3 *crop* обозначает процедуру извлечения случайного фрагмента заданного размера, а *resize* – процедуру масштабирования всего изображения к заданному размеру. Заметим, что при использовании *crop* масштаб объектов не меняется, поэтому для такого дообучения используется метрика DiceScore^{AV}(2592p, 2592p).

Как видно из табл. 3, увеличение разрешения изображений из БД LIVECell с 704p до 1472p не приводит к улучшению результатов дообучения на наборе данных Glioma C6 по сравнению с использованием энкодеров, обученных на ImageNet. Дальнейшее увеличение разрешения до 2592p и вовсе ухудшает метрику. Это может быть связано с потерей на изображениях LIVECell высокочастотных признаков после применения билинейной интерполяции, что делает сформированные веса менее пригодными для дообучения на Glioma C6.

Таким образом, веса ImageNet показывают лучшие результаты, поскольку образы объектов из ImageNet характеризуются высокой степенью априорной неопределенности их размеров.

Поскольку размер 1472p является предельно возможным для обучения UNet-MiT-B3 и из табл. 3 видно, что при равных условиях тестирования эта модель показывает результаты лучше, чем UNet-VGG19, были исследованы возможности использования различных аугментирующих техник при предобучении UNet-MiT-B3 для того, чтобы повысить метрику DiceScore^{AV}(2592p, 2592p) без непосредственного обучения на изображениях такого размера.

Табл. 3. Результаты дообучения моделей на Glioma C6. Для оценки DiceScore^{AV} в сеть подавалось разрешение, равное разрешению, использованному для дообучения на Glioma C6

Архитектура	Предобучение. База данных (разрешение)	Дообучение. Метод обработки (разрешение)	DiceScore ^{AV} (x, 2592p)
UNet-VGG19	ImageNet	resize (1472p)	0,904
UNet-VGG19	LIVECell (704p)	resize (1472p)	0,901
UNet-VGG19	LIVECell (1472p)	resize (1472p)	0,901
UNet-VGG19	ImageNet	crop (1472p)	0,903
UNet-VGG19	ImageNet	resize (2592p)	0,907
UNet-VGG19	LIVECell (704p)	resize (2592p)	0,903
UNet-VGG19	LIVECell (1472p)	resize (2592p)	0,902
UNet-VGG19	LIVECell (2592p)	resize (2592p)	0,900
UNet-MiT-B3	ImageNet	resize (1472p)	0,905
UNet-MiT-B3	LIVECell (704p)	resize (1472p)	0,904
UNet-MiT-B3	LIVECell (1472p)	resize (1472p)	0,904
UNet-MiT-B3	ImageNet	crop (1472p)	0,905
UNet-MiT-B3	ImageNet	crop (1472p)	0,903
UNet-MiT-B3	LIVECell (704p)	crop (1472p)	0,899

В этом случае адаптация размеров признаков LIVECell при предобучении может быть достигнута с использованием аугментации на основании случайного извлечения и масштабирования фрагмента

(random resized crop). Суть данного метода заключается в том, чтобы вырезать случайный фрагмент изображения произвольного размера, имеющего масштаб $0 < s \leq 1$ относительно исходного изображения, и затем

привести этот фрагмент к заданному разрешению. На рис. 5 представлен пример того, как выполняется данная процедура.

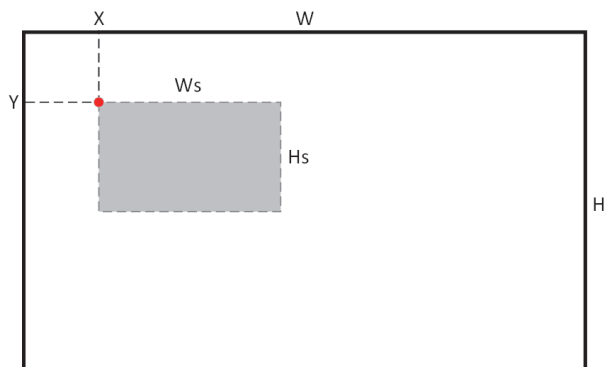


Рис. 5. Принцип извлечения фрагментов изображений в методе random resized crop, где W и H обозначают ширину и высоту исходного изображения, параметр s – процентное соотношение размера фрагмента к общему изображению, а X и Y – координаты случайной точки, выбранной в качестве верхнего левого угла фрагмента

Если исходить из того, что все полученные фрагменты приводились к максимально возможному для UNet-MiT-B3 разрешению 1472p, а отношение размера объектов к размеру изображения является одинаковым для баз данных LIVECell и Glioma C6, то для обеспечения согласованности размеров объектов во время предварительного обучения следует использовать масштабирующий коэффициент $s^* = 57\%$ (1472/2592) (для большей наглядности отношение размера объектов к размеру изображения является одинаковым для баз данных LIVECell и Glioma C6 далее используется описание масштабного коэффициента в процентах). Исследовались узкий и широкий диапазоны относительно оптимального: $s \pm 5\%$ и $s^* \pm 30\%$. В процессе аугментации s выбирался для каждого примера из заданного диапазона.

Оценка влияния различных техник аугментации LIVECell на распознавание Glioma C6 в максимальном разрешении приведена в табл. 4.

Как можно заметить из результатов, приведенных в табл. 4, использование метода аугментации на основании случайного извлечения и масштабирования фрагмента изображения обеспечивает положительный вклад в результаты. Широкий диапазон масштабирования показывает более высокое значение метрики $\text{DiceScore}^{\text{AV}}$ по сравнению с узким, вероятно, благодаря тому, что охватывает большее разнообразие размеров фрагментов изображений. Кроме того, добавление дополнительных техник аугментаций также позволяет увеличить значение метрики $\text{DiceScore}^{\text{AV}}$, достигнув наилучшего результата для UNet-MiT-B3 среди всех проведенных экспериментов, включая обучение 2592p с использованием весов ImageNet. Заметим при этом, что использование чрезмерно большого количества аугментирующих техник в конечном счете приводит к ухудшению результатов, что можно объяснить

смещением распределения обучающих данных от тестового множества [56]. Использование техник случайного формирования цепочек аугментирующих преобразований фиксированной длины (например, [57]), вероятно, может позволить еще улучшить результаты.

Табл. 4. Влияние параметров аугментации LIVECell при предварительном обучении UNet-MiT-B3 на точность сегментации Glioma C6 в максимальном разрешении. Обучение на Glioma C6 всегда проводится на фрагментах 1472p из исходных снимков (без использования масштабирования)

Размер фрагментов, %	Зеркальные отражения	Изменение резкости	Изменение контраста	$\text{DiceScore}^{\text{AV}}$ (2592p, 2592p)
-	-	-	-	0,899
57±5	-	-	-	0,903
57±30	-	-	-	0,904
57±30	+	-	-	0,903
57±30	+	+	-	0,906
57±30	+	-	+	0,906
57±30	+	+	+	0,903

Также был проведен анализ воздействия аугментации при дообучении моделей на БД Glioma C6. Результаты данного эксперимента приведены в табл. 5.

Табл. 5. Результаты дообучения моделей на фрагментах изображений размером 1472p из Glioma C6 с использованием различных методов аугментации. Предварительное обучение проводилось на LIVECell с использованием широкого диапазона масштабов фрагментов изображений ($57 \pm 30\%$)

Методы аугментации	$\text{DiceScore}^{\text{AV}}$ (2592p, 2592p)
-	0,904
Оба зеркальных отражения	0,904
Оба зеркальных отражения и повышение резкости	0,903
Оба зеркальных отражения и повышение контраста	0,902

Как видно, использование дополнительных методов аугментации изображений не привело к повышению значения метрики $\text{DiceScore}^{\text{AV}}$. Это может быть связано с высокой однородностью имеющейся выборки Glioma C6.

На рис. 6 представлены примеры бинарной сегментации клеток глиомы C6 на изображениях без подготовленной специалистами разметки. Маска сегментации была получена моделью UNet-MiT-B3.

На рис. 6а коричневыми являются плотные кластеры клеток, которые не были представлены в Glioma C6. Видно, что использование разных стартовых весов позволяет получать разные свойства обобщения: веса LIVECell (рис. 6г – е) позволяют исключать кластеры из рассмотрения, чего не удастся добиться при использовании лишь весов ImageNet (рис. 6б – в). Такое свойство является важным для проведения исследований, когда требуется оценивать жизненный цикл клеток, а также в случае, если возможно появление других чужеродных тел.

Кроме того, видно, что использование изображений с более высоким разрешением при дообучении на Glioma C6 повышает точность сегментации, а применение дополнительных методов аугментации в процессе предобучения на LIVECell значительно улучшает выявление кластеров.

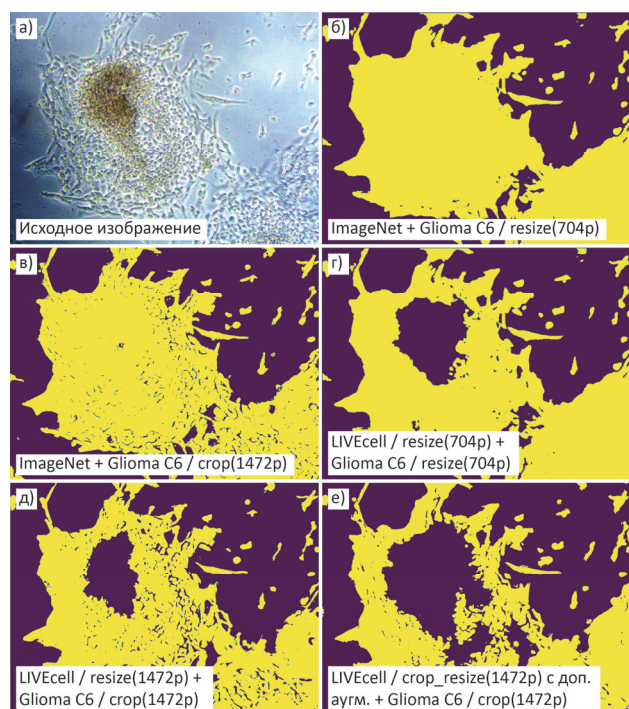


Рис. 6. Качественное сравнение результатов. Результат бинарной сегментации на неразмеченном изображении.

В подписях указаны обучающая и тестовая БД соответственно, а в скобках – разрешение

Влияние архитектуры энкодера. Максимальное значение метрики в разрешении 2592p (0,907) было достигнуто моделью UNet-VGG19, использующей веса ImageNet и обученной на изображениях Glioma C6 с разрешением 2592p. Однако, учитывая, что модели UNet-VGG19 (полностью сверточная) и UNet-MiT-B3 (с использованием слоев внимания) продемонстрировали близкие результаты при обучении в равных условиях, можно предположить, что при отсутствии аппаратных ограничений модель UNet-MiT-B3 могла бы достичь аналогичных или даже превосходящих результатов.

Предварительное обучение моделей на наборе данных LIVECell с использованием более сложных техник аугментации ([56–57]), вероятно, также могло бы оказать положительное влияние на конечные результаты.

Заключение

Нами была собрана БД фазово-контрастных снимков клеток глиомы C6 (Glioma C6). С её использованием разработана модель на основе модифицированной сети U-Net, выполняющая бинарную сегментацию собранной БД. Собранная база данных содержит изображения клеток в высоком разрешении, что позволило нам исследовать свойства методов трансферного обучения клеточных культур в этих условиях. Мы выявили значительную

зависимость эффективности переноса признаков, сформированных при обучении на микроскопических снимках, от согласованности размеров объектов во время обучения и дообучения. Мы предложили метод обеспечения необходимой согласованности размеров объектов, основанный на извлечении и масштабировании фрагментов изображений с применением различных техник аугментации. Этот подход был применен при предобучении моделей на LIVECell и позволил превзойти результаты, достигнутые при использовании весов ImageNet, для архитектуры энкодера, использующей слой внимания.

Благодарности

Разработка алгоритма сегментации клеток глиомы, включающая исследование методов трансферного обучения, выполнена при финансовой поддержке гранта Санкт-Петербургского научного фонда (№ 23-РБ-05-24 от 15.12.2023) (параграф 1, параграф 2, подпараграф 3.1, подпараграф 3.2).

Работа в части исследования общих свойств архитектур сегментации поддержана средствами федерального бюджета в рамках государственного задания ФГБУН Институт физиологии им. И.П. Павлова РАН (№1021062411653-4-3.1.8) (подпараграф 3.1).

Создание базы данных Glioma C6 выполнено при финансовой поддержке БРФФИ (ГКНТ) по гранту «Детекция опухолевых клеток нервной ткани с применением методов глубокого обучения» (договор № M24СПБГ-010 от 01.03.2024) (параграф 2, подпараграф 3.2).

References

- [1] Giakoumettis D, Kritis A, Foroglou N. C6 cell line: the gold standard in glioma research. Hippokratia 2018; 22(3): 105-112.
- [2] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In Book: Navab N, Hornegger J, Wells WM, Frangi AF, eds. Medical image computing and computer-assisted intervention – MICCAI 2015. 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III. New York: Springer International Publishing Switzerland; 2015: 234-241. DOI: 10.1007/978-3-319-24574-4_28.
- [3] 2018 Data Science Bowl. 2018. Source: <<https://www.kaggle.com/c/data-science-bowl-2018>>.
- [4] Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods 2021; 18: 100-106. DOI: 10.1038/s41592-020-01018-x.
- [5] Cutler KJ, Stringer C, Lo TW, et al. Omnipose: a high-precision morphology-independent solution for bacterial cell segmentation. Nat Methods 2022; 19: 1438-1448. DOI: 10.1038/s41592-022-01639-4.
- [6] Edlund C, Jackson TR, Khalid N, et al. LIVECell – A large-scale dataset for label-free live cell segmentation. Nat Methods 2021; 18: 1038-1045. DOI: 10.1038/s41592-021-01249-6.
- [7] Pachitariu M., Stringer C. Cellpose 2.0: how to train your own model. Nat Methods 2022; 19: 1634-1641. DOI: 10.1038/s41592-022-01663-4.
- [8] Li G, Xiao Z. Transfer learning-based neuronal cell instance segmentation with pointwise attentive path fusion. IEEE Access 2022; 10: 54794-54804. DOI: 10.1109/ACCESS.2022.3176956.

- [9] Guo F. Segmenting neuronal cells in microscopic images using cascade mask R-CNN. 2022 2nd Int Conf on Medical Imaging, Sanitation and Biological Pharmacy (MISBP 2022) 2022: 1-7.
- [10] Huynh HN, Ha QT, Tran AT, Tran HDT, Bui PA, Tran TN. Brain cell segmentation and detection from the LIVECell dataset using deep learning with the EfficientDet model. 2022 IEEE 4th Eurasia Conf on Biomedical Engineering, Healthcare and Sustainability (ECBIOS) 2022; 12-15. DOI: 10.1109/ECBIOS54627.2022.9945050.
- [11] Zhou Y, Li W, Yang G. SCTS: Instance segmentation of single cells using a transformer-based semantic-aware model and space-filling augmentation. 2023 IEEE/CVF Winter Conf on Applications of Computer Vision (WACV) 2023: 5944-5953. DOI: 10.1109/WACV56688.2023.00589.
- [12] Liao W, Li X, Wang Q, Xu Y, Yin Z, Xiong H. CUPre: Cross-domain unsupervised pre-training for few-shot cell segmentation. arXiv Preprint. 2023. Source: <https://arxiv.org/abs/2310.03981>. DOI: 10.48550/arXiv.2310.03981.
- [13] Pan F, Wang F. DCSN: A flexible and efficient lightweight network for dense cell segmentation. 2023 IEEE Int Conf on Medical Artificial Intelligence (MedAI) 2023: 334-343. DOI: 10.1109/MedAI59581.2023.00052.
- [14] Liu Y, Wang C, Wen Y, Huo Y, Liu J. Efficient segmentation algorithm for complex cellular image analysis system. IET Control Theory Appl 2023; 17: 2268-2279. DOI: 10.1049/cth2.12466.
- [15] Khalid N, et al. DeepCeNS: An end-to-end pipeline for cell and nucleus segmentation in microscopic images. 2021 Int Joint Conf on Neural Networks (IJCNN) 2021: 1-8. DOI: 10.1109/IJCNN52387.2021.9533624.
- [16] Khalid N, Koochali M, Rajashekhar V, et al. DeepMuCS: A framework for mono- & co-culture microscopic image analysis: From generation to segmentation. TechRxiv. 2022. Source: <https://d197for5662m48.cloudfront.net/documents/publicationstatus/163785/preprint_pdf/5e0aa768c6c9b2f35e11c54e8de4f71c.pdf>. DOI: 10.36227/techrxiv.19181552.v1.
- [17] Khalid N, et al. DeepMuCS: A framework for co-culture microscopic image analysis: From generation to segmentation. 2022 IEEE-EMBS Int Conf on Biomedical and Health Informatics (BHI) 2022: 1-4. DOI: 10.1109/BHI56158.2022.9926936.
- [18] Schilling MP, et al. AI2Seg: A method and tool for AI-based annotation inspection of biomedical instance segmentation datasets. 2023 45th Annual Int Conf of the IEEE Engineering in Medicine & Biology Society (EMBC) 2023: 1-6. DOI: 10.1109/EMBC40787.2023.10341074.
- [19] Ying Z, Li G, Ren Y, Wang R, Wang W. A new image contrast enhancement algorithm using exposure fusion framework. In Book: Felsberg M, Heyden A, Krüger N, eds. Computer analysis of images and patterns. 17th International Conference, CAIP 2017, Ystad, Sweden, August 22-24, 2017, Proceedings, Part II. 2017; 36-46. DOI: 10.1007/978-3-319-64698-5_4.
- [20] Tan M, Pang R, Le QV. EfficientDet: Scalable and efficient object detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2020: 1-10. DOI: 10.1109/CVPR42600.2020.01079.
- [21] Liu Z, et al. Swin transformer: Hierarchical vision transformer using shifted windows. arXiv Preprint. 2021. Source: <https://arxiv.org/abs/2103.14030>. DOI: 10.48550/arXiv.2103.14030.
- [22] Schwendy M, Unger RE, Parekh SH. EVICAN – A balanced dataset for algorithm development in cell and nucleus segmentation. Bioinformatics 2020; 36(12): 3863-3870. DOI: 10.1093/bioinformatics/btaa225.
- [23] Lin T-Y, Dollar P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2017: 936-944. DOI: 10.1109/CVPR.2017.106.
- [24] Zhang H, et al. ResNeSt: Split-attention network. 2022 IEEE/CVF Conf on Computer Vision and Pattern Recognition Workshops (CVPRW) 2022: 2736-2746. DOI: 10.1109/CVPRW56347.2022.00309.
- [25] Cai Z, Vasconcelos N. Cascade R-CNN: High quality object detection and instance segmentation. IEEE Trans Pattern Anal Mach Intell 2021; 43(5): 1483-1498. DOI: 10.1109/TPAMI.2019.2956516.
- [26] Yi J, Wu P, Hoepfner DJ, Metaxas D. Pixel-wise neural cell instance segmentation. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) 2018: 373-377. DOI: 10.1109/ISBI.2018.8363596.
- [27] Farhan R. Residual network with attention to neural cells segmentation. Iraqi J Sci 2023; 64(4): 2023-2036. DOI: 10.24996/ijss.2023.64.4.37.
- [28] Panigrahi S, Murat D, Gall AL, et al. Mistic, a general deep learning-based method for the high-throughput cell segmentation of complex bacterial communities. Elife 2021; 10: e65151. DOI: 10.7554/eLife.65151.
- [29] Piotrowski T, Rippel O, Elanzew A, et al. Deep-learning-based multi-class segmentation for automated, non-invasive routine assessment of human pluripotent stem cell culture status. Comput Biol Med 2021; 129: 104172. DOI: 10.1016/j.compbiomed.2020.104172.
- [30] Long F. Microscopy cell nuclei segmentation with enhanced U-Net. BMC Bioinformatics 2020; 21(8): 8. DOI: 10.1186/s12859-019-3332-1.
- [31] Sharan TS, Tripathi S, Sharma S, Sharma N. Encoder modified U-Net and feature pyramid network for multi-class segmentation of cardiac magnetic resonance images. IETE Tech Rev 2021; 39(5): 1092-1104. DOI: 10.1080/02564602.2021.
- [32] Chen J, et al. TransUNet: Transformers make strong encoders for medical image segmentation. arXiv Preprint. 2021. Source: <https://arxiv.org/abs/2102.04306>. DOI: 10.48550/arXiv.2102.04306.
- [33] Qiu P, et al. AgileFormer: Spatially agile transformer UNet for medical image segmentation. arXiv Preprint. 2024. Source: <https://arxiv.org/abs/2404.00122>. DOI: 10.48550/arXiv.2404.00122.
- [34] He K, et al. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016: 770-778. DOI: 10.1109/CVPR.2016.90.
- [35] Atliha V, Šešok D. Comparison of VGG and ResNet used as encoders for image captioning. 2020 IEEE Open Conf of Electrical, Electronic and Information Sciences (eStream) 2020: 1-4. DOI: 10.1109/eStream50540.2020.9108880.
- [36] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv Preprint. 2014. Source: <https://arxiv.org/abs/1409.1556>. DOI: 10.48550/arXiv.1409.1556.
- [37] Iglovikov V, Shvets A. TeraUNet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. arXiv Preprint. 2018. Source: <https://arxiv.org/abs/1801.05746>. DOI: 10.48550/arXiv.1801.05746.
- [38] Computer Vision Annotation Tool (CVAT). 2025. Source: <https://github.com/opencv/cvat>.
- [39] COCO: Common Objects in Context. Data format. 2025. Source: <https://cocodataset.org/#format-data>.

- [40] Volkov AV, Demina YD, Ksenevich YV, Yachnaya VO, Ermachenkova MK, Ilyukhin DA. Preparation of a database for training an artificial neural network for tumour cells segmentation of neural tissue [In Russian]. XXI Mezhdunarodnaya nauchnaya konferenciya "Molodezh V Nauke 2024", 29-31 October, 2024, Minsk, Belarus.
- [41] LIVECell dataset. Train sample. 2025. Source: <https://LIVECell-dataset.s3.eu-central-1.amazonaws.com/LIVECell_dataset_2021/annotations/LIVECell/LIVECell_coco_train.json>.
- [42] LIVECell dataset. Validation sample. 2025. Source: <https://LIVECell-dataset.s3.eu-central-1.amazonaws.com/LIVECell_dataset_2021/annotations/LIVECell/LIVECell_coco_val.json>.
- [43] LIVECell dataset. Test sample. 2025. Source: <https://LIVECell-dataset.s3.eu-central-1.amazonaws.com/LIVECell_dataset_2021/annotations/LIVECell/LIVECell_coco_test.json>.
- [44] Huang G, et al. Densely connected convolutional networks. 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2017: 4700-4708. DOI: 10.1109/CVPR.2017.243.
- [45] Chen Y, et al. Dual path networks. 31st Conference on Neural Information Processing Systems (NIPS 2017) 2017: 30.
- [46] Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. 36th Int Conf on Machine Learning (ICML 2019) 2019: 6105-6114.
- [47] Szegedy, C et al. Inception-v4, inception-ResNet and the impact of residual connections on learning. AAAI'17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence 2017: 4278-4284.
- [48] Xie E, et al. SegFormer: Simple and efficient design for semantic segmentation with transformers. NIPS'21: Proceedings of the 35th Int Conf on Neural Information Processing Systems 2021: 12077-12090.
- [49] Sandler M, et al. MobileNetV2: Inverted residuals and linear bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition 2018: 4510-4520. DOI: 10.1109/CVPR.2018.00474.
- [50] Xie S, et al. Aggregated residual transformations for deep neural networks. 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2017: 1492-1500. DOI: 10.1109/CVPR.2017.634.
- [51] Hu J, Shen L, Sun G. Squeeze-and-excitation networks. 2018 IEEE/CVF Conf on Computer Vision and Pattern Recognition 2018: 7132-7141. DOI: 10.1109/CVPR.2018.00745.
- [52] Lin M, et al. Neural architecture design for GPU-efficient networks. arXiv Preprint. 2020. Source: <<https://arxiv.org/abs/2006.14090>>. DOI: 10.48550/arXiv.2006.14090.
- [53] Radosavovic I, et al. Designing network design spaces. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2020: 10428-10436. DOI: 10.1109/CVPR42600.2020.01044.
- [54] Gao SH, et al. Res2Net: A new multi-scale backbone architecture. IEEE Trans Pattern Anal Mach Intell 2019; 43(2): 652-662. DOI: 10.1109/TPAMI.2019.293875.
- [55] Chollet F. Xception: Deep learning with depthwise separable convolutions. 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR) 2017: 1251-1258. DOI: 10.1109/CVPR.2017.195.
- [56] Hendrycks D, Mu N, Cubuk ED, Zoph B, Gilmer J, Lakshminarayanan B. AugMix: A simple data processing method to improve robustness and uncertainty. arXiv Preprint. 2019. Source: <<https://arxiv.org/abs/1912.02781>>. DOI: 10.48550/arXiv.1912.02781.
- [57] Cubuk ED, Zoph B, Shlens J, Le QV. Randaugment: Practical automated data augmentation with a reduced search space. 2020 IEEE/CVF Conf on Computer Vision and Pattern Recognition Workshops (CVPRW) 2020: 3008-3017. DOI: 10.1109/CVPRW50498.2020.00359.

Приложение А

В табл. 6 приведены некоторые характеристики Glioma C6 и смежных БД. Как видно, собранные нами данные являются трёхканальными изображениями, тогда как в LIVECell, которая имеет схожие визуальные параметры и на которой ведётся предобучение, изображения полутоновые. Также собранные снимки представляют собой изображения высокого разрешения.

Табл. 6. Характеристики БД

Характеристика	Glioma C6	LIVECell [6]	Cellpose [4]	Omnipose [5]	DSB 2018 [3]
Тип изображения	Трёхканальный	Полутоновый	Трёхканальный и полутоновый	Трёхканальный и полутоновый	Трёхканальный
Тип микроскопического снимка	Фазово-контрастный	Фазово-контрастный	Флуоресцентный и фазово-контрастный	Флуоресцентный и фазово-контрастный	Флуоресцентный
Разрешение изображений	2592×1952	520 × 704	до 720×576	до 2038×2038	~256×256 и 512×512
Тип клеток	Живые клетки	Живые клетки	Живые клетки и препараты	Живые клетки и препараты	Окрашенные препараты

Отметим, что несмотря на то, что в Glioma C6 представлены только живые клетки, в ходе работ был отснят и дополнительный материал, а в проведённом исследовании оценивается возможность применения системы на изображениях, нетипичных для смежных БД: на рис.6 на снимке присутствуют плотные кластеры клеток, которые могут быть мёртвыми или открепившимися.

Сведения об авторах

Илюхин Даниил Андреевич, 2000 года рождения, в 2024 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности 09.04.01 «Информатика и вычислительная техника». Работает инженером-исследователем в Институте физиологии им. И.П. Павлова РАН. Область научных интересов: компьютерное зрение, машинное обучение и искусственный интеллект.

E-mail: ne.free532@gmail.com

Ячная Валерия Олеговна, 1997 года рождения, в 2021 году окончила Санкт-Петербургский Государственный университет аэрокосмического приборостроения по специальности 09.04.01 «Информатика и вычислительная техника». Проходит обучение по программе аспирантуры в Санкт-Петербургском государственном университете аэрокосмического приборостроения по специальности 09.06.01 «Информатика и вычислительная техника». Работает младшим научным сотрудником в Институте физиологии им. И.П. Павлова РАН. Область научных интересов: компьютерное зрение. E-mail: yachnaya_valeria@infran.ru

Малашин Роман Олегович, 1987 года рождения, в 2011 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения. В 2014 году защитил кандидатскую диссертацию в Национальном исследовательском университете информационных технологий, механики и оптики по теме: «Методы структурного анализа изображений трехмерных сцен». С 2017 года является руководителем группы нейронных сетей и искусственного интеллекта в институте им. И.П. Павлова РАН. Область научных интересов включает компьютерное зрение, машинное обучение и искусственный интеллект. E-mail: malashinroman@mail.ru

Ермаченкова Маргарита Кареновна, 1999 года рождения, в 2024 году окончила программу магистратуры Национального исследовательского университета ИТМО по специальности 27.04.05 «Инноватика». Работает инженером-исследователем в Институте физиологии им. И.П. Павлова РАН. Область научных интересов: компьютерное зрение, машинное обучение и искусственный интеллект. E-mail: ermachenkova.rita@mail.ru

Волков Арсений Владимирович, 2003 года рождения, обучается на пятом курсе Белорусского государственного университета по специальности «Физика наноматериалов и нанотехнологий». Работает стажером младшего научного сотрудника в Институте физиологии НАН Беларуси. Область научных интересов: клеточное культивирование, машинное обучение и искусственный интеллект, нейрофизиология. E-mail: av.volkov.phys@gmail.com

Пашкевич Светлана Георгиевна, 1972 года рождения, выпускница кафедры физиологии человека и животных биологического факультета БГУ. В 2003 г. защитила кандидатскую диссертацию по теме «Роль системы блуждающего нерва в реализации висцеральных и соматических ноцицептивных рефлексов» по специальности «Физиология». Научные руководители: профессор, д.м.н. В.А. Кульчицкий (Институт физиологии НАН Беларуси), д.м.н. С.А. Поленов (Институт физиологии им. Павлова РАН). С 2024 года заведующий Центром мозга Института физиологии НАН Беларуси. Область научных интересов: экспериментальная нейрофизиология. E-mail: skypasht@mail.ru

Денисов Андрей Анатольевич, 1975 года рождения, кандидат биологических наук, заведующий лабораторией клеточной инженерии и нанобиотехнологий кафедры биофизики Белорусского государственного университета, ведущий научный сотрудник Центра мозга Института физиологии НАН Беларуси. Область научных интересов: нейробиофизика и нейротехнологии; нейросетевые методы обработки данных; сенсоры электрической активности нервной ткани. E-mail: an.denisov@gmail.com

ГРНТИ: 28.23.15

Поступила в редакцию 16 сентября 2024 г. Окончательный вариант – 05 декабря 2024 г.

Transfer learning methods for glioma C6 cell segmentation

D.A. Ilyukhin ¹, V.O. Yachnaya ^{1,2}, R.O. Malashin ^{1,2}, M.K. Ermachenkova ^{1,2}, A.V. Volkov ^{3,4},
S.G. Pashkevich ⁴, A.A. Denisov ^{3,4}

¹Pavlov Institute of Physiology, Russian Academy of Sciences,
Naberezhnaya Makarova 6, Saint-Petersburg, 199034, Russia;

²Saint Petersburg State University of Aerospace Instrumentation,
Bolshaya Morskaya 67, Saint Petersburg, 190000, Russia;

³Institute of Physiology, National Academy of Sciences of Belarus, Akademicheskaya 28, Minsk, 220072, Belarus;

⁴Belarusian State University, Prospekt Nezavisimosti 4, Minsk, 220030, Belarus

Abstract

In this paper, we present an algorithm for binary segmentation of glioma C6 cells using deep learning methods to simplify and speed up the analysis of this culture growth. The first of its kind dataset containing 30 microscopic phase-contrast images of glioma C6 cells is collected to design and test the algorithm. We explore the influence of the encoder architecture in the neural network segmenter on the accuracy of glioma cell segmentation. Transfer learning approaches using the LIVECell dataset of microscopic images and the large ImageNet dataset of non-specialized images are used since the collected dataset contains a relatively small number of images. Experiments show that pre-training the neural network on LIVECell provides a significant advantage in low-resolution glioma cell recognition, with encoders trained on ImageNet providing better results at higher resolution. The paper proposes ways to improve the generalizing abilities of LIVECell weights to work at high resolution by applying augmentation. We demonstrate that using different starting weights allows us to obtain different generalization properties beyond the training set, which can be useful when detecting, or excluding from consideration, other cells in an image.

Keywords: binary segmentation, microscopic images, computer vision, U-Net, neural networks, glioma C6.

Citation: Ilyukhin DA, Yachnaya VO, Malashin RO, Ermachenkova MK, Volkov AV, Pashkevich SG, Denisov AA. Transfer learning methods for glioma C6 cell segmentation. *Computer Optics* 2025; 49(5): 794-804. DOI: 10.18287/2412-6179-CO-1609.

Acknowledgements: The work was partly funded under a grant from the St. Petersburg Science Foundation and a government project of Pavlov Institute of Physiology, Russian Academy of Sciences (№1021062411653-4-3.1.8), and financially supported by BRFFR (SCST) under grant “Detection of tumor cells in nervous tissue using deep learning methods” (contract No. M24SPbG-010).

Author's information

Daniil Andreevich Ilyukhin (b. 2000) graduated from Saint Petersburg State University of Aerospace Instrumentation in 2024, majoring in Computer Science and Engineering. He currently works as a research engineer at the Pavlov Institute of Physiology of RAS. The research interest lies in the field of computer vision, machine learning and artificial intelligence. E-mail: ne.free532@gmail.com

Valeriya Olegovna Yachnaya (b. 1997) graduated from Saint Petersburg State University of Aerospace Instrumentation in 2021, majoring in Computer Science and Engineering. Currently she is a graduate student at Saint Petersburg State University of Aerospace Instrumentation and works as the assistant research worker at the Pavlov Institute of Physiology of RAS. Research interest is computer vision. E-mail: yachnaya_valeria@infran.ru

Roman Olegovich Malashin (b. 1987) graduated from Saint Petersburg State University of Aerospace Instrumentation in 2011. He received Phd degree in University of Information technology, mechanics and optics in 2014. The title of the Phd thesis was “Structural analysis of images of 3-D scenes”. From 2017 he is the head of the Artificial intelligence and neural networks group in Pavlov Institute of Physiology RAS. Research interest lies in the fields of computer vision, machine learning and artificial intelligence. E-mail: malashinroman@mail.ru

Margarita Karenovna Ermachenkova (b. 1999) graduated from ITMO University in 2024, majoring in Innovation Studies. She currently works as a research engineer at the Pavlov Institute of Physiology of RAS. The research interests lie in the fields of computer vision, machine learning, and artificial intelligence. E-mail: ermachenkova.rita@mail.ru

Arseniy Vladimirovich Volkov (b. 2003) studies at the Belarusian State University, majoring in Physics of Nanomaterials and Nanotechnologies. He currently works as junior research assistant at the Institute of Physiology of NAS of Belarus. The research interest lies in the field of cell cultivation, machine learning and artificial intelligence, neurophysiology. E-mail: av.volkov.phys@gmail.com

Svetlana Georgievna Pashkevich (b. 1972) graduated from the Department of Human and Animal Physiology, Faculty of Biology, BSU. In 2003 she defended her PhD thesis on the topic "The role of the vagus nerve system in the implementation of visceral and somatic nociceptive reflexes" (Physiology). Scientific supervisors: professor, MD V.A. Kulchitsky (Institute of Physiology, NAS of Belarus), MD S.A. Polenov (Pavlov Institute of Physiology, RAS). Since 2024, Head of the Brain Center, Institute of Physiology, NAS of Belarus. Research interests: experimental neurophysiology. E-mail: skypasht@mail.ru

Andrey Anatolyevich Denisov (b. 1975) graduated from Physics Department of Belarusian State University, PhD in Biology, Head of the Laboratory of Cell Engineering and Nanobiotechnology, Department of Biophysics, Belarusian State University; Leading Researcher, Brain Center, Institute of Physiology, National Academy of Sciences of Belarus. Research interests: neurobiophysics and neurotechnology; neural network methods of data processing; sensors of electrical activity of nervous tissue. E-mail: an.denisov@gmail.com

Received September 16, 2024. The final version – December 05, 2024.
