

УДК 378:004

DOI: 10.25686/978-5-8158-2474-4-2025-852-865

Доверенная самообучающаяся полимасштабная сегментация полнослайдовых гистологических изображений для диагностики опухолей

В. Рехамния, А. М. Трыкин, Е. П. Васильев, В. Е. Турлапов

Исследовательский центр в области искусственного интеллекта, Нижегородский государственный университет им. Н.И. Лобачевского, Нижний Новгород, Россия

Аннотация. Гистологические полнослайдовые изображения (WSI) являются самым точным и информационно богатым источником для диагностики и исследования опухолей мозга. Они содержат иерархию изображений кратных разрешений, подобную MIP-текстурам, и создают возможность использовать в процессе сегментации признаки разного масштаба одновременно. Использование искусственного интеллекта (ИИ) в этой области тормозится высокой трудоемкостью разметки таких изображений и высокими требованиями доверенности результата ИИ в медицине. Предложен доверенный алгоритм конвейера обработки WSIs, который решает указанные выше проблемы полностью на основе только самообучения, используя в своей структуре только объяснимые блоки: 1) автоэнкодер (АЕ) как инструмент выбора уровня WSI, соответствующего ткани и дополняющего клеточный уровень, имеющий максимальное разрешение; 2) кластеризации сегментируемых изображений по разномасштабным тканевым и клеточным признакам; 3) двухстадийный метод Оцу (TSMO) как инструмент выбора естественных порогов признака для сегментации объектов; 4) сегментации ядер клеток и оценки степеней поражения по геометрическим признакам формы. Исследование работоспособности алгоритма выполнено на неразмеченном датасете UPenn-GBM. Среднеквадратическая ошибка (MSE) АЕ составила 0,0016. Точность сегментации ядер клеток количественно оценена на размеченном датасете MoNuSeg. Она составила 77,46 % по метрике F1-score, что превосходит результаты аналогов, также использовавших только самообучение.

Ключевые слова: доверенный искусственный интеллект, диагностика опухолей, гистологические полнослайдовые изображения, полимасштабная сегментация WSI, кластеризация, сегментация тканей, сегментация ядра клетки, самообучаемая сегментация, автоэнкодер

Trusted unsupervised multiscale segmentation of whole-slide histological images for tumor diagnosis

W. Rehamnia, A. M. Trykin, E. P. Vasiliev, V. E. Turlapov

Research Center for Artificial Intelligence, Lobachevsky State University, Nizhny Novgorod, Russia

Abstract. Histological whole-slide images (WSI) are the most accurate and information-rich source for diagnostics and research of brain tumors. They contain a hierarchy of images of multiple resolutions, similar to MIP textures, and create the ability to use features of different scales simultaneously in the segmentation process. The use of artificial intelligence (AI) in this area is hampered by the high labor intensity of labeling such images and high requirements for the trustworthiness of the AI result in medicine. A trusted algorithm for the WSI processing pipeline is proposed, which solves the above problems entirely on the basis of self-training only, using only explainable blocks in its structure: 1) an autoencoder (AE) as a tool for selecting the WSI level corresponding to the tissue and complementing the cellular level with the maximum resolution; 2) clustering of segmented images by different-scale tissue and cellular features; 3) Two-Stage Multithreshold Otsu method as a tool for selecting natural thresholds of features for object segmentation. 4) segmentation of cell nuclei and assessment of damage degrees based on geometric shape features. The algorithm performance study was performed on the unlabeled UPenn-GBM dataset. The mean square error (MSE) AE was 0.0016. The accuracy of cell nucleus segmentation was quantitatively assessed on the labeled MoNuSeg dataset. It was 77.46 % by the F1-score metric, which exceeds the results of analogs that also used only unsupervised models.

Keywords: trusted artificial intelligence, tumor diagnostics, histological full-slide images, multiscale WSI segmentation, clustering, tissue segmentation, cell nucleus segmentation, self-learning segmentation, autoencoder

Введение

Использование результатов гистологии в полнослайдовой (WSI) форме стало стандартом современной цифровой патологии, позволяя проводить детальный многомасштабный анализ образцов тканей с высоким разрешением (Wright, et al., 2013) (Gu, et al., 2018). Вместе с тем гигапиксельные размеры WSI сами по себе создают проблемы для эффективности сегментации (Ronneberger, et al., 2015; Gu, et al., 2018). Достижение высокой точности сегментации эффективных моделей глубокого обучения (DL) требует обширных аннотаций на уровне пикселей, выполняемых экспертами. При этом результаты разметки одного датасета часто не обобщаются на различные клинические образцы и тем

более датасеты (Peng and Wang, 2021), (Bejnordi, et al., 2017). Непрозрачность моделей DL также не соответствует требованиям объяснимости и доверенности, которые обязательны для медицины (Tajbakhsh, et al., 2020). Наличие шума или смещения данных могут существенно снизить точность модели (Brixtel, et al., 2022; Roy, et al., 2024).

В результате появился интерес к разработке алгоритмов сегментации без учителя. Так, ученые Tajbakhsh и др. (2020) разработали методы самообучения для сегментации ядра клетки, достигшие оценки Dice score 74,77 %. Javed и др. (2024) предложили использовать для сегментации WSI взаимное обучение трансформеров (transformer) без учителя. Предложены и различные подходы неглубокого обучения. Например, Kochetov и др. (2024), Zhang и др. (2024) предложили алгоритм сегментации ядра на основе неглубокого обучения с оценкой F1 66,70 % и 70-75 % соответственно. Хотя эти методы более вычислительно эффективны и интерпретируемы, чем модели глубокого обучения, они не достигли необходимого уровня устойчивости и точности, необходимых в клинической практике. Их важным недостатком является также ограниченность только сегментацией ядра на уровне самого высокого разрешения WSI.

Использование машинного обучения (МО) в поиске изображений по их содержанию стало темой интереса в недавних исследованиях. Ozturk (2020) предложили метод разметки высокоуровневых признаков на основе стекового автокодировщика для поиска медицинских изображений по содержанию. Guo и др. (2020) сосредоточились на поиске изображений с сохранением конфиденциальности (PPIS) с использованием сверточных нейронных сетей (CNN) для безопасной классификации и поиска по закодированным медицинским изображениям большого размера. Singh с соавторами (2021) предложили модель с четырьмя сверточными слоями для поиска изображений по содержанию в данных о знаменитостях с использованием глубоких CNN. Kanwal и др. (2020) предложили метод, который объединяет симметрию, оценки FAST, фильтрацию формы и положения с CNN для поиска изображений большого размера, и достигли высокоточных результатов на сложных датасетах. Komura и др. (2022) исследовали универсальное кодирование гистологии паннаковых заболеваний с помощью глубоких текстурных оценок. Samalan с соавторами (2020) разработали систему поиска изображений на основе контента для изображений барабанной перепонки с использованием глубокого обучения. Zhuo и Zhou (2020) предложили метод дискриминантного представления низкой размерности для поиска изображений дистанционного зондирования Земли с высоким разрешением. Деерак и Амеер (2020) исследовали поиск опухолей в изображениях магнитно-резонансной томографии (МРТ) мозга, используя оценку сходства по контрастным потерям при кодировании в GoogLeNet. Kumar и др. (2021) представили U-Net-подобную нейронную сеть, реализующую поиск изображений по объектам присутствующим на изображении. Улучшены характеристики точности, достоверности и полноты. Monowar с соавторами (2022) представили AutoRet, самообучаемую рекуррентную сеть для поиска изображений на основе контента. В совокупности эти исследования демонстрируют разнообразные применения и достижения в методах нейронных сетей для поиска изображений по содержанию в различных медицинских и немедицинских областях.

Сегментация WSI без учителя использует, как правило, передовые методы машинного обучения для анализа гистологических изображений без необходимости обширного ручного аннотирования. Предложенные методы варьируются от слабообучаемых, самообучаемых до полностью самообучаемых подходов. Gadermaier и др. (2019) исследовали четыре метода и их комбинации для образования нового метода сегментации без учителя. Первый метод, SDS, использует модель U-Net для точной обучаемой сегментации. Второй метод преобразует входное изображение из размеченного датасета, имитируя замену красителя на альтернативный перед подачей на U-Net, и таким образом решает проблему разметки датасета с другим красителем. Третий метод, SDU, обсуждавшийся в их более раннем исследовании (Gadermaier, et al.), включает создание синтетических датасетов (с готовой аннотацией). Далее некоторый размеченный датасет может быть аугментирован с помощью генеративно-состязательной сети (GAN). Одним из ограничений данного исследования является использование синтетических изображений, основанных на визуальных предубеждениях GAN.

В статье Sebai и др. (2020) представлен целый фреймворк глубокого обучения на основе RCNN для обнаружения митозов на гистопатологических изображениях, включающий спектр подходов: строго обучаемые, слабообучаемые и без учителя. В работе Li и др. (2022) предложен фреймворк сегментации раковых клеток без учителя для гистопатологических изображений, использующий архитектуру глубокой сети U-Net. Авторы использовали самообучение (контрастное обучение) для этапа аугментации данных и показали, что оно более эффективно для производительности модели и снижения сложности, чем альтернативные стратегии аугментации (по крайней мере, в среде без учителя).

В исследовании van Nierop и др. (2024) представлена модель сегментации клеток без учителя, которая оптимизирует взаимную информацию между парами данных, достигая точной сегментации клеток без разметки экспертом. Их модель превзошла SOTA обучаемых и самообучаемых моделей, достигнув оценки Жаккара 0,71 на собственном наборе данных. Cisternino с соавторами (2024) использовали самообучение на огромном наборе данных из 1,7 миллиона изображений для сегментации тканей и прогноза экспрессии РНК. Их стратегия превзошла существующие методологии, улучшив оценку силуэта на 43 %, коррелируя морфологию ткани с экспрессией генов. Этот метод превзошел несколько моделей с учителем, продемонстрировав свою эффективность в исследованиях живых клеток.

Хотя предыдущие исследования успешно решили проблему разметки данных, они по-прежнему страдают из-за высоких вычислительных ресурсов и проблем с доверенностью. Поэтому Zhang и др. (2024) предложили более быстрое и интерпретируемое решение, использующее для сегментации ядер несколько преобразований цветовых моделей с k-средними и нечеткими c-средними (FCM). Кроме того, в комплексном обзоре Wei и др. (2024) выявили несколько будущих пробелов, включая интерпретацию глубокой кластеризации, выбор гиперпараметров (DL и параметры кластеризации), многовидовую глубокую кластеризацию (как показано в работе Chen и др. (2023)). Эти результаты вдохновляют на дальнейшие исследования более совершенных алгоритмов для полностью самообучаемой сегментации гистопатологических данных без потери ее точности.

Постановка задачи

Из Введения следует, что наилучшим современным решением для сегментации гистологических WSI-изображений является подход, применяющий объяснимый и доверенный искусственный интеллект, максимально использующий предобработку, учитывающую все уровни WSI. То есть подход должен быть и полимасштабным и не требующим ручной разметки (т.е. unsupervised). Использование низких (тканевых) уровней разрешения WSI для повышения точности сегментации ядер в клетках мотивировано достижениями Kochetov и др. (2024) и Zhang с соавторами (2024), а объяснимость и доверенность результатов конвейера может быть достигнута посредством использования наглядных высокоуровневых признаков сегментации.

Говоря о доверенности подхода мы полагаемся на наше системное видение понятия доверенного искусственного интеллекта (Trusted Artificial intelligence, TAI), удовлетворяющее, тем не менее, существующим стандартам. Предлагается системное определение понятия доверенного ИИ, базированное на трёх основных свойствах-требованиях, понимаемых как категории:

Доверенный ИИ – это искусственный интеллект, удовлетворяющий трем категориальным требованиям: 1) управляемость человеком; 2) безопасность; 3) объяснимость.

В случае медицины или здравоохранения эти три свойства должны быть реализованы для треугольника «Врач–ИИ–Пациент», требующего доверенных отношений.

Все остальные требования, раскрытые в стандартах, выводятся из указанных трёх, раскрывая категории. Например,

- Человек: врач, пациент, семья, коллектив, население, правовые акты предприятий, управлений, министерств, субъектов государства,...
- Врач: человек, бригада скорой помощи, медицинское отделение, клиника, функциональное подразделение здравоохранения, клинические рекомендации...

- Пациент: человек, семья, коллектив предприятия или отрасли, общественная организация, население территории, граждане государства.

В данной работе сделана попытка обеспечить решению задачи сегментации WSI объяснимость и безопасность для представителей всех участников тройки «Врач–ИИ–Пациент» и контролируемость результата человеком. Наши решения в направлении построения алгоритма сегментации опухолей по WSI-данным гистологических изображений, который мог бы существенно повысить доверенность результата, предлагаются в следующих разделах.

Теория

1. Датасеты

В данном исследовании мы использовали набор данных UPenn-GBM, включающий гистологические, MPT-данные и данные пациентов с глиобластомой, предоставленные Bakas и др. (2022). Мы ограничились его гистологической частью, включающей 71 WSIs в формате NDPI. Эти слайды получены с 9 различными уровнями увеличения, хотя для повышения вычислительной эффективности библиотека OpenSlide Python интерпретирует их как 18 уровней. Примеры фрагментов, извлеченных из этих слайдов, показаны на рисунке 1.

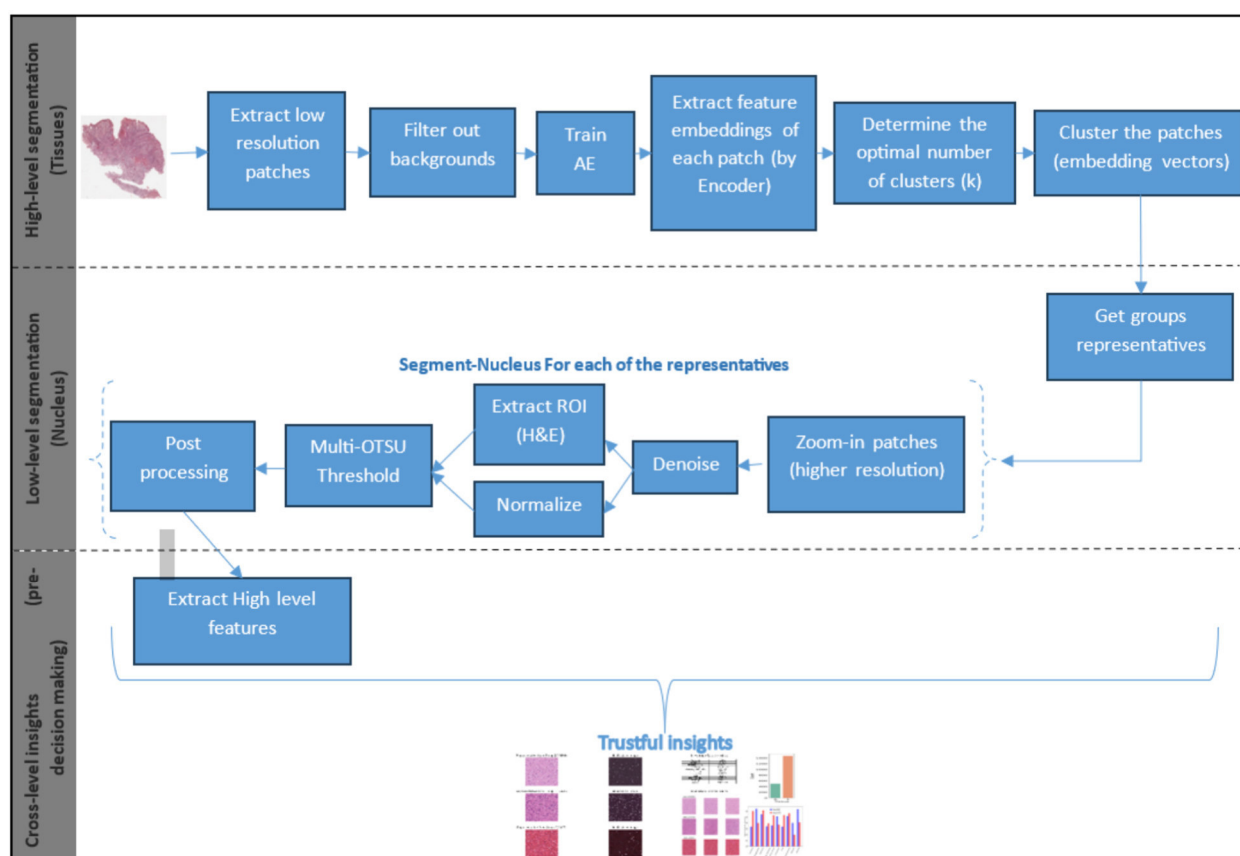


Рисунок 1. Структура алгоритма доверенной обработки: 1) уровень сегментации тканей (верхний ряд блоков); 2) уровень сегментации ядер (второй ряд блоков) – максимальное разрешение; 3) экстракция признаков высокого уровня, объединяющих оба уровня, для принятия доверенного решения

Учитывая, что набор данных UPenn-GBM не содержит какой-либо разметки, для оценки точности конвейера сегментации ядра и ее переносимости на другие датасеты дополнительно потребовался размеченный датасет. Для этой цели мы использовали датасет MoNuSeg, представленный Kumar и др. (2017). Этот набор данных содержит хорошо аннотированные изображения тканей, полученные из различных органов в различных больницах.

Тестовый набор состоит из 14 изображений TIF (по два образца гистопатологии из семи органов: молочной железы, почки, печени, предстательной железы, мочевого пузыря, толстой кишки, желудка), каждое размером $1024 \times 1024 \times 3$, с их аннотациями в формате XML.

2. Предлагаемый алгоритм

Предлагается алгоритм доверенной обработки WSI, структура которого представлена на рисунке 1. Эта структура должна обеспечивать решение трех групп задач: сегментация тканей; сегментация ядер клеток; кросс-уровневые решения.

Сегментация тканей (низкое разрешение). На этом этапе распознаются ткани, представленные в данных, и отображаются их области. Начинают с разбиения нижнего уровня WSI на более мелкие патчи, отфильтровывая их фоновую часть на основе статистических порогов, после чего автоэнкодер (АЕ) обучается реконструировать эти патчи для уровней тканей. После обучения кодирующая часть АЕ извлекает скрытое пространство признаков размером 256 (2x2x64) для каждого патча (встраивание признаков). Патчи кластеризуются с помощью алгоритма k-средних на основе этих признаков (гиперпараметр k определяется оптимальным срезом иерархической кластерной дендрограммы). Наконец, используя подход учета неопределенности (основанный на расстоянии до центроидов), для каждой ткани определяют N наиболее репрезентативных патчей (ближайших к центроиду по атрибутам).

Сегментация ядер (высокое разрешение). На данном этапе выполняется сегментация ядер клеток. Начинаем с шага шумоподавления, который удаляет артефакты из патчей. Восстановленный канал окрашивания гематоксилином объединяется с нормализованным патчем. Процесс нормализации описан в Масенко и др. (2009). После объединения выполняется пороговая пост-обработка (Post-processing) оптимизированным двухстадийным методом multi-Otsu (TSMO) (Huang, 2009), который дает вариант маски сегментации ядра, зашумленный квантованными порогами. Шум маски сегментации корректируется постобработкой операциями матморфологии (closing, opening, dilation). Для выбора лучшей маски исследованы три варианта пост-обработки результата TSMO, формирующие три разных маски сегментации: 1) первый использует операции closing, затем opening; 2) второй – одну операцию dilation; 3) третий – медианное размытие первого варианта маски.

Кросс-уровневые решения (принятие предварительного решения, подконтрольного специалисту согласно требованиям ТАИ). На этом этапе извлекаются суммарные признаки из выходных данных двух конвейеров (тканевого – текстурные признаки и ядерного – морфологические) и объединяются с визуализациями, чтобы сделать более быстрым и точным контроль со стороны врача-невропатолога.

3. Подготовка и детали реализации эксперимента

Размеры патчей и уровни WSI устанавливаются эмпирически для обеспечения пространственной согласованности на разных уровнях. Для получения четкой сегментации ядра начинаем с максимального разрешения (или MaxSize) 1024x1024 ($2^{10} \times 2^{10}$) пикселя (MaxSize=1024). Затем уменьшаем разрешение, с каждым уровнем вдвое, до уровня, на котором становится различима только ткань, что занимает Nsteps=6 шагов. Линейный размер патча на этом уровне составляет 16 пикселей; он получается путём подстановки Nsteps=6 в формулу (1):

$$\text{LinearSize}(N\text{steps}) = \text{MaxSize} / (2^{N\text{steps}}). \quad (1)$$

Предлагаемая структура реализована на Python с использованием фреймворка PyTorch. Эксперименты проведены на персональном компьютере с графическим процессором NVIDIA GTX 1060. Для повышения эффективности обучения при высокоуровневой сегментации кодер включает несколько слоёв из предобученной модели ResNet18 (обученной на наборе данных ImageNet), а декодер представляет собой простую сверточную нейронную сеть (со случайными начальными весами). Сеть АЕ обучалась с использованием оптимизатора Adam с размером батча 128, скоростью обучения 0,001 и уменьшением веса $1e-05$, а также с использованием нормализации пакета и ранней остановки (после 10 шагов) в качестве стратегий регуляризации для повышения точности и обобщения. Кроме того, использовали асинхронный режим предварительной выборки патчей с 12 рабочими процессами для эффективного устранения узких мест ввода-вывода, что значительно ускоряет обучение, особенно при слабом графическом процессоре и большом размере пакета (Lin et al., 2018). Проиллюстрированные выше этапы тканевой сегментации исследованы экспериментально. Например, эксперимент показывает, что двухстадийный многоклассовый метод Оцу, как правило, быстрее и гораздо более устойчив к различию фрагментов, чем метод k-средних, который медленнее и иногда требует настройки гиперпараметра k для повышения точности. Набор данных UPenn-GBM разделён на

три подмножества: обучение (70 %) (124494 патча), валидация (15 %) (26678 патчей) и тестирование (15 %) (26678 патчей).

4. Метрики оценки

Среднеквадратическая ошибка (MSE) оценивает ошибку прогнозирования модели для непрерывных результатов. Она рассчитывается по формуле (2), где более низкая MSE предполагает улучшение качества модели (James et al., 2013):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (2)$$

Оценка по Дайсу (DSC), также известная как F1 в контексте классификации, представляет собой гармоническое среднее значение Precision и Recall, что обеспечивает баланс между этими двумя метриками, и особенно полезна в случаях неравномерного распределения классов. Она рассчитывается по формуле (3) или (4), где более высокая оценка предполагает улучшение качества модели (Sokolova and Lapalme, 2009):

$$Dice = 2 |A \cap B| / |A + B| = 2 \cdot TP / (2 \cdot TP + FP + FN) \quad (3)$$

или

$$F1 = 2 \text{Precision} \cdot \text{Recall} / (\text{Precision} + \text{Recall}) = 2 \cdot TP / (2 \cdot TP + FP + FN). \quad (4)$$

Коэффициент сходства Жаккара (JSC), также известный как «пересечение над объединением» (IoU), предсказанного $|B|$ и истинного $|A|$ объектов рассчитывается по формуле (5), где большее значение IoU свидетельствует об улучшении точности модели (Everingham et al., 2010):

$$J(A,B) = |A \cap B| / |A \cup B|. \quad (5)$$

MSE используется для оценки ошибки сегментации тканей автоэнкодером, тогда как Дайс и коэффициент Жаккара применяются для оценки сгенерированных масок ядер при сегментации ядер клеток.

5. Признаки, экстрагированные из уровней WSI, для сегментации гистологических изображений

В качестве признаков сегментации исследовались две группы признаков: 1) текстурные – на основе Gray Level Co-occurrence Matrix (GLCM) плюс цвет окрашивания; 2) морфологические. Текстурные признаки эффективны как для сегментации тканей (включая их состояние), так и для детектирования ядра на фоне клетки. Морфологические признаки наиболее эффективны для определения состояния ядра по его форме. Список признаков получил в итоге следующие 11 позиций.

Текстурные признаки (вычисляются по значениям элементов GLCM-матрицы):

1. *Contrast*: измеряет контраст интенсивности между пикселем и его соседом по всему изображению. Высокий контраст указывает на большую разницу между значениями соседних пикселей.

2. *Homogeneity*: оценивает близость распределения элементов в GLCM-матрице к диагонали. Высокая однородность означает наличие существенных областей с однородной по интенсивности заливкой.

3. *Intensity*: средняя интенсивность пикселей в ядре, отражающая яркость.

4. *Staining Intensity*: интенсивность окрашивания в ядре, часто указывающая на концентрацию конкретного биомаркера или красителя.

5. *Number of Nuclei (num nuclei)*: количество отдельных ядер, обнаруженных в сегментированной области.

Морфологические признаки:

6. *Size*: площадь области ядра, измеренная в пикселях.

7. *Circularity*: мера того, насколько форма ядра близка к идеальному кругу; рассчитывается по формуле (6), где значение 1,0 соответствует идеальному кругу:

$$\text{Circularity} = 4\pi \times \text{Area} / \text{Perimeter}^2. \quad (6)$$

8. *Density*: количество ядер на единицу площади, отражающее плотность их упаковки, рассчитывается по формуле (7):

$$\text{Плотность} = (\text{Количество ядер}) / (\text{Площадь области интереса}). \quad (7)$$

9. *Eccentricity*: отношение расстояния между фокусами эллипса, который наилучшим образом соответствует ядру, к длине его большой оси. Значение близкое к 1 соответствует вытянутой форме, а 0 — идеальному кругу:

$$e = \sqrt{1 - (b/a)^2}, \quad (8)$$

где e — эксцентриситет эллипса; a и b — длина наибольшей/наименьшей оси эллипса.

10. *Spread*: пространственное распределение ядер, описывающее, насколько широко они разбросаны в области; рассчитывается как стандартное отклонение евклидовых попарных расстояний между центроидами ядер.

Результаты экспериментов

1. Сегментация тканей

Обучение было прекращено с помощью техники раннего останова после 27 эпох, поскольку потери валидации перестали понижаться. На рисунке 2 показано, что 16-я эпоха дала наилучшую ошибку валидации 0,00166, в то время как наилучшая ошибка обучения, равная 0,00138, была достигнута последней эпохой. Фаза тестирования демонстрирует ту же потерю валидации.

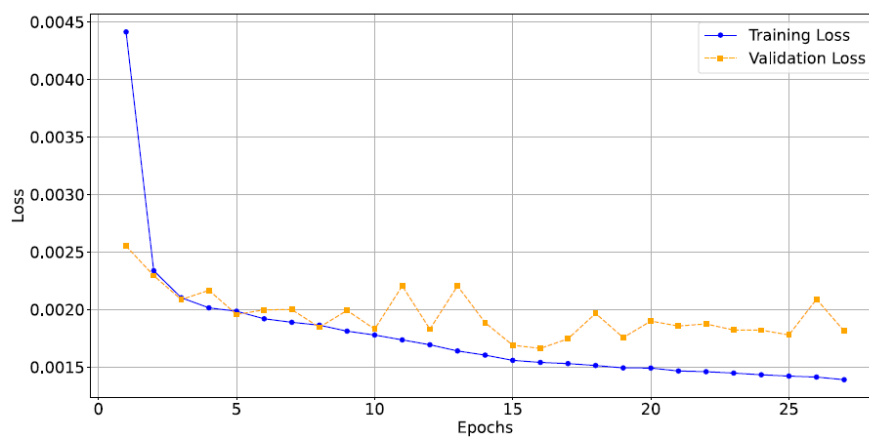


Рисунок 2. Графики потерь и валидации при обучении Training vs validation loss

В ходе обучения, всякий раз как ошибка валидации улучшается, сохраняется рисунок, аналогичный рисунку 3, на котором сравниваются 5-6 случайных участков (парами: исходное разрешение и его реконструкция) для визуальной оценки эффективности модели в задаче захвата важных характеристик ткани.

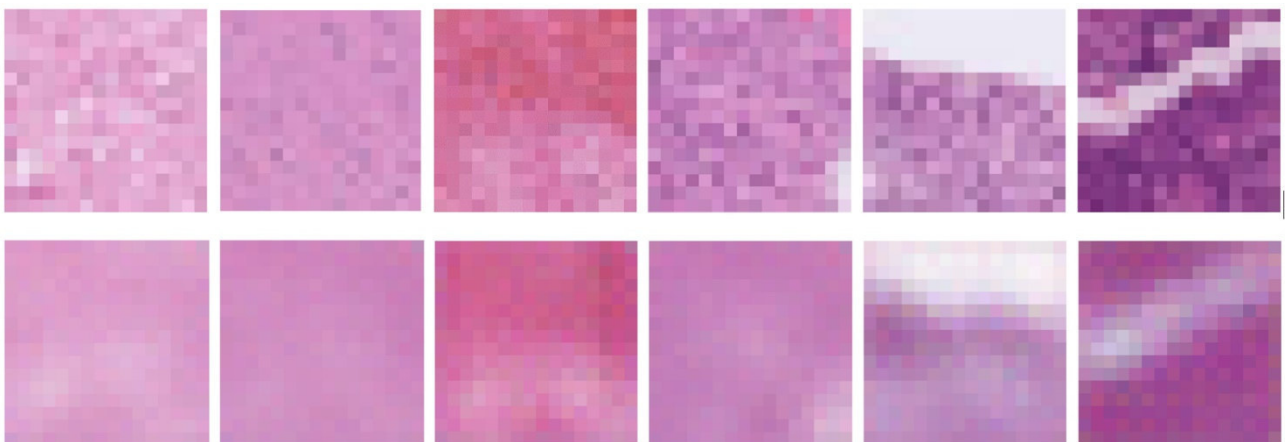


Рисунок 3. Сопоставление 6 пар патчей (фрагментов): исходное разрешение (*верхний ряд*) и его реконструкция (*нижний ряд*) патчей после одной эпохи обучения

После этого фаза глубокой кластеризации использует сглаженное латентное пространство, сформированное кодирующей частью 256 (2x2x64). Для определения подходящего количества групп,

как упоминалось, использована иерархическая дендрограмма кластеризации с использованием метода лучшего отсечения дерева, который предполагает выбор отсечения по наибольшему расстоянию на дендрограмме. На рисунке 4 представлен пример отсечения по наибольшему расстоянию с двумя тканями.

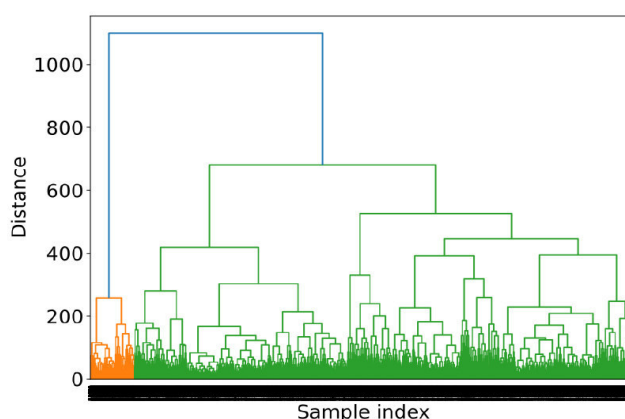


Рисунок 4. Дендрограмма иерархической кластеризации WSI с разделением тканей на два класса (выделены оранжевым и зеленым цветом) по признаку наибольшего расстояния

2. Сегментация ядер клеток

Как показано выше, сегментация тканей включает несколько этапов предварительной обработки и создает три различные маски сегментации. На рисунке 5 показаны этапы и результаты семантической сегментации ядер и три маски, сформированные в идеальном сценарии (без шума).

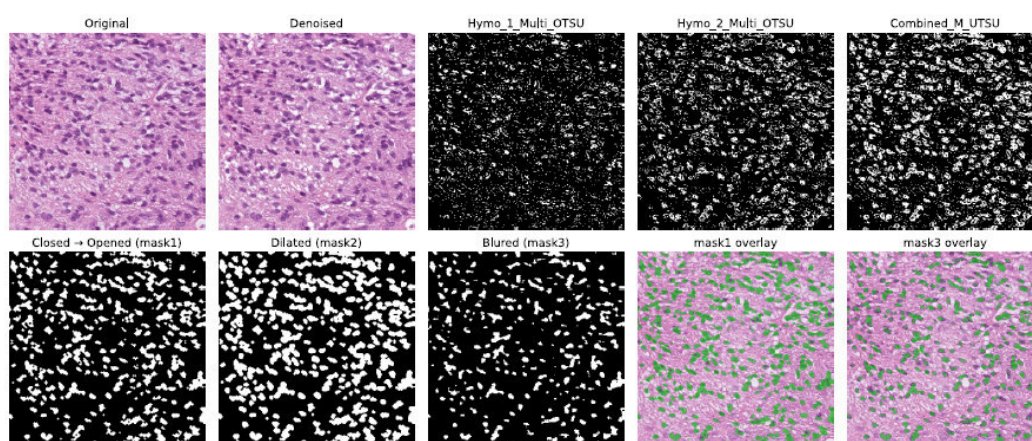


Рисунок 5. Визуализация этапов сегментации ядер клеток (без шума)

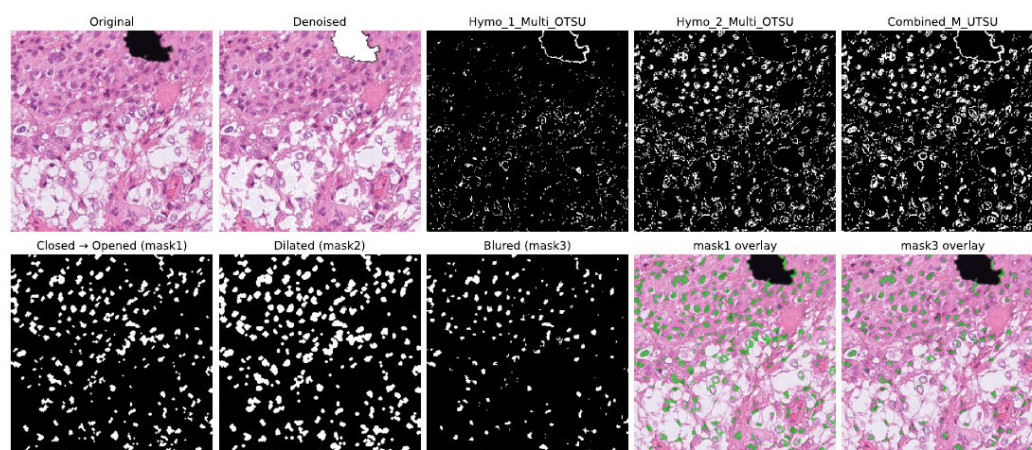


Рисунок 6. Визуализация этапов сегментации ядер клеток (с шумом 1-го типа)

Некоторые WSI из UPenn-GBM содержат значимый шум, поэтому перед дальнейшей обработкой к входному фрагменту применяется алгоритм шумоподавления. Анализ фрагментов выявил два вида шума: чёрную область (рис. 6) и серый шум (рис. 7).

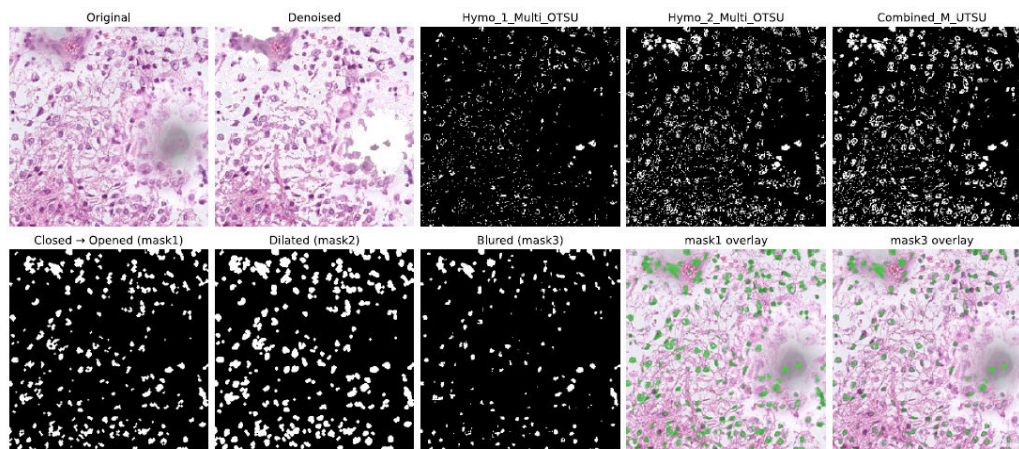


Рисунок 7. Визуализация этапов сегментации ядер клеток (с шумом 2-го типа)

Проведены эксперименты, определившие влияние такого шума на сегментацию ядер клеток и качество удаления шума. Результаты показали, что в отсутствие шумоподавления алгоритм может детектировать шумовые области как большие ядра. А шумоподавление удаляет оба вида шума, не жертвуя важными деталями изображения. Рисунок 7 демонстрирует пример, где метод успешно подавил шум и обнаружил на его месте два скрытых ядра.

3. Сравнение результатов алгоритма с недавними публикациями

Важно также сравнить достигнутую нами точность с точностью современных методов на популярных контролируемых наборах данных. Мы сделали это на датасете MoNuSeg. В таблице 1 показана точность сегментации ядер клеток каждого из 7 органов представленных MoNuSeg, достигнутая тремя нашими масками в метрике F1-score.

Таблица 1. Точность F1-сегментации ядер клеток для 7 органов из датасета monuseg тремя нашими масками

Organ	Mask1	Mask2	Mask3
Bladder	0.7943	0.7323	0.7622
Brain	0.8159	0.7696	0.8115
Breast	0.7418	0.6426	0.7426
Colon	0.7617	0.6408	0.7483
Kidney	0.7842	0.6927	0.7595
Lung	0.7428	0.5957	0.7605
Prostate	0.7703	0.7074	0.7398
Average	0.7746	0.6867	0.7614

Mask1 доминирует для большинства органов, и орган Brain имеет здесь лучшую оценку ($F1=0,8159$), а орган Breast – самую низкую ($F1=0,7418$). Точность Mask3 немного превосходит Mask1 для Breast (0.7426 против 0.7418) и более заметно для Lung (0.7605 против 0.7428). В итоге Mask1 является рекомендуемой.

На рисунке 8 показаны ground truth маски сегментации 7 органов из этого датасета, размеченные экспертами, и 3 маски, построенные нашим алгоритмом.

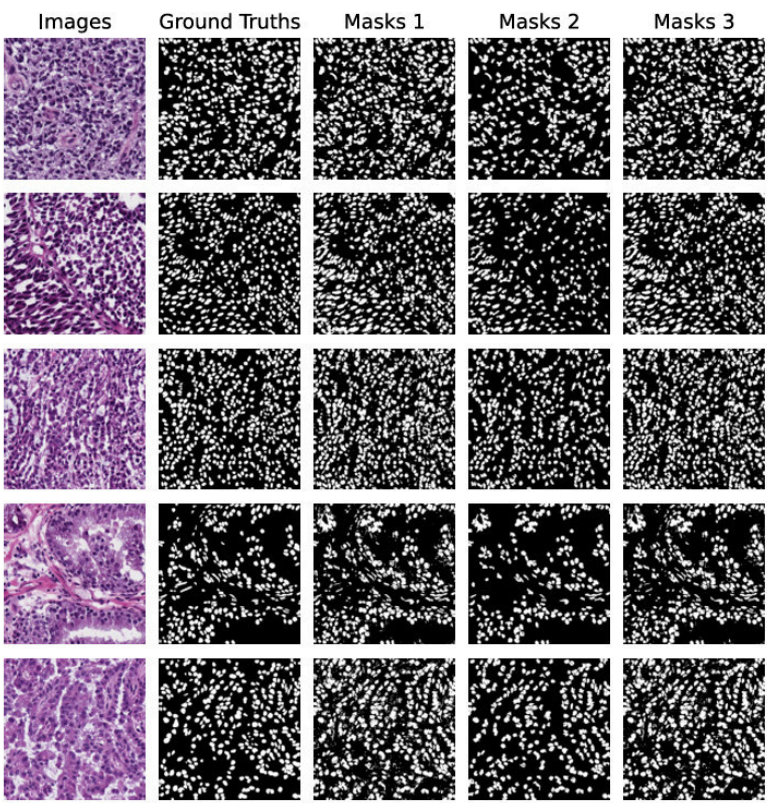


Рисунок 8. Сравнение на 5 образцах тканей датасета MoNuSeg, ручной разметки ядер (Ground Truth) и 3 масок сегментации, сгенерированных нашим алгоритмом

В таблице 2 представлены результаты недавних исследований сегментации ядра также на датасете MoNuSeg, за исключением одного (Kochetov et al., 2024), в котором для оценки использовался другой датасет. Исследования отсортированы в хронологическом порядке, но Supervised и Unsupervised раздельно. Для сравнения использована общая метрика Dice Score.

Таблица 2. Сравнение с недавними исследованиями сегментации ядра

Study	Method	Learning approach	F1(Dice)
Naylor, et al., 2018	max {Mask R-CNN, U-Net, FCN, <u>DIST</u> }	Supervised	78.63
Valanarasu, et al., 2021	LoGo: Local-Global training strategy	Supervised	79.56
Li, et al., 2023	LViT-LW: Language meets Vision Transformer in Medical Image Segmentation	Supervised	80.66
Roy, et al., 2024	GRU-Net: Supervised learning with Gaussian attention and dense skip connections	Supervised	80.35
Showrav and Hasan, 2024	Hi-gMISnet: Generalized segmentation using pGAN with dual-mode attention	Supervised	82.5
Wang, et al., 2024b	U-HRMLP: Multi-scale feature fusion to refine segmentation boundaries	Supervised	80.83
Wang, et al., 2024a	UDTransNet: Learnable skip connections with DAT and DRA modules	Supervised	79.47
Tajbakhsh, et al., 2020	Self-supervised approaches for nucleus segmentation	Self-supervised	74.77
Kochetov, et al., 2024	UNSEG: Bayesian-like framework to segment cells and their nuclei without requiring training data	Unsupervised	66.70
Zhang, et al., 2024	YCbCr color space with: k-means > fuzzy c-means	Unsupervised	75.53
The present study	Unsupervised Poly-scale Histology Segmentation	Unsupervised	77.46

Обсуждение результатов

1. Сравнение результатов UPHS-алгоритма и подходов с обучением

Поскольку наш алгоритм сегментации гистологических изображений может быть охарактеризован как алгоритм, реализованный полностью без учителя и обеспечивающий в процессе сегментации

полимасштабное использование WSI, введем для него краткое наименование UPHS (Unsupervised Poly-scale Histology Segmentation).

Сравнение UPHS с подходами с учителем. Значение оценки F1, полученное с помощью UPHS, сопоставимо с 7 представленными в таблице 2 supervised-подходами. Небольшое снижение оценки обусловлено неконтролируемым характером метода, не требующим разметки ядер для их сегментации. Кроме того, хотя решения на основе глубокого обучения (DL) высокоэффективны, однако часто обучение на одном датасете не переносится в полной мере на другие датасеты. Например, хотя модель GRU-Net (Roy, et al., 2024) превзошла все, кроме одного, решения SOTA на наборе данных MoNuSeg (80,35 %, табл. 2), но при обучении на наборе данных TNBC показала резкое снижение точности до 65,98 % в тестировании на MoNuSeg. В отличие от этого, настоящее исследование продемонстрировало большую устойчивость в сохранении точности.

Сравнение UPHS с подходами без учителя. Таблица 2 показывает, что UPHS превзошел все 3 приведенных в таблице подхода без учителя. Он превосходит байесовский фреймворк UNSEG (Kochetov, et al., 2024) примерно на 11 %, цветовое пространство YCbCr с k-means (Zhang, et al., 2024) примерно на 2 % и подходы с самообучением (Tajbakhsh, et al., 2020) примерно на 3 %.

2. Кроссуровневая аналитика

Обобщённая кроссуровневая WSI аналитика строится следующим образом. На этапе кроссуровневой аналитики для каждой обнаруженной ткани выбирается группа из N (в данном исследовании N принято равным 5) наиболее репрезентативных патчей, которая характеризуется центроидом тканевых признаков. Выбор определяется по критерию удаленности вектора признаков патча от центроида ткани. Выбранные патчи отображаются на соответствующие патчи уровня ядер (с помощью модуля «Open Slide» Python 3.x). Пример выбора для двух тканей показан на рисунке 9.

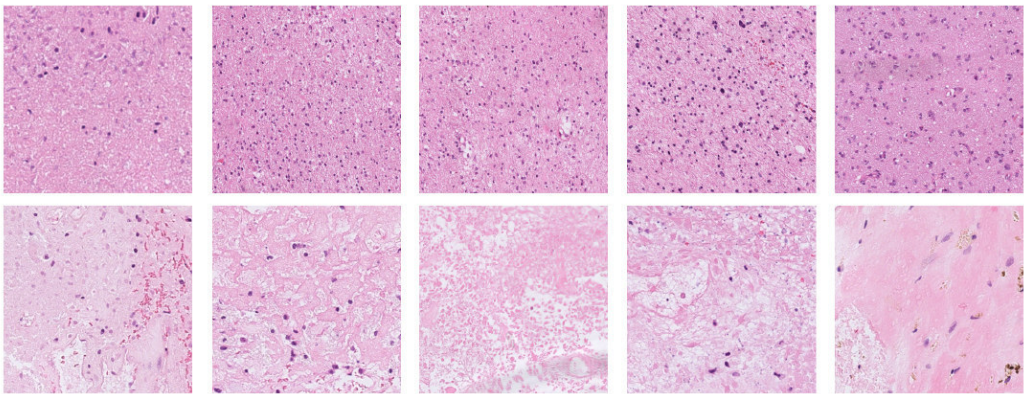


Рисунок 9. Пример 5 репрезентативных патчей для двух тканей

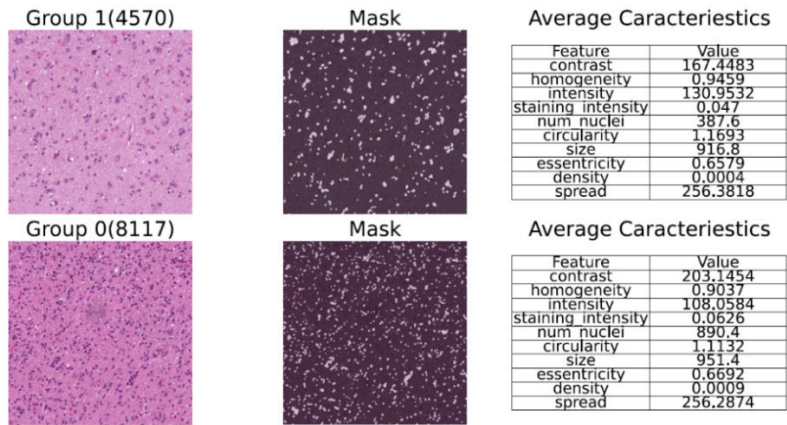


Рисунок 10. Пример репрезентативных патчей 2 тканей (по одному из группы), а также их маски сегментации ядер и векторы признаков, в которые добавлено число ядер и их параметры

Далее выполняются сегментация ядер и расчет их морфологических признаков. На рисунке 10 показаны по одному из пяти (группы) репрезентативных патчей для 2 тканей, а также маски сегментации ядер и вектор признаков каждого патча с указанием числа сегментированных ядер и их усредненных по результатам сегментации параметров.

Признаки патчей могут быть далее использованы для обучения детектированию типа опухоли и степени поражения.

Заключение

Предложен доверенный (стабильный, безопасный, объяснимый и контролируемый человеком) метод сегментации без учителя для анализа гистологических WSIs (полнослайдовых изображений), названный UPHS (Unsupervised Polyscale Histology Segmentation). Отказ от использования учителя устраняет необходимость в ручной разметке датасетов, снижает требования к вычислительным ресурсам при сохранении конкурентоспособной точности семантической сегментации, предоставляет объяснимую информацию всего из двух репрезентативных уровней и включает учет наличия шума в высокоуровневую сегментацию. Эксперименты показали, что наш алгоритм конкурентоспособен, несмотря на свою простоту. Достигнута среднеквадратическая ошибка (MSE) реконструкции уровня тканей равная 0,0016, число признаков для дальнейшей классификации опухоли снижено до 10, получена наилучшая среди unsupervised подходов точность сегментации ядер с оценкой $F1=77,46\%$. Дальнейшие исследования будут сосредоточены на повышении точности сегментации ядер, исследовании эффективности признаков сегментации в задаче классификации опухоли, расширении тестирования надежности алгоритма.

Данные и код, подтверждающие результаты данного исследования, доступны по запросу.

Источник финансирования

Работа выполнена при поддержке Министерства экономического развития Российской Федерации (соглашение о предоставлении гранта № 139-15-2025-004 от 17 апреля 2025 г., ИГК 000000Ц313925P3X0002)

Список литературы

1. Bakas, S., Sako, C., Akbari, H., Bilello, M., Sotiras, A., Shukla, G., Rudie, J.D., Santamaría, N.F., Kazerooni, A.F., Pati, S., et al., **2022**. The university of pennsylvania glioblastoma (UPenn-GBM) cohort: advanced MRI, clinical, genomics, & radiomics. *Scientific data* 9, 453.
2. Bejnordi, B.E., Veta, M., Van Diest, P.J., Van Ginneken, B., Karssemeijer, N., Litjens, G., Van Der Laak, J.A., Hermesen, M., Manson, Q.F., Balkenhol, M., et al., **2017**. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *Jama* 318, 2199–2210.
3. Brixel, R., Bougleux, S., L'ezoray, O., Caillot, Y., Lemoine, B., Fontaine, M., Nebati, D., Renouf, A., **2022**. Whole slide image quality in digital pathology: review and perspectives. *IEEE Access* 10, 131005–131035.
4. Camalan, S., Niazi, M.K.K., Moberly, A.C., Teknos, T., Essig, G., Elmaraghy, C., Taj-Schaal, N., Gurcan, M.N., **2020**. Otomatch: Contentbased eardrum image retrieval using deep learning. *Plos one* 15, e0232776.
5. Chen, J., Mao, H., Woo, W.L., Peng, X., **2023**. Deep multiview clustering by contrasting cluster assignments, in: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 16752–16761.
6. Cisternino, F., Ometto, S., Chatterjee, S., Giacomuzzi, E., Levine, A.P., Glastonbury, C.A., **2024**. Self-supervised learning for characterising histomorphological diversity and spatial rna expression prediction across 23 human tissue types. *Nature Communications* 15, 5906.
7. Deepak, S., Ameer, P., **2020**. Retrieval of brain mri with tumor using contrastive loss based similarity on googlenet encodings. *Computers in biology and medicine* 125, 103993.
8. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A., **2010**. The pascal visual object classes (voc) challenge. *International journal of computer vision* 88, 303–338.
9. Gadermayr, M., Gupta, L., Appel, V., Boor, P., Klinkhammer, B.M., Merhof, D., **2019**. Generative adversarial networks for facilitating stainindependent supervised and unsupervised segmentation: a study on kidney histology. *IEEE transactions on medical imaging* 38, 2293–2302.
10. Gadermayr, M., Gupta, L., Klinkhammer, B., Boor, P., Merhof, D. Unsupervisedly training gans for segmenting digital pathology with automatically generated annotations. *arxiv* **2018**. arXiv preprint arXiv:1805.10059.
11. Gu, F., Burlutskiy, N., Andersson, M., Wil'en, L.K., **2018**. Multi-resolution networks for semantic segmentation in whole slide images, in: *Computational Pathology and Ophthalmic Medical Image Analysis: First International Workshop, COMPAY 2018, and 5th International Workshop, OMIA 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16-20, 2018, Proceedings* 5, Springer, pp. 11–18.

12. Guo, C., Jia, J., Choo, K.K.R., Jie, Y., **2020**. Privacy-preserving image search (ppis): Secure classification and searching using convolutional neural network over large-scale encrypted medical images. *Computers & Security* 99, 102021.
13. Huang, D.-Y., & Wang, C.-H. (2009). Optimal multi-level thresholding using a two-stage Otsu optimization approach. *Pattern Recognition Letters* 30 (2009), 275-284.
14. James, G., Witten, D., Hastie, T., Tibshirani, R., et al., **2013**. An introduction to statistical learning. volume 112. Springer.
15. Javed, S., Mahmood, A., Qaiser, T., Werghi, N., Rajpoot, N., **2024**. Unsupervised mutual transformer learning for multi-gigapixel whole slide image classification. *Medical Image Analysis* 96, 103203.
16. Kanwal, K., Ahmad, K.T., Khan, R., Abbasi, A.T., Li, J., **2020**. Deep learning using symmetry, fast scores, shape-based filtering and spatial mapping integrated with cnn for large scale image retrieval. *Symmetry* 12, 612.
17. Kochetov, B., Bell, P.D., Garcia, P.S., Shalaby, A.S., Raphael, R., Raymond, B., Leibowitz, B.J., Schoedel, K., Brand, R.M., Brand, R.E., et al., **2024**. Unseg: unsupervised segmentation of cells and their nuclei in complex tissue samples. *Communications Biology* 7, 1062.
18. Komura, D., Kawabe, A., Fukuta, K., Sano, K., Umezaki, T., Koda, H., Suzuki, R., Tominaga, K., Ochi, M., Konishi, H., et al., **2022**. Universal encoding of pan-cancer histology by deep texture representations. *Cell Reports* 38.
19. Kumar, N., Verma, R., Sharma, S., Bhargava, S., Vahadane, A., Sethi, A., **2017**. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE transactions on medical imaging* 36, 1550–1560.
20. Kumar, S., Jain, A., Kumar Agarwal, A., Rani, S., Ghimire, A., **2021**. [retracted] object-based image retrieval using the u-net-based neural network. *Computational Intelligence and Neuroscience* 2021, 4395646.
21. Li, Y., Wang, Y., Zhou, H., Wang, H., Jia, G., Zhang, Q., **2022**. Du-net based unsupervised contrastive learning for cancer segmentation in histology images, in: *International Conference on Intelligent Robotics and Applications*, Springer. pp. 201–210.
22. Li, Z., Li, Y., Li, Q., Wang, P., Guo, D., Lu, L., Jin, D., Zhang, Y., Hong, Q., **2023**. Lvit: language meets vision transformer in medical image segmentation. *IEEE transactions on medical imaging*.
23. Lin, H., Chen, H., Dou, Q., Wang, L., Qin, J., Heng, P.A., **2018**. Scannet: A fast and dense scanning framework for metastatic breast cancer detection from whole-slide image, in: *2018 IEEE winter conference on applications of computer vision (WACV)*, IEEE. pp. 539–546.
24. Macenko, M., Niethammer, M., Marron, J.S., Borland, D., Woosley, J.T., Guan, X., Schmitt, C., Thomas, N.E., **2009**. A method for normalizing histology slides for quantitative analysis, in: *2009 IEEE international symposium on biomedical imaging: from nano to macro*, IEEE. pp. 1107–1110.
25. Monowar, M.M., Hamid, M.A., Ohi, A.Q., Alassafi, M.O., Mridha, M.F., **2022**. Autoret: A self-supervised spatial recurrent network for contentbased image retrieval. *Sensors* 22, 2188.
26. Naylor, P., La'e, M., Reyat, F., Walter, T., **2018**. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE transactions on medical imaging* 38, 448–459.
27. van Nierop, W.L., Schneider, J.N., van der Sommen, F., et al., **2024**. Unsupervised cell segmentation by invariant information clustering, in: *Medical Imaging 2024: Image Processing*, SPIE. pp. 195–206.
28. Ozt'urk, S., **2020**. Stacked auto-encoder based tagging with deep features for content-based medical image retrieval. *Expert Systems with Applications* 161, 113693.
29. Peng, J., Wang, Y., **2021**. Medical image segmentation with limited supervision: a review of deep network models. *IEEE Access* 9, 36827–36851.
30. Rehamnia, W., Turlapov, V., **2024**. XAI and weak annotation toward better accuracy and interpretability in the nested classification of brain tumors. URL: <http://dx.doi.org/10.25206/978-5-8149-3873-2-2024-675-689>.
31. Ronneberger, O., Fischer, P., Brox, T., **2015**. U-net: Convolutional networks for biomedical image segmentation, in: *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III* 18, Springer. pp. 234–241.
32. Roy, A., Pramanik, P., Ghosal, S., Valenkova, D., Kaplun, D., Sarkar, R., **2024**. Gru-net: Gaussian attention aided dense skip connection based multiresnet for breast histopathology image segmentation, in: *Annual Conference on Medical Image Understanding and Analysis*, Springer. pp.300–313.
33. Sebai, M., Wang, X., Wang, T., **2020**. Maskmitosis: a deep learning framework for fully supervised, weakly supervised, and unsupervised mitosis detection in histopathology images. *Medical & Biological Engineering & Computing* 58, 1603–1623.
34. Showrav, T.T., Hasan, M.K., **2024**. Hi-gmisnet: generalized medical image segmentation using dwt based multilayer fusion and dual mode attention into high resolution pgan. *Physics in Medicine & Biology* 69, 115019.
35. Singh, P., Hrisheeksha, P., Singh, V.K., **2021**. Cbir-cnn: content-based image retrieval on celebrity data using deep convolution neural network. *Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science)* 14, 257–272.
36. Sokolova, M., Lapalme, G., **2009**. A systematic analysis of performance measures for classification tasks. *Information processing & management* 45, 427–437.
37. Tajbakhsh, N., Jeyaseelan, L., Li, Q., Chiang, J.N., Wu, Z., Ding, X., **2020**. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. *Medical image analysis* 63, 101693.

-
38. Valanarasu, J.M.J., Oza, P., Hacihaliloglu, I., Patel, V.M., **2021**. Medical transformer: Gated axial-attention for medical image segmentation, in: Medical image computing and computer assisted intervention–MICCAI 2021: 24th international conference, Strasbourg, France, September 27–October 1, 2021, proceedings, part I 24, Springer. pp. 36–46.
39. Wang, H., Cao, P., Yang, J., Zaiane, O., **2024a**. Narrowing the semantic gaps in u-net with learnable skip connections: The case of medical image segmentation. *Neural Networks* 178, 106546.
40. Wang, T., Zhang, K., Wang, W., Ma, M., Zhang, Y., Zhao, H., Liu, G., **2024b**. U-hrmlp: Refining segmentation boundaries in histopathology images, in: 2024 IEEE International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1–5.
41. Wei, X., Zhang, Z., Huang, H., Zhou, Y., **2024**. An overview on deep clustering. *Neurocomputing* , 127761.
42. Wright, A.M., Smith, D., Dhurandhar, B., Fairley, T., Scheiber-Pacht, M., Chakraborty, S., Gorman, B.K., Mody, D., Coffey, D.M., **2013**. Digital slide imaging in cervicovaginal cytology: a pilot study. *Archives of pathology and laboratory medicine* 137, 618–624.