

---

УДК 378:004

DOI: 10.25686/978-5-8158-2474-4-2025-816-827

## Обзор наборов данных электронной микроскопии мозга для задач сегментации

**М. Э. Курakin, Н. А. Соколов, А. А. Серебрякова**

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

**Аннотация.** В данной работе представлен систематический обзор публично доступных размеченных наборов данных для электронной микроскопии (ЭМ). Приведен детальный анализ ключевых характеристик существующих датасетов, включая разрешение изображений, плотность аннотаций, баланс классов и согласованность разметки. Результаты исследования демонстрируют существенную неоднородность качества и полноты существующих разметок, что создает значительные сложности для обучения универсальных моделей сегментации. Работа представляет ценность для исследователей в области компьютерного зрения и нейробиологии, предлагая методическую основу для оценки качества ЭМ-данных и выявления наиболее перспективных направлений для создания новых специализированных датасетов.

**Ключевые слова:** электронная микроскопия, сегментация изображений, обзор датасетов, гистограммный анализ, нейробиология, компьютерное зрение, мозг мыши

## Brain electron microscopy dataset review

**M. E. Kurakin, N. A. Sokolov, A. A. Serebryakova**

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

**Abstract.** This paper presents a systematic review of publicly available annotated datasets for electron microscopy (EM). We conduct a detailed analysis of key characteristics of existing datasets, including image resolution, annotation density, class balance, and labeling consistency. The study reveals significant heterogeneity in the quality and completeness of existing annotations, which poses major challenges for training universal segmentation models. The work provides valuable insights for researchers in computer vision and neurobiology by offering a methodological framework for assessing EM data quality and identifying promising directions for developing new specialized datasets.

**Keywords:** electron microscopy, image segmentation, datasets overview, histogram analysis, neurobiology, computer vision, mouse brain

### Введение

Современные методы исследования биологических тканей зачастую содержат алгоритмы машинного обучения, в том числе и глубокие нейронные сети. Однако первые нейросетевые модели, появившиеся в середине XX века, такие как перцептрон Розенблата, предложенный в 1957 году, не получили широкого распространения как из-за нехватки вычислительных мощностей, так и данных для обучения. С появлением специализированных графических адаптеров в начале XXI века нейросетевые технологии стали активно развиваться, но только с появлением большого количества размеченных данных их начинают активно использовать. Так, с появлением AlexNet в 2012 году сверточные нейронные сети начали активно использоваться для обработки изображений, а появившаяся архитектура U-Net [1] в 2015 году стала основой для сегментации медицинских изображений.

Среди биологических тканей особый интерес представляют наборы данных изображений клеток головного мозга мыши. Для такой задачи, как исследование структуры строения и работы мозга, одним из нескольких основных методов получения данных являются электронные микроскопы (ЭМ), а для обработки этих данных часто применяются методы сегментации при помощи сверточных нейронных сетей (СНС). К сожалению, успешность применения методов глубокого обучения снижается нехваткой данных для обучения, поэтому в этой статье будет проведен обзор и анализ доступных датасетов мозга мыши.

### Описание датасетов

Датасеты естественным образом встречаются в работах по архитектурам нейросетей, для аккумуляции датасетов создаются облачные хранилища и сайты с СУБД, а также исследователи собирают в свои таблички и публикуют в статьях, как, например, в статье авторов *Aswath A. и др.* [2].

**Набор данных EPFL**, или Lucchi, или Lucchi++, или ITMM, (рис. 1) стал одним из первых общедоступных датасетов для 3D-сегментации нейронных структур в электронной микроскопии. Он

появился в 2012 году и до сих пор служит эталоном для проверки алгоритмов автоматического анализа митохондрий. Данные, пример которых представлен на рисунке 1 $a$ , были получены с помощью серийной электронной микроскопии (SEM) и включали аннотированные срезы, взятые из области гиппокампа CA1 мозга мыши, что позволило исследователям тестировать методы машинного обучения на реальных биологических структурах.

Lucchi++ остаётся важным историческим бенчмарком и до сих пор используется в образовательных целях и для отладки новых моделей [3-7].

Размер датасета составляет 1065x2048x1536 вокселей, что при размере вокселя (разрешении) 5x5x5 нм составляет 5325x10240x7680 нм.

Размеченная часть датасета состоит из двух непересекающихся областей, названных как test и train с разрешением по 1024x768x165 вокселей. Для EPFL датасета существуют три версии разметки: первоначальная EPFL [8], улучшенная Lucchi++ [9] (пример на рисунке 1 $b$ ) и расширенная ITMM [10]. Первые две реализации разметки содержат только разметку митохондрий, расширенная версия содержит меньшее количество слоев, но разметку на 6 классов: митохондрии, мембранны клеток, постсинаптические уплотнения (PSD), везикулы, аксоны, границы митохондрий. Шестиклассовая разметка всё ещё находится в процессе работы и на момент публикации содержит 60 слоев из набора данных train и 10 из test с разрешением по 1024x768.

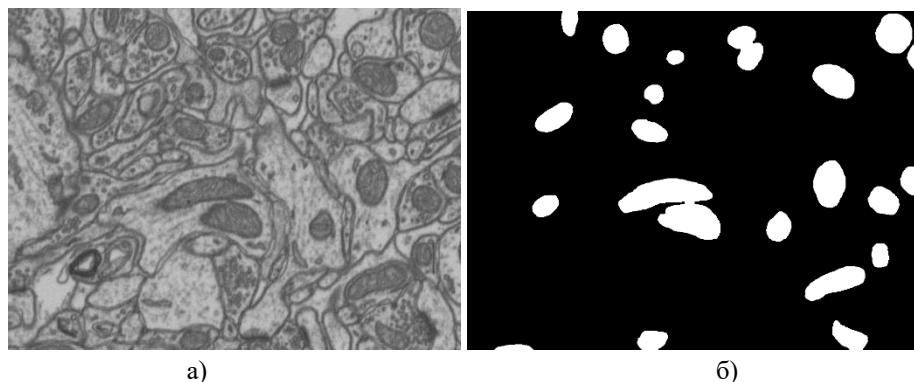


Рисунок 1. Пример данных Lucchi++: а – слой данных; б – разметка митохондрий

**Набор данных Kasthuri** (рис. 2) содержит аннотации митохондрий 3-цилиндрового объема коры мыши [11]. Ткань представляет собой плотный нейропиль млекопитающих из слоев 4 и 5 первичной соматосенсорной коры S1, полученный с помощью серийной секционной электронной микроскопии (ssEM).

Датасет состоит из двух стеков размерами 1463×1613×85 и 1334×1553×75 вокселей. Размер вокселя 3x3x30 нм. Данный датасет соответствует ткани 4389x4839x2550 и 4002x4659x2250 нм соответственно.

Было обнаружено две версии разметки, содержащие маски для митохондрий: первоначальная и улучшенная (см. рис. 2 $b$ ).

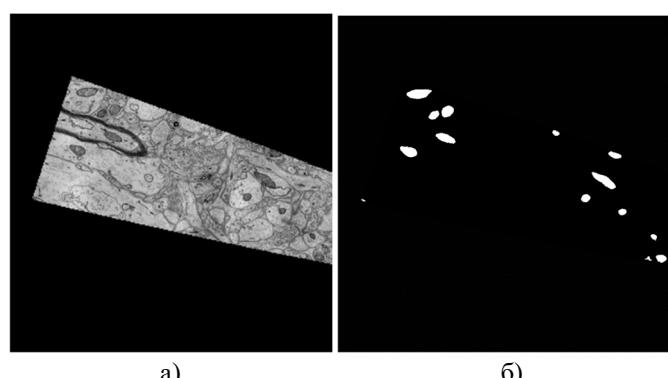


Рисунок 2. Пример данных Kasthuri++: а – слой данных; б – разметка митохондрий

**Набор данных AC4** (рис. 3) был использован в ISBI 2013 EM Challenge [12] в качестве набора для обучения и тестирования. Состоит из стеков, полученных при помощи ЭМ из коры головного мозга мыши (см. рис. 3а). Набор данных состоит из двух стеков, обозначенных как *train* и *test*.

Размер стеков составляет  $1024 \times 1024 \times 100$  вокселей, что при размере вокселя (разрешении)  $6 \times 6 \times 30$  нм соответствует двум экземплярам ткани по  $6144 \times 6144 \times 3000$  нм.

Разметка содержит экземплярную (instance) маркировку нейритов (рис. 3б).

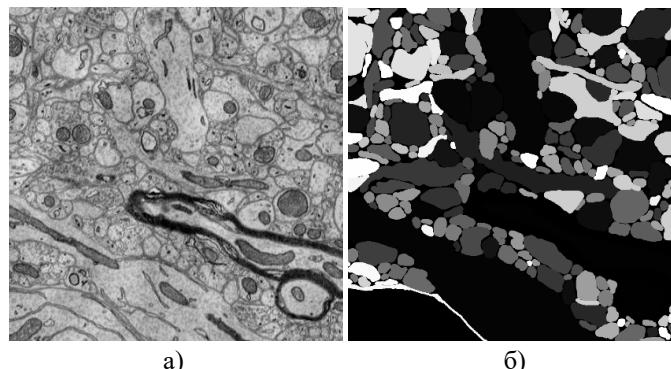


Рисунок 3. Пример данных AC4: а – слой данных; б – разметка клеток после трансформации в полный диапазон 0-255 для визуального восприятия

**Набор данных MitoEM** (рис. 4) представлен в статье [13].

Размер датасета составляет  $512 \times 512 \times 128$  вокселей, что при размере вокселя (разрешении)  $8 \times 8 \times 30$  нм составляет  $4096 \times 4096 \times 3840$  нм.

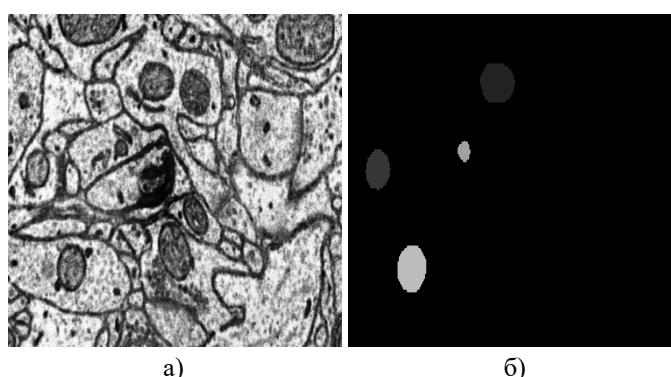


Рисунок 4. Пример данных MitoEM: а – слой данных; б – разметка митохондрий

**Набор данных AxonEM** (рис. 5), представленный в статье [14], состоит из 4 наборов по  $1536 \times 1536 \times 90$  вокселей, что при размере вокселя в  $30 \times 30 \times 30$  нм составляет  $46080 \times 46080 \times 2700$  нм.

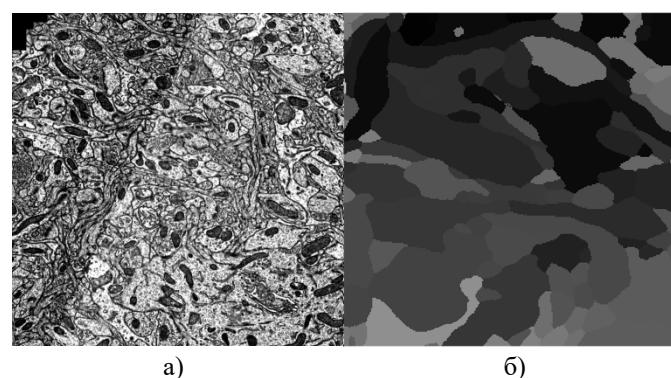


Рисунок 5. Пример данных AxonEM: а – слой данных; б – разметка клеток

**Набор данных CDeep3M-S** (рис. 6) представлен в статье [15].

Размер датасета составляет  $1024 \times 1024 \times 21$  воксель, что при размере вокселя (разрешении)  $2,4 \times 2,4 \times 2,4$  нм составляет  $2458 \times 2458 \times 50$  нм.

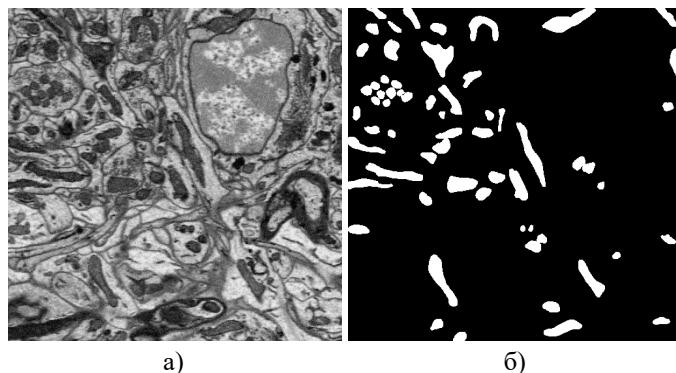


Рисунок 6. Пример данных CDeep3M-S: а – слой данных; б – разметка митохондрий

**Набор данных NucMM** (рис. 7), представленный в статье [16], состоит из 8 размеченных наборов по  $192 \times 192 \times 192$  вокселя, что при разрешении вокселя  $4 \times 4 \times 30$  нм составляет  $768 \times 768 \times 5760$  нм.

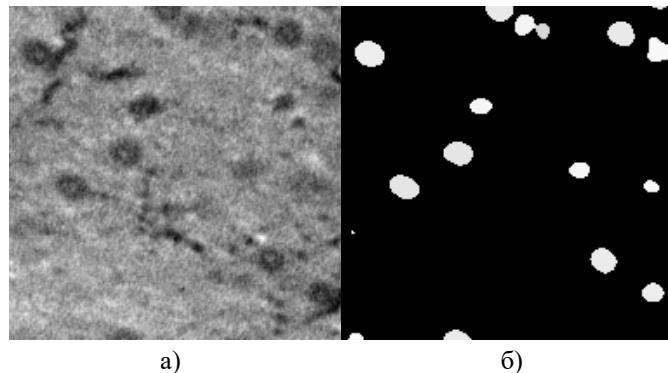


Рисунок 7. Пример данных NucMM: а – слой данных; б – разметка ядер

После фильтрации датасетов с сайта <https://openorganelle.janelia.org/datasets> по снимкам мозга мыши было найдено 16 датасетов: 1-4. Mouse nucleus accumbens; 5-7. Cultured mouse cerebellar granule neurons; 8-9. Mouse dorsal striatum; 10-11. Mouse deep nuclei; 12. Mouse cerebral cortex; 13-15. Mouse hippocampus; 16. Mouse choroid plexus.

**Mouse nucleus accumbens** состоит из ЭМ-снимков ткани прилегающего ядра мыши. Состоит из 4 датасетов.

#### Датасет Mouse nucleus accumbens jrc\_mus-nacc-1

Размер датасета составляет  $2361 \times 2442 \times 563$  вокселя, что при размере вокселя (разрешении)  $4 \times 4 \times 4$  нм составляет  $9444 \times 9768 \times 2252$  нм (рис. 8).

Размеченная часть датасета представляет собой фрагмент исходного слоя объемом  $300 \times 300 \times 300$  вокселей. Разметка состоит из следующих 20 классов, которые можно разбить на три группы: 1) основные клеточные структуры (тело клетки, плазматическая мембрана, внеклеточное пространство); 2) органеллы (митохондрия целиком, внутренний матрикс митохондрии, мембранны митохондрии, рибосомы, ассоциированные с митохондриями, эндоплазматический ретикулум целиком, внутреннее содержимое эндоплазматического ретикулума, мембранны эндоплазматического ретикулума); 3) цитоплазму и включения (цитоплазма, везикулы, содержимое везикул, мембранны везикул, микротрубочки, внутренние участки микротрубочек, внешние участки микротрубочек) (рис. 8б).

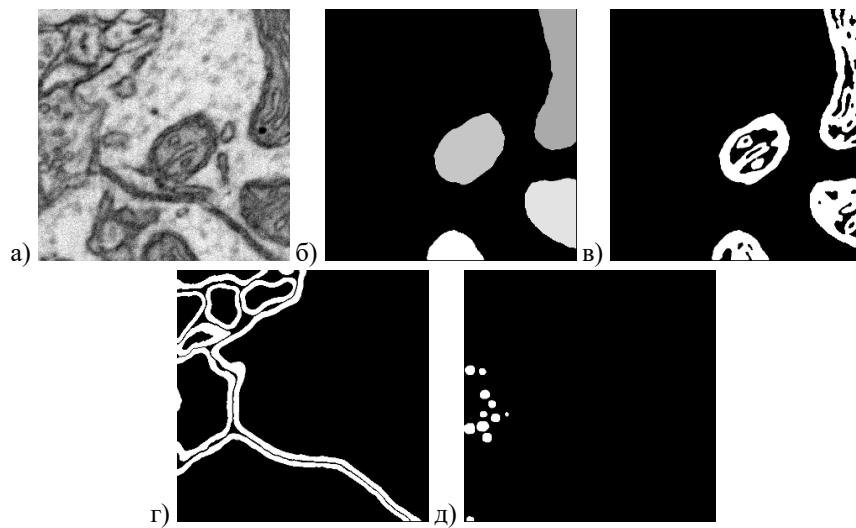


Рисунок 8. Пример данных Mouse nucleus accumbens jrc\_mus-nacc-1: а – слой данных; б – разметки митохондрий; в – разметка границ митохондрий; г – разметка мембран; д – разметка везикул

Исходный датасет имеет объем 2361x2442x563voxеля, в то время как размер предоставленной разметки составляет 600x600x600 voxелей. Первоначальное несоответствие в размерах, особенно тот факт, что глубина разметки превышает глубину исходных данных, объясняется разным разрешением voxелей. Это означает, что размеченный кубический участок не соответствует voxелям оригинала по масштабу. Для корректного совмещения требовалось провести пространственное согласование: уменьшить размер разметки в 2 раза по всем осям, преобразовав её до 300x300x300 voxелей. Поскольку в данных не было координат размеченной области, её пришлось искать вручную. При анализе первого слоя была обнаружена область, визуально похожая на разметку, но соответствующая размеру 300x300x300 voxелей, что и подтвердило необходимость масштабирования и позволило сопоставить разметку с исходным изображением.

#### Датасет Mouse nucleus accumbens jrc\_mus-nacc-2

Размер датасета составляет 2596x2520x564 voxеля, что при размере voxеля (разрешении) 4x4x2,96 нм составляет 10384x10080x1669 нм (рис. 9).

Маркировка содержит набор автоматической разметки (с помощью алгоритма лапласиана Гаусса (LoG)) ядер [17] объемом 324x315x70 voxелей (см. рис. 9 (справа)). Такая же автоматическая разметка применена к датасетам с разметкой ядер (см. ниже).



Рисунок 9. Пример данных Mouse nucleus accumbens jrc\_mus-nacc-2 слоя (слева) и разметки ядра (справа)

#### Датасет Mouse nucleus accumbens jrc\_mus-nacc-3

Размер датасета составляет 5108x5250x732 voxеля, что при размере voxеля (разрешении) 4x4x3,83 нм составляет 20432x21000x2804 нм (рис. 10).

Маркировка содержит набор автоматической разметки ядер [17] объёмом 638x656x91 вокселей.

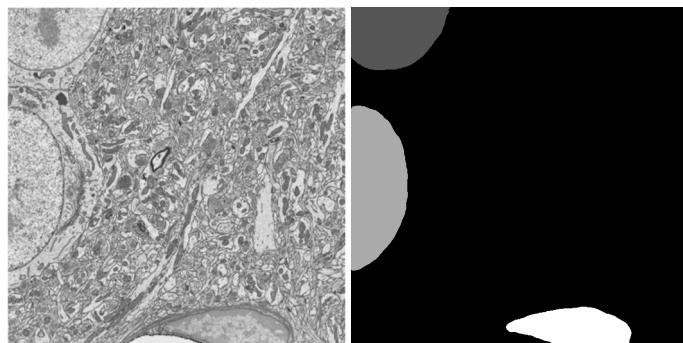


Рисунок 10. Пример данных Mouse nucleus accumbens jrc\_mus-nacc-3 слоя (слева) и разметки ядра (справа)

#### **Датасет Mouse nucleus accumbens jrc\_mus-nacc-4**

Размер датасета составляет 5083x5123x763 вокселя, что при размере вокселя (разрешении) 4x4x3,99 нм составляет 20332x20492x3044 нм (рис. 11).

Маркировка содержит набор автоматической разметки ядер [17] объемом 635x640x95 вокселей.

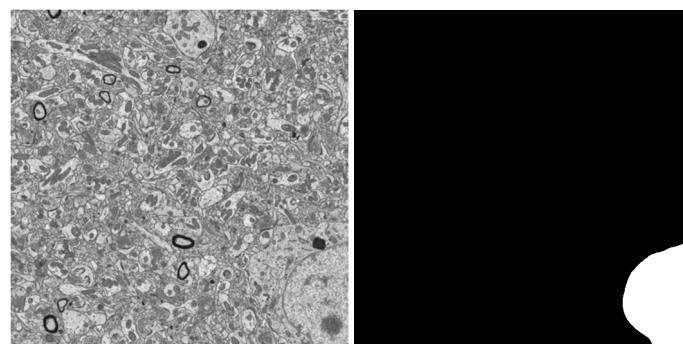


Рисунок 11. Пример данных Mouse nucleus accumbens jrc\_mus-nacc-4 слоя (слева) и разметки ядра (справа)

**Cultured mouse cerebellar granule neurons** состоит из ЭМ-снимков ткани гранулярных нейронов мозжечка мыши, включает 3 датасета.

#### **Cultured mouse cerebellar granule neurons jrc\_mus-granule-neurons-1**

Размер датасета составляет 13000x1360x4350 вокселей, что при размере вокселя (разрешении) 4x4x3,76 нм составляет 52000x5440x16356 нм (рис. 12).

Маркировка содержит набор автоматической разметки ядер [17] объемом 812x85x271 воксель.

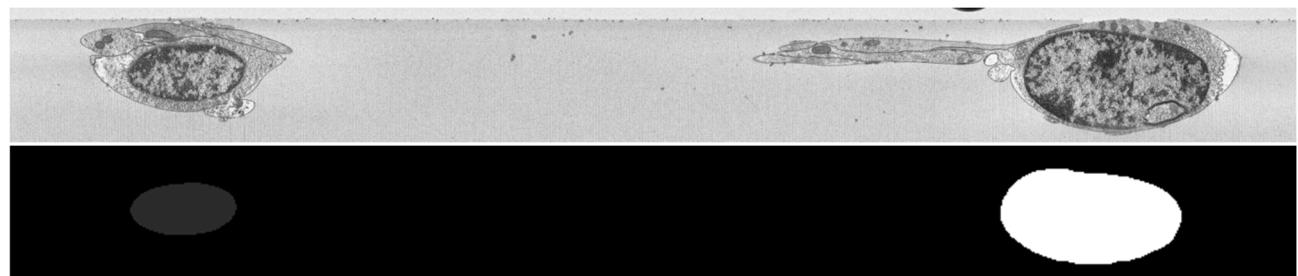


Рисунок 12. Пример данных Cultured mouse cerebellar granule neurons jrc\_mus-granule-neurons-1 слоя (вверху) и разметки ядра (внизу)

### Cultured mouse cerebellar granule neurons jrc\_mus-granule-neurons-2

Размер датасета составляет 2262x537x3946 вокселей, что при размере вокселя (разрешении) 16x16x15,72 нм составляет 36192x8592x62031 нм (рис. 13).

Маркировка содержит набор автоматической разметки ядер [17] объёмом 565x134x986 вокселей.

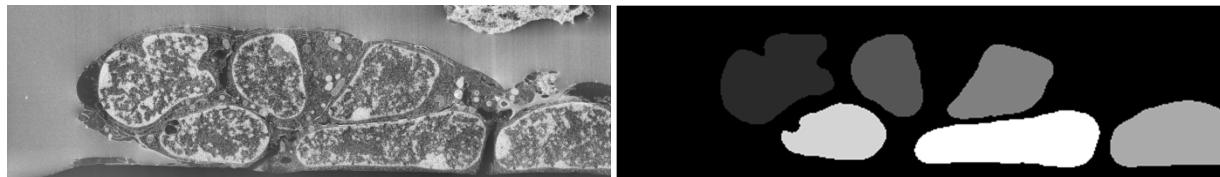


Рисунок 13. Пример данных Cultured mouse cerebellar granule neurons jrc\_mus-granule-neurons-2 слоя (слева) и разметки ядра (справа)

### Cultured mouse cerebellar granule neurons jrc\_mus-granule-neurons-3

Размер датасета составляет 7800x3000x3759 вокселей, что при размере вокселя (разрешении) 4x4x5,37 нм составляет 31200x12000x20186 нм (рис. 14).

Маркировка содержит набор автоматической разметки ядер [17] объемом 487x187x234 вокселя.

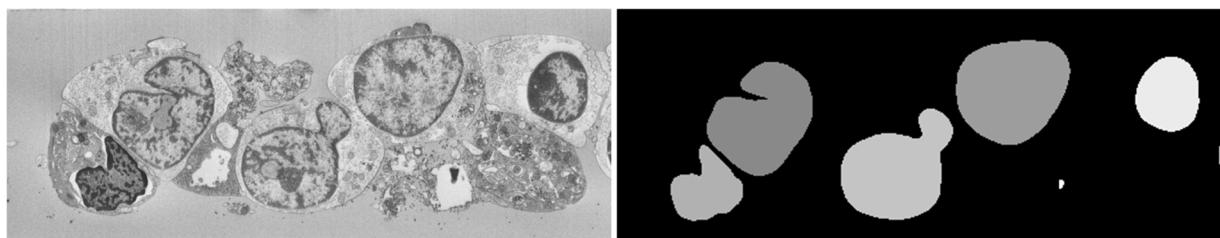


Рисунок 14. Пример данных Cultured mouse cerebellar granule neurons jrc\_mus-granule-neurons-3 слоя (слева) и разметки ядра (справа)

**Mouse dorsal striatum** состоит из снимков ЭМ ткани полосатого тела мыши, включает два датасета.

### Mouse dorsal striatum jrc\_mus-dorsal-striatum

Размер датасета составляет 2554x2564x1353 вокселя, что при размере вокселя (разрешении) 4x4x4,44 нм составляет 10216x10256x6007 нм (рис. 15a). Не содержит разметку.

### Mouse dorsal striatum jrc\_mus-dorsal-striatum-2

Размер датасета составляет 1793x1597x1591 вокселя, что при размере вокселя (разрешении) 4x4x4 нм составляет 7172x6388x6363 нм (рис. 15б). Не содержит разметку.

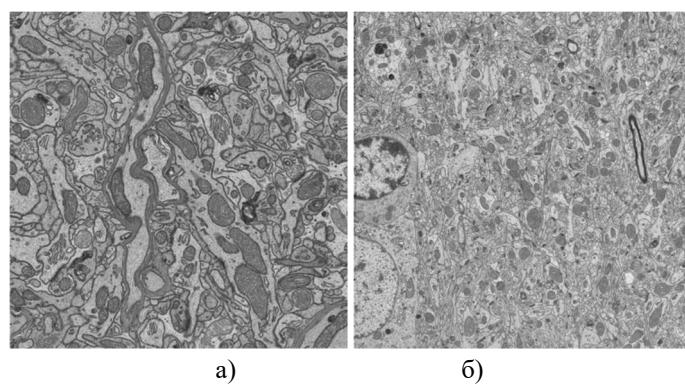


Рисунок 15. Пример данных Mouse dorsal striatum: а – jrc\_mus-dorsal-striatum; б – striatum jrc\_mus-dorsal-striatum-2

**Mouse deep nuclei** состоит из снимков ЭМ-ткани глубоких ядер мозжечка мыши, включает два датасета.

**Mouse deep nuclei jrc\_mus-cerebellum-4**

Размер датасета составляет  $6326 \times 6291 \times 1088$  воксели, что при размере вокселя (разрешении)  $4 \times 4 \times 4$  нм составляет  $25304 \times 25164 \times 4352$  нм (рис. 16a). Не содержит разметку.

**Mouse deep nuclei jrc\_mus-cerebellum-5**

Размер датасета составляет  $1793 \times 1597 \times 1591$  воксель, что при размере вокселя (разрешении)  $4 \times 4 \times 4$  нм составляет  $7172 \times 6388 \times 6363$  нм (рис. 16б). Не содержит разметку.

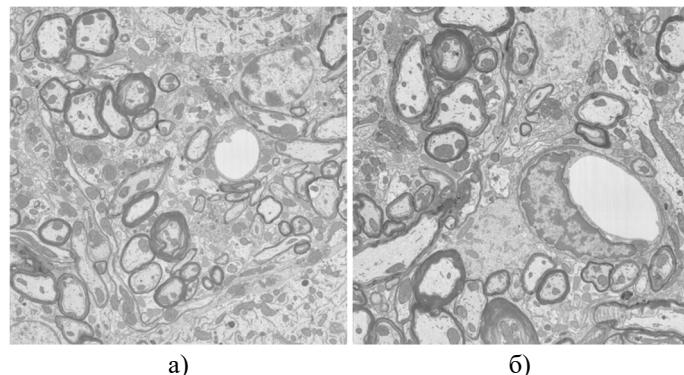


Рисунок 16. Пример данных Mouse deep nuclei: а – jrc\_mus-cerebellum-4; б – jrc\_mus-cerebellum-5

**Mouse cerebral cortex** состоит из снимков ЭМ тканей коры больших полушарий мыши.

Размер датасета составляет  $6351 \times 6395 \times 5910$  воксели, что при размере вокселя (разрешении)  $4 \times 4 \times 4$  нм составляет  $25404 \times 25580 \times 23640$  нм (рис. 17). Не содержит разметку.

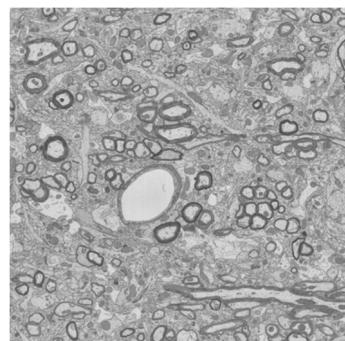


Рисунок 17. Пример данных Mouse cerebral cortex

**Mouse hippocampus** состоит из снимков ЭМ-ткани гиппокампа мыши, включает три датасета.**Mouse hippocampus jrc\_mus-hippocampus-1**

Размер датасета составляет  $17672 \times 17829 \times 24548$  воксели, что при размере вокселя (разрешении)  $8 \times 8 \times 8$  нм составляет  $196384 \times 142632 \times 141376$  нм (рис. 18). Содержит автоматическую разметку ядер [17] (рис. 18, справа).

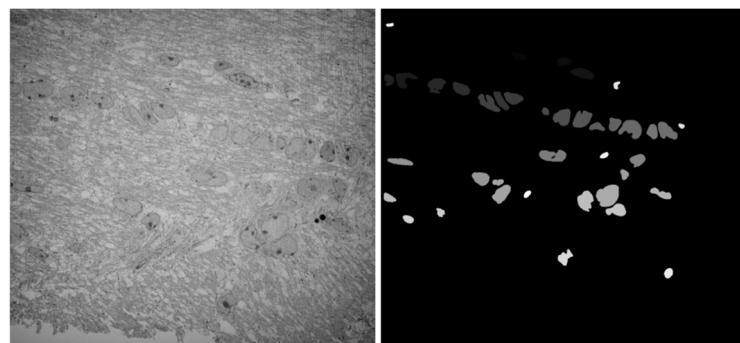


Рисунок 18. Пример данных Mouse hippocampus jrc\_mus-hippocampus-1 (слева) и разметки ядра (справа)

### **Mouse hippocampus jrc\_mus-hippocampus-2**

Размер датасета составляет 5622x5612x6882 voxеля, что при размере voxеля (разрешении) 8x8x8 нм составляет 44976x44896x55056 нм (рис. 19a). Не содержит разметку.

### **Mouse hippocampus jrc\_mus-hippocampus-3**

Размер датасета составляет 7613x10160x5634 voxеля, что при размере voxеля (разрешении) 8x8x8 нм составляет 60904x81280x45072 нм (рис. 19б). Не содержит разметку.

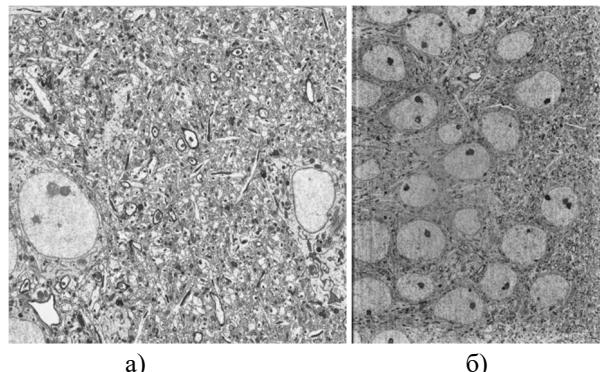


Рисунок 19. Пример данных: а – Mouse hippocampus jrc\_mus-hippocampus-2; б – Mouse hippocampus jrc\_mus-hippocampus-3

### **Mouse choroid plexus** состоит из снимков ЭМ ткани сосудистого сплетения мыши.

Размер датасета составляет 7500x7500x7560 voxелей, что при размере voxеля (разрешении) 8x8x8 нм составляет 60000x60000x60480 нм (рис. 20). Не содержит разметку.

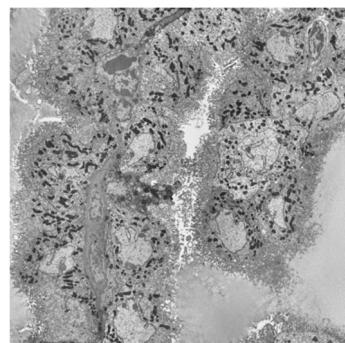


Рисунок 20. Пример данных Mouse choroid plexus

Среди рассмотренных с сайта 16 датасетов изображений тканей мозга мыши ручная разметка доступна только в Mouse nucleus accumbens, охватывающая множество классов, но аннотации покрывают не весь слой, а только определенную область. Остальные датасеты содержат либо автоматическую разметку ядер, либо не имеют разметки. Это ограничивает их применение для задач сегментации клеточных структур.

### **Анализ датасетов**

В рамках данного исследования проведен анализ датасетов, включающий оценку их объема и временной динамики формирования. Построен график роста числа размеченных датасетов и размеченных пикселей по годам (рис. 21). За 12 лет развития области в публичный доступ было выложено 9 размеченных наборов данных, менее одного набора данных за год. Рост количества публично выложенной разметки сильно меньше, чем количество выложенных неразмеченных датасетов и количество предложенных архитектур. Присутствует положительная динамика появления новых датасетов, особенно в последние годы, что свидетельствует об актуальности темы исследования мозга при помощи ЭМ (см. рис. 21). Заметим, что на графике присутствуют только те датасеты, которые смогли успешно скачать. Их несколько меньше, чем в обзоре [2], так как часть ссылок в 2025 году не открылись, поэтому исследователям в каждый момент времени доступны не все описанные датасеты. Несмотря на

появляющиеся время от времени инструменты полуавтоматической разметки, получение качественной разметки так и остается сложной задачей.

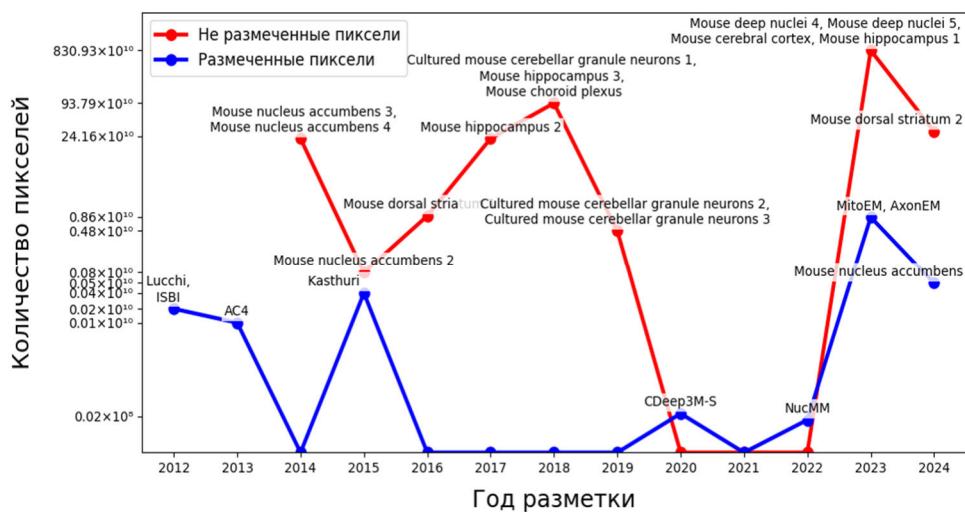


Рисунок 21. График появления размеченных датасетов

В ходе анализа ключевых публичных датасетов для сегментации компартментов мозга на изображениях, полученных с электронного микроскопа, были построены диаграммы распределения пикселей органелл (рис. 22).

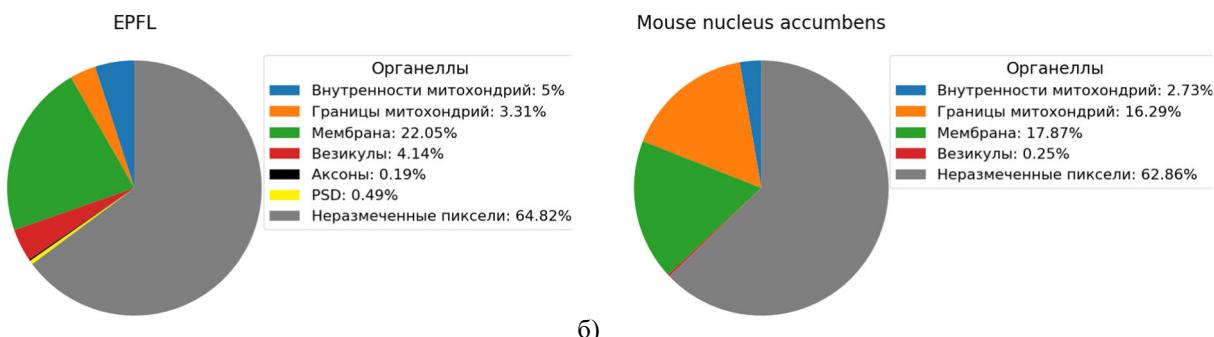


Рисунок 22. Диаграммы, построенные по наборам данных: а – EPFL; б – Mouse nucleus accumbens

Важным аспектом анализа данных является оценка баланса классов. Как видно из представленных диаграмм (рис. 22), для всех проанализированных датасетов наблюдается значительный дисбаланс в распределении пикселей по классам. Например, класс митохондрий на несколько порядков превышает по объему такие классы, как аксоны или PSD. Анализ общего количества размеченных пикселей (табл. 1) также показывает значительную несбалансированность классов.

Таблица 1. Общее количество размеченных пикселей для классов

Класс	Количество размеченных пикселей	Датасеты
Митохондрии	186 932 838	Lucchi, Kasthuri, mnac, CDeep3M-S
Границы митохондрий	5 958 553	ITMM, mnac
Мембранны	15 227 809	ITMM, mnac
Везикулы	2 021 510	ITMM, mnac
Аксоны	88 467	ITMM
PSD	229 141	ITMM
Нейтриты	84 322 554	AC4
Клетки	1 917 180	NucMM

Кроме баланса, ещё одной из ключевых характеристик класса в наборе данных является его цвет (или значения его интенсивности на изображении). Для анализа этой характеристики построены гистограммы датасетов, содержащих несколько классов (рис. 23).

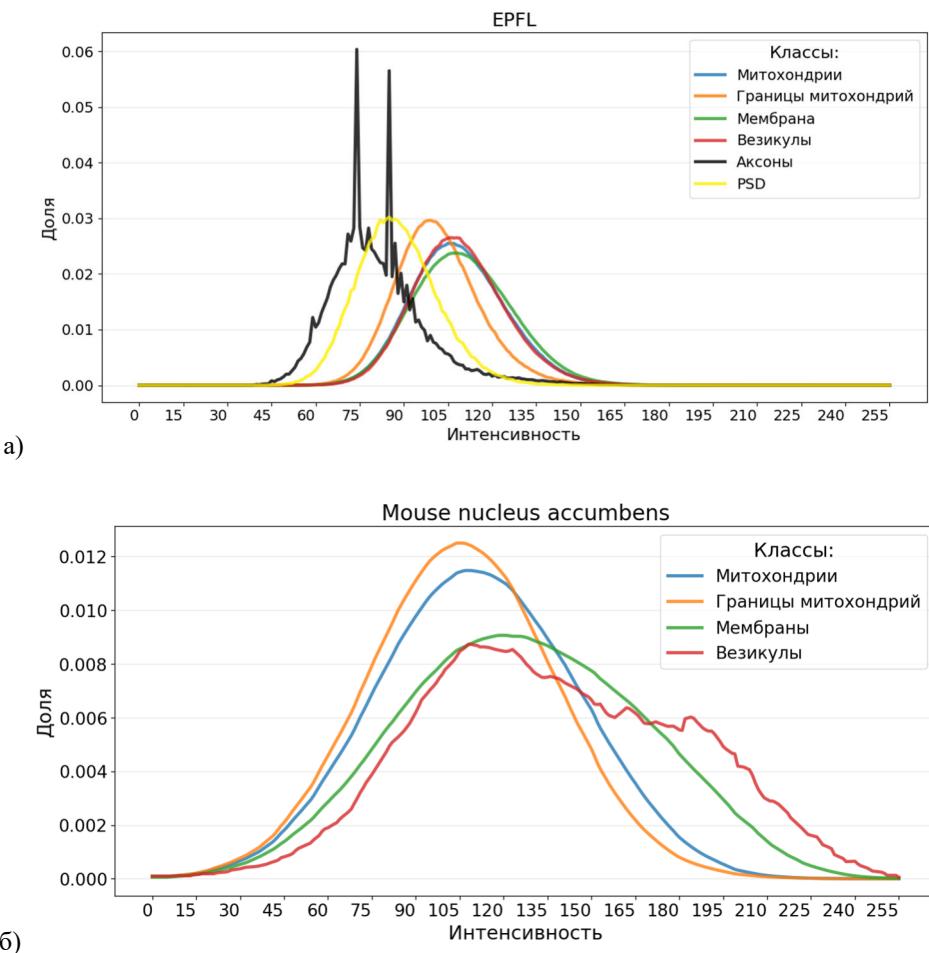


Рисунок 23. Гистограммы, построенные по наборам данных: а – EPFL; б – Mouse nucleus accumbens

Полученные графики имеют визуальное сходство между различными структурами, которое сохраняется между различными датасетами, что не позволяет классифицировать данные по значениям интенсивности. Это указывает на то, что такое распределение является общей характеристикой ЭМ-изображений рассматриваемых биологических тканей и не является существенной характеристикой. Пилообразность некоторых графиков связана с малой представленностью класса в наборе данных.

Схожесть гистограмм одного класса на разных наборах данных мозга мыши показывает наличие единого подхода для получения данных при помощи ЭМ, что дает возможность создать универсальный алгоритм как для сегментации, так и для генерации данных.

### Заключение

Данные – это фундамент, на котором строится глубокое обучение, это стратегический актив, определяющий возможности и ограничения искусственного интеллекта. Ценность данных определяется их объемом, репрезентативностью, чистотой и балансом.

Проведенный анализ публичных наборов данных по электронной микроскопии однозначно подтверждает проблему данной работы: основным препятствием для эффективного и широкого применения нейронных сетей в автоматической сегментации электронно-микроскопических изображений мозга является острая нехватка качественных, объемных и разнообразных размеченных датасетов. Несмотря на значительный прогресс в разработке специализированных архитектур, методов снижения зависимости от разметки и инструментов для работы с данными, все эти решения в конечном

итоге упираются в проблему данных. Трудоемкость и медлительность ручной экспертной разметки, особенно для многоклассовой сегментации, дисбаланс классов, сложность создания реалистичных синтетических данных и отсутствие детально размеченных датасетов для многих структур мозга остаются ключевыми барьерами. Преодоление этой зависимости от данных является критически важной задачей для дальнейшего развития надежных, точных и универсально применимых методов автоматической сегментации в нейробиологии.

### **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 000000Ц313925Р3Х0002).

### **Список литературы**

1. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015). 2015: 234-241.
2. Aswath A., et al. Segmentation in large-scale cellular electron microscopy with deep learning: A literature survey. Medical image analysis, 2023, 89, 102920.
3. Marcén A. C., Iglesias A., Lapeña R., Pérez F., Cetina C. A Systematic Literature Review of Model-Driven Engineering Using Machine Learning, in IEEE Transactions on Software Engineering, Sept. 2024, 50(9), 2269-2293, doi: 10.1109/TSE.2024.3430514.
4. Kazimi B., Ruzaeva K., Sandfeld S. Self-Supervised Learning with Generative Adversarial Networks for Electron Microscopy, 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle, WA, USA, 2024, pp. 71-81, doi: 10.1109/CVPRW63382.2024.00012.
5. Ali M., Benfante V., Basirinia G., Alongi P., Sperandeo A., Quattrochi A., Giannone A.G., Cabibi D., Yezzi A., Di Raimondo D., et al. Applications of Artificial Intelligence, Deep Learning, and Machine Learning to Support the Analysis of Microscopic Images of Cells and Tissues. J. Imaging, 2025, 11, 59. <https://doi.org/10.3390/jimaging11020059>
6. Lun Che, Zhongping He, Kaiyuan Zheng, Tianyu Si, Meiling Ge, Hong Cheng, Lingrong Zeng. Deep learning in alloy material microstructures: Application and prospects, Materials Today Communications, 2023, 37, 107531, ISSN 2352-4928, <https://doi.org/10.1016/j.mtcomm.2023.107531>.
7. Azhagu Madhavan Sivalingam, Darshitha D. Sureshkumar, Vijayalakshmi Pandurangan. Cerebellar pathology in forensic and clinical neuroscience, Ageing Research Reviews, 2025, 106, 102697, ISSN 1568-1637, <https://doi.org/10.1016/j.arr.2025.102697>.
8. Lucchi A., Smith K., Achanta R., Knott G., Fua P. Electron microscopy dataset. 2023. <https://www.epfl.ch/labs/cvlab/data/data-em/>.
9. Casser V., Kang K., Pfister H., Haehn D. Fast mitochondria detection for connectomics, Proceedings of Machine Learning Research, 2020, 1, 10.
10. Getmanskaya A.A., Sokolov N.A. Turlapov V.E. Multiclass U-Net Segmentation of Brain Electron Microscopy Data Using Original and Semi-Synthetic Training Datasets. Program Comput Soft, 2022, 48, 164–171.
11. Kasthuri N., Hayworth K. J., Berger D. R., et al. Saturated reconstruction of a volume of neocortex, Cell, 2015, 162(3), 648–661.
12. Arganda-Carreras I., Turaga S. C., Berger D. R., et al. Crowdsourcing the creation of image segmentation algorithms for connectomics. Frontiers in Neuroanatomy, 2015, 9, 142.
13. Wei Donglai, Lin Zudi, Franco-Barranco Daniel, et al. MitoEM Dataset: Large-Scale 3D Mitochondria Instance Segmentation from EM Images. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2020: 66-76.
14. Wei D. et al. AxonEM Dataset: 3D Axon Instance Segmentation of Brain Cortical Regions. Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. MICCAI 2021. Lecture Notes in Computer Science, 12901. Springer, Cham, 2021. [https://doi.org/10.1007/978-3-030-87193-2\\_17](https://doi.org/10.1007/978-3-030-87193-2_17)
15. Haberl M.G., Churas C., Tindall L., et al. CDeep3M—Plug-and-Play cloud-based deep learning for image segmentation. Nature Methods, 2018, 15 (9), 677-680.
16. Lin Z., Wei D., Petkova M.D., et al. NucMM dataset: 3d neuronal nuclei instance segmentation at sub-cubic millimeter scale. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2021, 164-174.
17. Xu C.S., Pang S., Shtengel G. et al. An open-access volume electron microscopy atlas of whole cells and tissues. Nature, 2021, 599, 147–151. <https://doi.org/10.1038/s41586-021-03992-4>