

Научная статья
УДК 004.891.3
doi: 10.17586/2713-1874-2023-3-58-65

ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ КЛИНИЧЕСКИХ РЕШЕНИЙ ПРИ ДИАГНОСТИКЕ РЕДКИХ ЗАБОЛЕВАНИЙ

Екатерина Александровна Жабровец^{1,3}, Анастасия Сергеевна Лукина^{2,3}

¹ООО «Тензор», Санкт-Петербург, Россия, gvenny99@gmail.com, <https://orcid.org/0000-0001-5504-165X>

²АО «Вертекс», Санкт-Петербург, Россия, anastasiyalukina@inbox.ru, <https://orcid.org/0009-0000-1503-8675>

³Университет ИТМО, Санкт-Петербург, Россия

Язык статьи – русский

Аннотация: Появление программных решений, способных аккумулировать и обрабатывать информацию, извлеченную из историй болезней пациентов с редко встречающимися заболеваниями, способствует повышению уровня осведомленности врачей, помогает избежать клинических ошибок и оптимизировать организацию процесса лечения больных. Настоящее исследование посвящено разработке информационного обеспечения клинических решений при диагностике редких заболеваний на примере опухолей шеи доброкачественного характера, локализующихся в области сонной артерии (каротидных хемодектом). Обоснован состав компонентов и определены требования к системе информационного обеспечения диагностики редких заболеваний, спроектирована база данных системы, разработана и опубликована версия библиотеки для работы с радиологическими изображениями опухоли шеи и датасетом, представляющая средства для анализа пользовательских данных о пациентах и классификации. Набор данных пациентов с каротидной опухолью и схожими заболеваниями сформирован и интегрирован в библиотеку (возможно его обновление за счет добавления информации о новых пациентах). Реализованы такие операции обработки радиологических изображений, как применение стандартных преобразований (Hounsfield Scale, Erosion, Dilation) и rescaling. Итоговая программная библиотека может быть использована IT-исследователями в области работы с медицинскими данными, в частности, для заимствования и самостоятельной доработки методов извлечения медицинской информации и распознавания радиологических изображений опухоли шеи.

Ключевые слова: база данных, каротидная параганглиома, опухоль каротидного тела, радиологические изображения, хемодектома

Исследование выполнено при финансовой поддержке Университета ИТМО НИР № 622274 «Разработка технологии диагностики опухолей шеи».

Ссылка для цитирования: Жабровец Е.А., Лукина А.С. Информационное обеспечение клинических решений при диагностике редких заболеваний // Экономика. Право. Инновации. 2023. № 3. С. 58–65. <http://dx.doi.org/10.17586/2713-1874-2023-3-58-65>.

INFORMATION SUPPORT OF CLINICAL DECISIONS IN THE DIAGNOSIS OF RARE DISEASES

Ekaterina A. Zhabrovets^{1,3}, Anastasia S. Lukina^{2,3}

¹Tensor LLC», Saint Petersburg, Russia, gvenny99@gmail.com, <https://orcid.org/0000-0001-5504-165X>

²Vertex JSC, Saint Petersburg, Russia, anastasiyalukina@inbox.ru, <https://orcid.org/0009-0000-1503-8675>

³ITMO University, Saint Petersburg, Russia

Article in Russian

Abstract: The emergence of software solutions capable of accumulating and processing information extracted from the case histories of patients with rare diseases helps to increase the level of physicians' awareness, helps to avoid clinical errors and optimize the organization of the patient treatment process. The present study is devoted to the development of information support for clinical solutions in the diagnosis of rare diseases on the example of benign neck tumors localized in the carotid artery (carotid chemodectomas). The composition of the components is substantiated and the requirements for the information support system for diagnosing rare diseases are determined, the database of the system is designed, a version of the library for working with radiological images of a neck tumor and a dataset is developed and published, representing tools for analyzing user data about patients and classification. A data set of patients with carotid tumor and similar diseases has been formed and integrated into the library (it can be updated by adding information about new patients). Implemented such radiological image processing operations as applying standard transformations (Hounsfield Scale, Erosion, Dilation) and rescaling. The resulting software library can be used by IT researchers in the field of working

with medical data, in particular, to borrow and independently refine methods for extracting medical information and recognizing radiological images of a neck tumor.

Keywords: carotid body tumor, carotid paraganglioma, chemodectoma, database, radiological imaging

The study was financially supported by ITMO University, Research Work No. 622274 «Development of Technology for Diagnosing Neck Tumors».

For citation: Zhabrovets E.A., Lukina A.S. Information Support of Clinical Decisions in the Diagnosis of Rare Diseases. *Ekonomika. Pravo. Innovacii*. 2023. No. 3. pp. 58–65. (In Russ.). <http://dx.doi.org/10.17586/2713-1874-2023-3-58-65>.

Введение. Появление программных решений, способных аккумулировать и обрабатывать информацию, извлеченную из историй болезней пациентов с редко встречающимися заболеваниями, способствует повышению уровня осведомленности врачей, помогает избежать клинических ошибок и оптимизировать организацию процесса лечения больных.

Настоящее исследование посвящено разработке информационного обеспечения клинических решений при диагностике редких заболеваний на примере опухолей шеи доброкачественного характера, локализующихся в области сонной артерии (каротидных хемодектом).

Концептуальное описание каротидных хемодектом (краткий обзор). Каротидная хемодектома, известная также как параганглиома, гломусная опухоль, опухоль каротидного тельца, является доброкачественной опухолью [1–4]. Доля опухоли каротидного гломуса в общем количестве опухолей шеи и головы составляет 0,5–0,6% [2, 5]. Она более чем в два раза чаще встречается у женщин, чем у мужчин. Самым типичным местом расположения хемодектомы является каротидный синус. Он представляет собой место, где происходит расширение сонной артерии перед ее непосредственным разделением. Проявления болезни обычно происходят в возрасте от 20 до 50 лет [2, 5]. Также существуют свидетельства, что хемодектома имеет связь с некоторыми заболеваниями, передающимися по наследству. Исследователи отмечают возможную связь с мутацией генов сукцинатдегидрогеназы (SDHB, SDHC, SDHD), которая может быть передана от отца к детям [6].

По мере роста хемодектомы человек начинает испытывать различного рода дискомфорт. В зависимости от конкретных случаев больные могут страдать от нарушения глотательной функции, головных болей,

резких повышений артериального давления, частых головокружений, болей в затылке, бессонницы. В силу особенностей расположения, параганглиома давит на нервные структуры, у человека могут развиваться такие признаки, как сухой кашель, осиплость голоса, тошнота [1–5].

Своевременное диагностирование каротидной хемодектомы шеи является довольно трудной задачей для специалистов. На данном этапе развития медицинской науки в обнаружении каротидных хемодектом ошибаются в 20–90% случаев [2, 5]. Причиной этому служит уже упомянутая крайняя редкость заболевания. Следствием ошибочной или поздней диагностики являются ошибки лечения и организации медицинской помощи пациентам. Таким образом, разработка информационного обеспечения клинических решений при диагностике и лечении редких заболеваний, в частности, каротидных хемодектомах, является актуальной. Наиболее важными компонентами информационного обеспечения являются: база медицинских данных о редких заболеваниях, возможность обработки изображений, а также подключение к международным системам хранения медицинских данных.

Материалы и методы исследования. Исследовано основано на анализе международных информационных систем поддержки клинических решений и известных аналогов компонентов проектируемого информационного обеспечения.

Основная часть.

Обзор существующих систем хранения медицинских данных. Система PubMed является бесплатным ресурсом, который поддерживает поиск и извлечение биомедицинских данных и медицинской литературы [7]. База данных PubMed содержит более 35 миллионов цитирований и текстов медицинской литературы. Содержащаяся информация не

включает в себя полные журнальные статьи, тем не менее ссылки на последние при наличии доступа всегда приводятся. Ресурс разработан по инициативе Национального центра биотехнологической информации Соединенных Штатов Америки. Он поддерживает поиск по трем базам знаний:

1) MEDLINE – хранилище журнальных выписок и метаданных генетических и химических исследований.

2) PubMed Central (PMC) – второй по величине компонент ресурса, содержащий полные тексты статей из журналов, отобранные Национальной библиотекой медицины США для архивации (включает как новые, так и исторические материалы), а также статьи индивидуальных исследователей.

3) Bookshelf – архивы книг, отчетов, баз данных и прочих документов, связанных с науками о жизни.

Библиотечный ресурс Cochrane Library является одним из продуктов некоммерческой организацией Cochrane, созданной в Великобритании с целью повышения уровня доступности медицинских знаний для всех людей, заинтересованных в процессе принятия решений в области здравоохранения [8]. Он представляет собой коллекцию следующих баз данных:

1) Cochrane Database of Systematic Reviews – результаты рецензированных систематических исследований и протоколов.

2) Cochrane Clinical Answers – информационные заметки о собственных исследованиях сообщества. Каждая такая заметка содержит какой-либо клинический вопрос, ответ на него, а также те или иные связанные данные по клиническому случаю, которые потенциально могут быть интересны врачам (это могут быть описания, табличные данные или графики).

Система UpToDate является платным ресурсом для медиков, предоставляющим информацию о результатах реальных исследований [9]. База данных проекта содержит около десяти тысяч рецензированных статей врачей со всего мира.

Поисковая система National Library of Medicine по запросу «chemodectoma» (<https://openi.nlm.nih.gov/gridquery?q=chemodectoma&it=xg&m=401&n=500>) выдает набор

из 1570 разрозненных изображений из статей соответствующей тематики.

Таким образом, на основных информационных медицинских ресурсах, обладающих довольно обширными информационными медицинскими базами, не удалось найти датасеты по хемодектому, что подтверждает актуальность создания отдельной тематической базы в рамках предметной области.

Обзор существующих библиотек для работы с радиологическими изображениями. Для реализации функций обработки радиологических изображений хемодектом целесообразно проанализировать возможности современных библиотек для работы с радиологическими изображениями.

TorchIO – open-source библиотека с открытым исходным кодом, предназначенная для загрузки, обработки, преобразования медицинских изображений и их сэмпирования в форме батчей, в основе которой лежат решения библиотеки глубокого обучения Pytorch [10]. Библиотека включает в себя множество преобразований интенсивности изображения и пространственных преобразований. Эти преобразования представляют собой как типичные операции компьютерного зрения, так и специфические для данной области – моделирование артефактов интенсивности, которое могло возникнуть из-за неоднородности (смещения) магнитного поля МРТ.

В рамках библиотеки также реализован графический интерфейс, в котором можно делать все преобразования над трехмерным изображением вручную. В рассматриваемый программный пакет интегрировано несколько датасетов с медицинскими изображениями, таких как:

1) Information eXtraction from Images (IXI) – содержит около 600 различных изображений для здоровых людей.

2) EPISURG – клинический датасет с нормализованными изображениями пациентов с эпилепсией, перенесших резекцию на мозге.

3) RSNAMICCAI – датасет соревнования Kaggle по классификации опухолей мозга.

Существенный минус данной библиотеки заключается в том, что она работает с архивами изображений, то есть напрямую работать с файлами dicom не позволяет. Другим недостатком с точки зрения

исследования является отсутствие датасета с опухолями шеи.

Библиотека MedPy – это открытая Python-библиотека, предоставляющая базовые функции для чтения, записи и прочего рода манипуляций с медицинскими изображениями [11]. В рамках доступного функционала присутствует большое число фильтров для преобразований, такие как:

1) Фильтр Отсу – используется для нахождения оптимального значения для разделения изображения на фон и не фон.

2) Фильтр среднего – усредняет значения яркостей для крайних пикселей для устранения шума.

3) Преобразование Хафа – преобразование, использующееся для выделения геометрических объектов на изображении.

4) Стандартизатор значений интенсивности (применяется операция стандартизации для выравнивания шкалы яркостей пикселей).

5) Анизотропная диффузия – разновидность сглаживающего фильтра.

Библиотека располагает довольно большим количеством метрик, от манхэттенского расстояния до расстояния Минковского.

Библиотека Pydicom – это библиотека, позволяющая манипулировать с файлами формата dicom, а также их метаданными. Можно как прочитать заданный файл, так и создать свое собственное изображение [12]. Из метаданных может быть также извлечены массивы яркостей пикселей для непосредственной работы с графическим изображением.

Результаты исследования. Проектирование системы информационного обеспечения клинических решений при диагностике редких заболеваний. Для системы информационного обеспечения клинических решений при диагностике редких заболеваний были сформулированы следующие требования:

1) Возможность регистрации и авторизации для врача.

2) Просмотр данных пациентов.

3) Возможность актуализации информации.

4) Возможность добавления новых данных пользователя.

В рамках проектирования была составлена схема базы данных. Она включает в себя двенадцать таблиц:

1) «Операции» – содержит перечень всех уникальных операций, проводимых в стационаре.

2) «Лекарство» – содержит записи обо всех доступных в рамках стационара лекарствах.

3) «Исследование» – содержит информацию о лабораторных исследованиях, проводимых в стационаре.

4) «Операции в рамках ИБ» – хранит записи о проведенных конкретному пациенту операциях.

5) «Медикаментозное лечение» – хранит информацию о приемах лекарств пациентом.

6) «Исследование в рамках ИБ» – содержит записи о лабораторных исследованиях конкретного пациента, включая радиологические изображения.

7) «История болезни» – хранит истории болезни пациентов.

8) «Заболевание» – представляет собой перечень уникальных заболеваний.

9) «Отделение» – сведения о структурном подразделении стационара.

10) «Пациент» – сведения о пациенте.

11) «Лечебное учреждение» – основная информация о стационаре.

12) «Пользователь» – информация о пользователе системы.

На рисунке 1 представлена схема базы данных проектируемой системы.

По результатам выполнения данного этапа работы был разработан прототип системы хранения данных пациентов с доброкачественными опухолями шеи, который в дальнейшем будет использоваться в качестве справочного ресурса, в частности, разработаны интерфейсы регистрации в системе, форма входа на сайт, вкладка с пациентами, окно редактирования информации. Расширение системы предусматривает возможность обращения к существующим системам хранения медицинских данных, а также к библиотеке для работы с радиологическими изображениями.

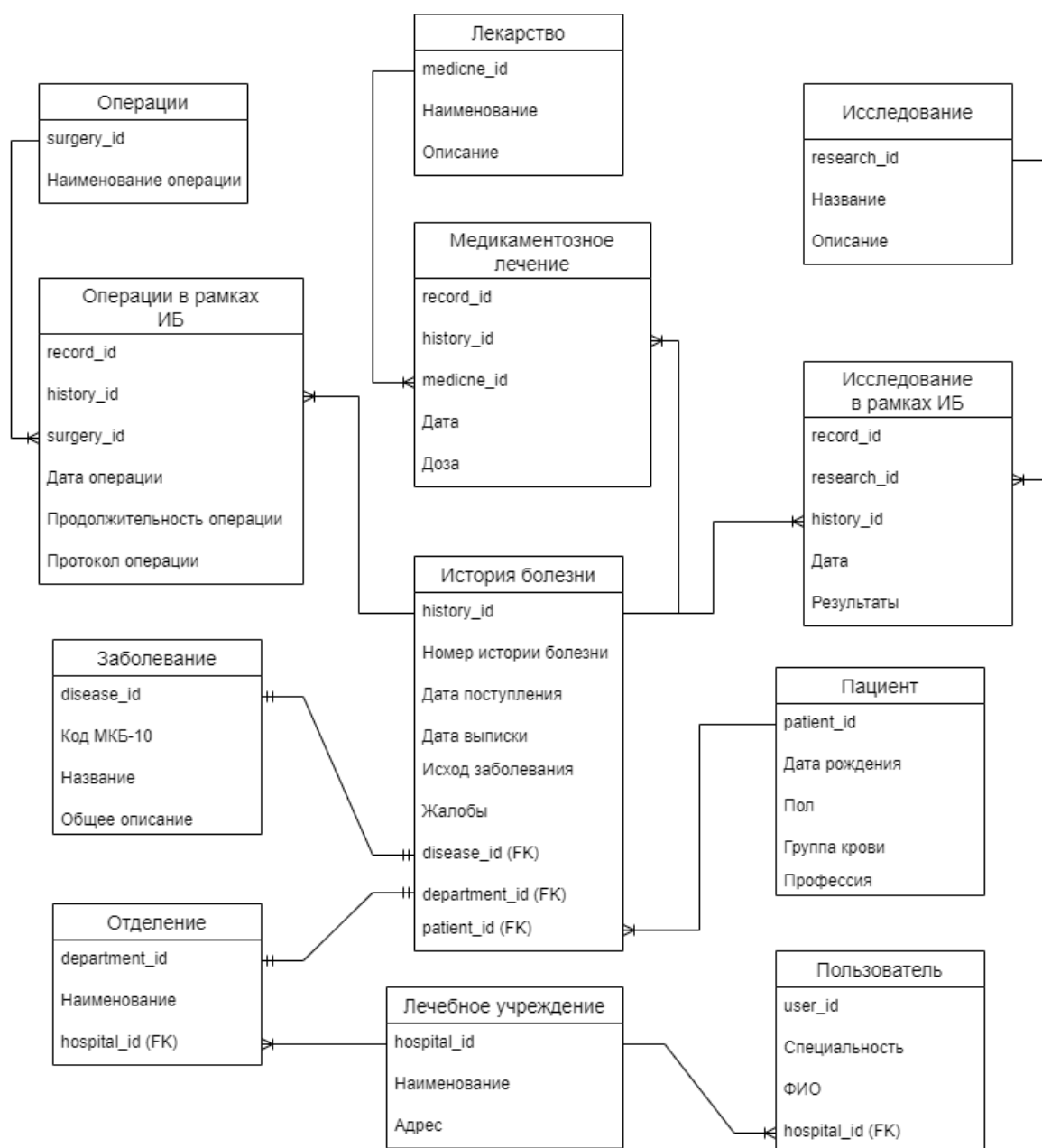


Рисунок 1 – Схема базы данных системы

Разработка версии библиотеки для работы с радиологическими изображениями.

Чаще всего изображения опухолей шеи, в частности каротидных хемодектом, представляют собой файлы формата dicom, поэтому было решено сделать упор, в первую очередь, именно на работу с данным типом.

Сформулированы следующие требования к конечному функционалу:

- 1) Чтение файлов из директории с преобразованием в многомерные массивы (как в виде единичных изображений, так и сэмплов).
- 2) Формирование набора данных.

3) Применение стандартных преобразований (Hounsfield Scale, Erosion, Dilation).

4) Rescaling изображения в соответствии с интересами исследователя (выделение фрагментов того или иного «вещества» при помощи Hounsfield Scale).

5) Преобразование изображения с применением радиологических окон.

6) Выделение областей по заданным маскам/ при помощи вычисляемых масок (с применением методов кластеризации).

В рамках данной работы реализованы пункты 1–4. Структура основного функционала библиотеки представлена на рисунке 2.

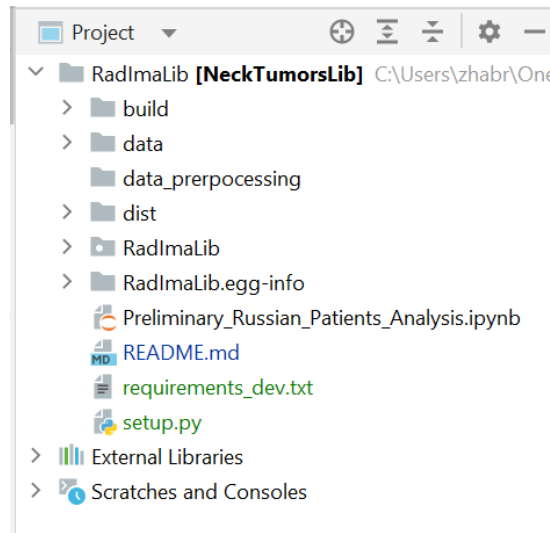


Рисунок 2 – Структура проекта библиотеки RadImaLib

Внутри папки RadImaLib находится файл `utils.py`, который содержит основные функции для работы с изображениями, в том числе:

- добавления расширения `.dcm` для файлов в директории с МРТ, загруженными с диска;

- нормализации к шкале Хаунсфильда;
- чтения `dicom`-файлов из директории.

Файл `requirements-dev` содержит перечень необходимых библиотек, используемых в рамках проекта для автоматической сопутствующей установки при загрузке библиотеки.

В папке `data` лежат датасеты, доступные для выгрузки через функции.

Публикация библиотеки на TestPyPI. В корневой директории проекта создан файл `setup.py`. Он содержит информацию для публикации библиотеки в открытый доступ (наименование, перечень сторонних необходимых пакетов, описание и так далее). Все

библиотеки программного языка Python сохраняются в онлайн-каталог Python Package Index (PyPI). PyPI – это ресурс, в котором происходит непосредственная публикация разрабатываемых библиотек с последующей возможностью их скачивания конечными пользователями [13]. Существует тестовая версия данного сайта, имеющая название TestPyPI. В ней можно отслеживать, как будет выглядеть библиотека после релиза, а также как она будет функционировать при скачивании. Перед началом работы на TestPyPI необходимо зарегистрировать своего пользователя. Далее создаются скрытые папки `dist`, `build` и `RadImaLib.egg-info`. Внутри `dist` находятся файлы, отвечающие за архивацию основного содержимого программного пакета.

Была произведена непосредственно публикация при помощи библиотеки `twine` [14]. После деплоя на странице пользователя отображается уже готовый проект, как представлено на рисунке 3.

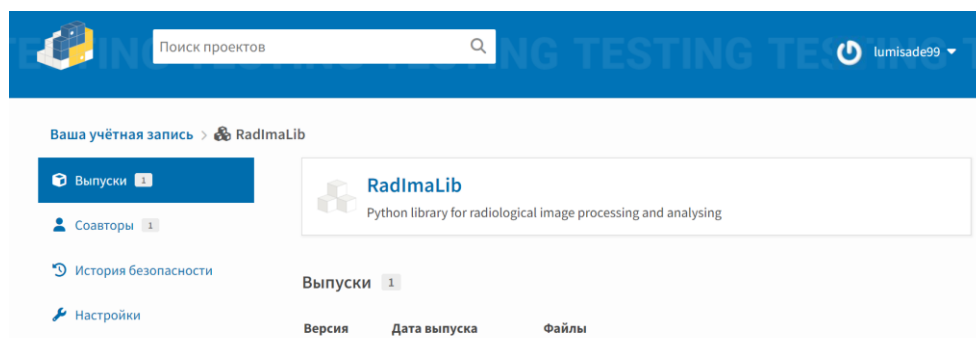


Рисунок 3 – Результат публикации библиотеки RadImaLib для работы с радиологическими изображениями опухолей шеи

По завершении всех процедур, библиотека готова к запуску и работе.

Таким образом, проведен анализ существующих библиотек для предобработки и анализа радиологических изображений, а также разработана ранняя версия собственной библиотеки RadImaLib с интегрированным датасетом и функциями обработки изображений.

Заключение. В современном мире развитие персонализированных методов поддержания здоровья играет огромную роль в силу индивидуальной специфики и редкости некоторых заболеваний, в частности, рассмотренном в рамках данной работы. Проведенное исследование подтверждает, что использование современного информационного обеспечения создает основу для повышения информированности лечащих врачей и обоснованности клинических решений при диагностике и лечении редких заболеваний. В результате исследования обоснован состав компонентов и определены требования к системе информа-

ционного обеспечения диагностики редких заболеваний, спроектирована база данных системы, разработана и опубликована версия библиотеки для работы с радиологическими изображениями опухоли шеи и датасетом, представляющая средства для анализа пользовательских данных о пациентах и классификации. Набор данных пациентов с каротидной опухолью и схожими заболеваниями сформирован и интегрирован в библиотеку (возможно его обновление за счет добавления информации о новых пациентах). Реализованы такие операции обработки радиологических изображений, как применение стандартных преобразований (Hounsfield Scale, Erosion, Dilation) и rescaling. Итоговая программная библиотека может быть использована IT-исследователями в области работы с медицинскими данными, в частности, для взаимодействия и самостоятельной доработки методов извлечения медицинской информации и распознавания радиологических изображений опухоли шеи.

Список источников

1. Вачев А.Н., Дмитриев О.В., Зелёва О.В., Зельтер П.М. Каротидная хемодектома в практике оториноларинголога: клиническое наблюдение // Кубанский научный медицинский вестник. 2020. № 27 (2). С. 135–143. DOI: 10.25207/1608-6228-2020-27-2-135-143
2. Абдулгасанов Р.А., Газимагомедов З.И. Современные принципы диагностики и лечения параганглиом (хемодектом) шеи // Комплексные проблемы сердечно-сосудистых заболеваний. 2016. № 1. С. 71–75.
3. Пинский С.Б., Дворниченко В.В., Репета О.Р. Редкое наблюдение множественной злокачественной параганглиомы шеи // Сибирский медицинский журнал (Иркутск). 2009. № 3. С. 131–134.
4. Forbes J., Menezes R.G. Anatomy, Head and Neck: Carotid Bodies. 2023 Jul 25. In: Stat Pearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan. PMID: 32965908. (In Eng.).
5. Дружинин Д.С., Пизова Н.В. Каротидная хемодектома: дифференциальная диагностика по данным ультразвукового исследования // Опухоли головы и шеи. 2012. № 1. С. 46–50.
6. Baysal B.E., Willett-Brozick J.E., Lawrence E.C. et al. Prevalence of SDHB, SDHC, and SDHD Germline Mutations in Clinic Patients with Head and Neck Paragangliomas // J. Med. Genet. 2002. № 39. С. 178–183. (In Eng.).

References

1. Vachev A.N., Dmitriev O.V., Zeleva O.V., Zelter P.M. Carotid Chemodectoma in the Practice of an Otorhinolaryngologist: a Clinical Observation. *Kubanskiy nauchnij medicinskiy vestnik*. 2020. No. 27 (2). pp. 135–143. (In Russ.). DOI: 10.25207/1608-6228-2020-27-2-135-143
2. Abdulgasanov R.A., Gazimagomedov Z.I. Modern Principles of Diagnostics and Treatment of Neck Paragangliomas (Chemodectomas). *Kompleksnyye problemy serdechno-sosudistykh zabolevanij*. 2016. No. 1. pp. 71–75. (In Russ.).
3. Pinsky S.B., Dvornichenko V.V., Repeta O.R. Rare Case of Multiple Malignant Paraganglioma of the Neck. *Sibirskiy medicinskiy zhurnal (Irkutsk)*. 2009. No. 3. pp. 131–134. (In Russ.).
4. Forbes J., Menezes R.G. Anatomy, Head and Neck: Carotid Bodies. 2023 Jul 25. In: Stat Pearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan. PMID: 32965908.
5. Druzhinin D.S., Pizova N.V. Carotid Chemodectoma: Differential Diagnosis Based on Ultrasound Data. *Opuholi golovy i shei*. 2012. No. 1. pp. 46–50. (In Russ.).
6. Baysal B.E., Willett-Brozick J.E., Lawrence E.C. et al. Prevalence of SDHB, SDHC, and SDHD Germline Mutations in Clinic Patients with Head and Neck Paragangliomas. *J. Med. Genet*. 2002. No. 39. pp. 178–183.

7. PubMed // National Library of Medicine. Официальный сайт [Электронный ресурс]. – Режим доступа: <https://pubmed.ncbi.nlm.nih.gov/> (In Eng.).
8. Cochrane Library // Cochrane. Официальный сайт [Электронный ресурс]. – Режим доступа: <https://www.cochrane.org/> (In Eng.).
9. UpToDate // Wolterskluwer. Официальный сайт [Электронный ресурс]. – Режим доступа: <https://www.uptodate.com/> (In Eng.).
10. Fernando Pérez-García, Rachel Sparks, Sébastien Ourselin. TorchIO: A Python Library for Efficient Loading, Preprocessing, Augmentation and Patch-based Sampling of Medical Images in Deep Learning // *Computer Methods and Programs in Biomedicine*. 2021. Т. 208. С. 106236. (In Eng.). DOI: 10.1016/j.cmpb.2021.106236
11. MedPy [Электронный ресурс]. – Режим доступа: <https://pypi.org/project/MedPy/> (In Eng.).
12. Chen H. DICOM Processing and Segmentation in Python [Электронный ресурс]. – Режим доступа: <https://www.raddq.com/dicom-processing-segmentation-visualization-in-python/> (In Eng.).
13. PyPI [Электронный ресурс]. – Режим доступа: <https://pypi.org/> (In Eng.).
14. Eisinga K. How to Create a Python Library [Электронный ресурс]. – Режим доступа: <https://medium.com/analytics-vidhya/how-to-create-a-python-library-7d5aea80cc3f> (In Eng.).
7. PubMed. *National Library of Medicine. Official site*. Available at: <https://pubmed.ncbi.nlm.nih.gov/>
8. Cochrane Library. *Cochrane. Official site*. Available at: <https://www.cochrane.org/>
9. UpToDate. *Wolterskluwer. Official site*. Available at: <https://www.uptodate.com/>
10. Fernando Pérez-García, Rachel Sparks, Sébastien Ourselin. TorchIO: A Python Library for Efficient Loading, Preprocessing, Augmentation and Patch-based Sampling of Medical Images in Deep Learning. *Computer Methods and Programs in Biomedicine*. 2021. Vol. 208. P. 106236. DOI: 10.1016/j.cmpb.2021.106236
11. MedPy. *Electronic resource*. Available at: <https://pypi.org/project/MedPy/>
12. Chen H. DICOM Processing and Segmentation in Python. *Electronic resource*. Available at: <https://www.raddq.com/dicom-processing-segmentation-visualization-in-python/>
13. PyPI. *Electronic resource*. Available at: <https://pypi.org/>
14. Eisinga K. How to Create a Python Library. *Electronic resource*. Available at: <https://medium.com/analytics-vidhya/how-to-create-a-python-library-7d5aea80cc3f>