

Набор данных МРТ головного мозга (рассеянный склероз) для самотестирования ИИ-сервисов, репрезентативный для ДЗМ

Данный набор данных (датасет) содержит 172 МРТ изображений головного мозга и результаты разметки признаков рассеянного склероза. Данные исследования были собраны в отделениях лучевой диагностики лечебных учреждений города Москвы в период с 23.11.2016 по 26.06.2021.

DISCLAIMER

Набор данных предназначен для следующих целей:

- разработка, обучение и тестирование точности программных продуктов (использующих в том числе методы компьютерного зрения), выявляющих признаки, характерные для рассеянного склероза;
- информирование медицинского сообщества и общественности в целом.

Лицензия **позволяет свободно делиться (обмениваться)** набором данных, то есть копировать и распространять материал на любом носителе и в любом формате, **при обязательном соблюдении следующих условий:**

- указано авторство, а именно:
 - авторы;
 - их организации;
 - правообладатель (копирайт);
 - постоянная ссылка на оригинальный набор данных.
- указана ссылка на лицензию.

Лицензия **запрещает**, в том числе:

- использовать набор данных в коммерческих целях;
- распространять переработанный, преобразованный набор данных или новые наборы данных, созданные на основе этого набора;
- накладывать ограничения поверх существующих ограничений, указанных в лицензии, например:
 - предоставлять платный доступ к набору данных;
 - искусственно сдерживать распространение набора данных техническими методами.

Общая информация

Название набора данных

Набор данных МРТ головного мозга (рассеянный склероз) для самотестирования ИИ-сервисов, репрезентативный для оборудования ДЗМ

Внутренний код

SCLEROSIS21_172

Классы разметки

1|0

Ключевые слова

магнитно-резонансная томография, МРТ, рассеянный склероз, признаки демиелинизации, головной мозг, само-тестирование

Язык

Английский, русский

Финансирование

Внутреннее финансирование

Версия набора данных

1.0

Постоянная ссылка

https://mosmed.ai/datasets/mri_scler_172

Дата публикации

22.10.2022

Аффилиация и авторы

Авторы

- Евтеева Ксения Борисовна¹
- Низовский Константин Андреевич¹
- Хоружая Анна Николаевна¹
- Анищенко Леся Николаевна¹
- Андрейченко Анна Евгеньевна¹
- Морозов Сергей Павлович¹
- Кремнева Елена Игоревна¹

Аффилиация

1. Государственное бюджетное учреждение здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы» (далее – НПКЦ ДиТ ДЗМ)

Структура набора данных

```
.
|-- dataset_registry.csv
|-- labels.xlsx
|-- LICENSE
|-- README_EN.md
|-- README_RU.md
|-- README_EN.pdf
|-- README_RU.pdf
'-- scanners
    |-- scanner_1
    |   |-- study_1
    |   |   |-- series_instance1
    |   |   |   | sop_instance1
    |   |   |   | ...
    |   |   |   ' sop_instanceN
    |   |   |   ...
    |   |   '-- series_instanceN
    |   |       | sop_instance1
    |   |       | ...
    |   |       ' sop_instanceN
    |   |-- ...
    |   '-- study_N
    |       |-- series_instance1
    |       |   | sop_instance1
    |       |   | ...
    |       |   ' sop_instanceN
    |       |   ...
    |       '-- series_instanceN
    |           | sop_instance1
    |           | ...
    |           ' sop_instanceN
    |-- ...
    |-- ...
    |-- ...
    '-- scanner_N
        |-- study_1
        |   |-- series_instance1
        |   |   | sop_instance1
        |   |   | ...
        |   |   ' sop_instanceN
        |   |   ...
        |   '-- series_instanceN
        |       | sop_instance1
        |       | ...
        |       ' sop_instanceN
        |-- ...
        '-- study_N
            |-- series_instance1
            |   | sop_instance1
            |   | ...
            |   ' sop_instanceN
            |   ...
            '-- series_instanceN
                | sop_instance1
                | ...
                ' sop_instanceN
```

- `README_EN.md` и `README_RU.md` содержат общую информацию о наборе данных в формате `Markdown` на английском и русском языках соответственно; та же информация в формате `PDF` представлена в `README_EN.pdf` и `README_RU.pdf`.
- `LICENSE` содержит описание лицензии Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

License.

- В директории `scanners` находятся 43 поддиректории, соответствующие различным аппаратам ДУ, находящимся в отделениях лучевой диагностики лечебных учреждений города Москвы
- В каждой директории ДУ находится по 4 поддиректории, соответствующие различным исследованиям, всего 172 исследования
- В каждой директории исследования находятся поддиректории, соответствующие различным сериям
- В каждой директории серии находятся поддиректории, соответствующие различным `DICOM` файлам

Список ДУ

Датасет содержит по 4 исследования для каждого устройства.

В датасет включены данные со следующих устройств:

Производитель	Модель
GE Healthcare	Signa Voyager
GE Medical Systems, LLC	Signa 0.2T
GE Medical Systems, LLC	Signa 1.5T
GE Medical Systems, LLC	Brivo mr 355 1,5 T
GE Medical Systems, LLC	Signa HD
GE Medical Systems, LLC	Signa Explorer
GE Medical Systems, LLC	Excelart Vantage Atlas-X
General Electric (GE)	Signa HD
Siemens	Magnetom Avanto
Siemens	MAGNETOM Aera
Siemens	Magnetom Skyra
Siemens	Magnetom Essenza
Toshiba Medical Systems Corporation	Excelart Vantage Titan 3T
Toshiba Medical Systems Corporation	Excelart Vantage Atlas-M
Toshiba Medical Systems Corporation	Excelart Vantage Atlas-X
Toshiba	EXCELART Vantage Atlas-X
Philips	Panorama HFO
Philips	ingenuity
Philips	Achieva 3.0 T
Philips	Ingenia 3.0T
Philips	Ingenia 1.5T
Philips	Ingenia Ambition S
Philips	Intera Achieva
«С.П. Гелпик»	ТМРпм-"Ренекс"
Hitachi	Aperto 2009

Обзор данных

Параметр	Значение
Количество исследований, ед.	172
Количество пациентов, чел.	172
Распределение по полу, % (М/ Ж)	56/ 116
Распределение по возрасту, лет (мин./ медиана/ макс.)	18/ 46/ 85
Распределение по классам, % (наличие целевой патологии/отсутствие целевой патологии)	50/ 50

Особенности подготовки исследований

- Одно исследование относится к одному пациенту.
- Каждое исследование анонимизировано в соответствии со стандартом DICOM PS3.15 2021b (part E1.1 Table E.1-1 Basic Profile), удалены приватные теги, теги (0008, 0023), (0008, 0033) оставлены пустыми.

Методика разметки

- МРТ-исследования были размечены в соответствии с выводами о наличии целевой патологии, сделанными на основе текстовых заключений врачей-рентгенологов

При разметке использовалось собственное программное обеспечение НПКЦ ДИТ ДЗМ.

Описание структуры файла разметки

Файл разметки имеет имя `labels.xlsx` и представляет собой таблицу с двумя столбцами: `study_uid` соответствует уникальному идентификатору исследования и имени файла исследования, `pathology` соответствует наличию/отсутствию целевой патологии. 0 -

патология не выявлена, 1 - в текстовом описании имеются признаки, указывающие на наличие целевой патологии.

Правила использования и распространения

Лицензия

Copyright © 2020 Государственное бюджетное учреждение здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы».

Набор данных доступен под лицензией Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0) License. За подробной информацией обратитесь к файлу [LICENSE](#) или пройдите по [ссылке](#).

Цитирование

Рекомендованная форма для цитирования:

В печати: Ksenia B Evteeva, Konstantin A Nizovskiy, Anna N Khoruzhaya, Lesya N Anishchenko, Anna E Andreychenko, Sergey P Morozov, Elena I Kremneva. MRISclerosisCa-172: Open-Access MRI Sclerosis Database.

Распространение

Данный датасет не должен распространяться без указания:

- авторов;
- аффилиций;
- правообладателя (копирайта);
- постоянной ссылки на оригинальный набор данных;
- ссылки на лицензию.