README.md 2025-09-25

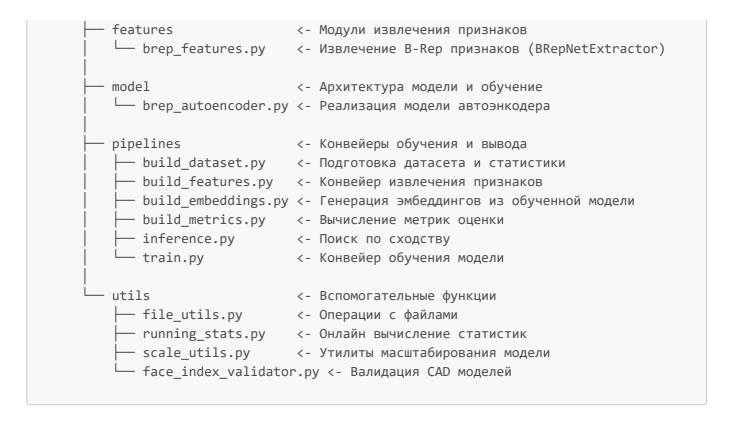
BRepMocAutoencoder

Автоэнкодер для B-Rep представлений CAD моделей с поиском похожих объектов по эмбеддингам.

Структура проекта

```
— LICENSE <- Лицензия с открытым исходным кодом
Makefile
                     <- Makefile c командами типа `make data` или `make train`
├─ data
    - external
                    <- Данные из внешних источников
    — interim <- Промежуточные преобразованные данные 
— processed <- Финальные готовые датасеты для моделирования
   L— raw
                     <- Исходные неизменяемые данные
├─ docs
                     <- Файлы документации
 — models
                      <- Обученные и сериализованные модели (best.ckpt)
                      <- Jupyter блокноты для экспериментов и тестирования
    ├─ notebook_setup.py <- Скрипт настройки для блокнотов
    └─ test_model.ipynb <- Блокнот для тестирования модели
— references     <- Словари данных, руководства и пояснительные материалы
                     <- Сгенерированные отчёты в HTML, PDF, LaTeX, Excel и т.д.
- reports
   occwl <- Библиотека-обёртка OpenCASCADE для CAD операций
— base.py <- Базовые классы и утилиты
— compound.py <- Операции со сложными формами
— edge.py <- Обработка рёбер
— face.py <- Обработка граней
 — occwl
    <- Другие модули обработки CAD
  - src
                     <- Исходный код проекта
    ├─ __init__.py
                              <- Делает src Python модулем
    — config.py
                              <- Хранение полезных переменных и конфигурации
     — dataset.py
                             <- Классы датасетов для обучения
    plots.py
                              <- Код для создания визуализаций
    — metrics.py
                              <- Метрики оценки для поиска по сходству
     — searcher.py
                              <- Функционал поиска по сходству
```

README.md 2025-09-25



Структура датасета

```
data
  ├─ external <- Данные из внешних источников
      ├── s2.0.0_step_all_features.json <- Конфигурация признаков STEP
        - segment_names.json <- Имена сегментов
- feature lists/ <- Списки признаков
       — feature_lists/
└─ all ison
                                      <- Списки признаков
          ─ all.json
                                      <- Все признаки
          - no_edge_*.json
- no_face_*.json
          └─ no_face_*.json
                                      <- Без различных признаков граней
        - kernels/
                                      <- Конфигурации ядер
                                   <- Асимметричные ядра
<- Простое ребро
          -- asymmetric*.json
           — simple_edge.json
          └─ winged_edge*.json
                                      <- Крылатые рёбра
  — interim
                   <- Промежуточные преобразованные данные
    - processed <- Финальные готовые датасеты для моделирования
       — dataset stats.json <- Статистики датасета
       — features/ <- Извлечённые признаки

├─ brep/ <- В-Rep признаки (.прz файлы)
       — embeddings/ <- Эмбеддинги моделей (.npz файлы)
— steps/ <- STEP файлы CAD моделей (.stp)
                   <- Исходные неизменяемые данные
```

В конвейрах обработки уже заложены пути по умолчанию необходимо только либо перенести файлы CAD-моделей в папку steps, либо в build_features.py указать step_path_dir директорию, где находятся CAD модели

README.md 2025-09-25

Основные компоненты

Конвейеры обработки (Pipelines)

- build_features.py Извлечение B-Rep признаков из CAD моделей
- build_dataset.py Подготовка датасета и вычисление статистик
- train.py Обучение автоэнкодера
- build_embeddings.py Генерация эмбеддингов из обученной модели
- inference.py Поиск похожих моделей по эмбеддингам
- build_metrics.py Вычисление метрик качества

Запуск конвейеров

```
# Извлечение признаков
python -m src.pipelines.build_features

# Подготовка датасета
python -m src.pipelines.build_dataset

# Обучение модели
python -m src.pipelines.train

# Генерация эмбеддингов
python -m src.pipelines.build_embeddings

# Поиск похожих моделей
python -m src.pipelines.inference

# Вычисление метрик
python -m src.pipelines.build_metrics
```

Для получения справки по параметрам каждого конвейера используйте флаг --help.