

Руководство по установке Сервиса инференса MarQus.

Сервис инференса нейросетевых моделей MarQus для автоматизации процессов распознавания и сортировки различных типов ТБО.

2023

Аннотация

В настоящем документе представлено руководство по установке экземпляра сервиса инференса нейросетевых моделей MarQus.

Развёртывание локально.

Сервис инференса нейросетевых моделей для автоматизации процессов распознавания и сортировки различных типов ТБО (marqus-video-recotracker) представляет собой набор микросервисов, реализованных и запускаемых в платформе Docker.

Системные требования

Требования к предустановленному ПО:

- Операционная система: Любой дистрибутив Linux с ядром 5.4 или выше
- CPU: Intel(R) Core(TM) i7-7700K CPU @ 4.20GHz and newer
- RAM: минимум 8 GB
- HDD/SSD: минимум 64 GB
- GPU: Nvidia 1080Ti или новей
- GPU: память не ниже 11264 МБ
- Compiler: GCC 7.3.1 или новей
- CUDA: 11.0 или новей
- Docker: 20.10 или новей

Установка и запуск

Для установки и запуска сервиса необходимо:

1. Скачать образ, используя команду

```
docker pull harbor.zebrains.team/mlearning/marqus-video-recotracker:latest
```

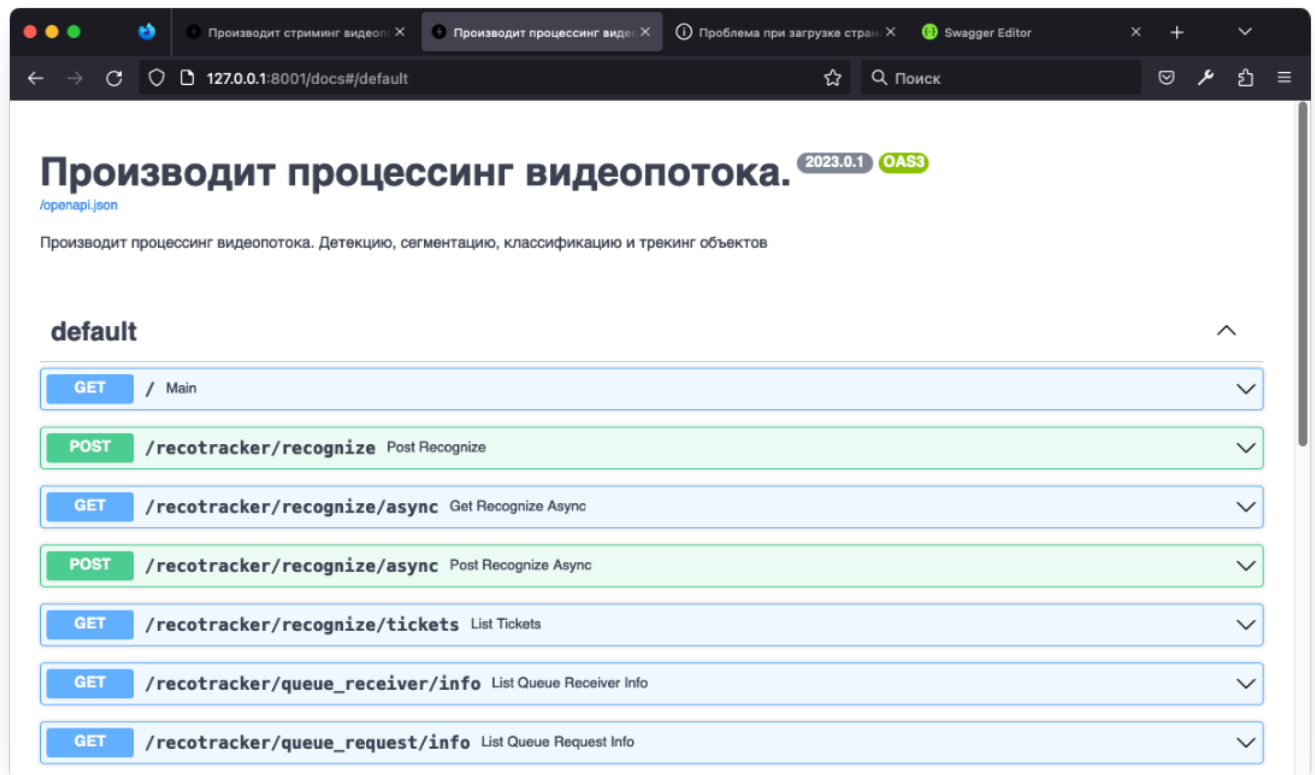
2. Запустить новый контейнер из образа

```
docker run --name marqus-video-recotracker -p 8001:8001/tcp -p 8002:8002/tcp harbor.zebrains.team/mlearning/marqus-video-recotracker:latest
```

3. Проверить в браузере

Открыть ссылку <http://127.0.0.1:8001>

Должны увидеть API сервиса marqus-video-recotracker:



4. Проверить эндпоинт `/recotracker/info`

В Response body должны увидеть информацию о доступности cuda.

```
"torch_info": {  
  "cuda.is_available": true  
},
```

Готово, решение запущено.

Если в результате увидим

```
"torch_info": {  
  "cuda.is_available": false  
},
```

это означает, что cuda недоступна. Сервис всё равно будет работать, но производительность сервиса упадёт на порядок.