

Правительство Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»
(НИУ ВШЭ)

Программное обеспечение
«ФРЕЙМВОРК АВТОМАТИЗАЦИИ СИНТЕЗА МОДЕЛЕЙ
КОМПЬЮТЕРНОГО ЗРЕНИЯ AutoOD»

Описание функциональных характеристик

на 10 листах

Москва 2024

ОГЛАВЛЕНИЕ

1. Общие сведения.....	3
1.1. Обозначение и наименование программы.....	3
1.2 Назначение программы	3
1.3 Языки программирования и дополнительное программное обеспечение	3
2. Функциональное назначение.....	5
3. Используемые технические средства	6
4. Входные данные и выходные данные программы	7
4.1 Модуль №1 Предобработка набора данных.....	7
4.2 Модуль №2 Рекомендация архитектуры	8
4.3 Модуль №3 Подготовка окружения и конфигурация моделей.....	8
4.4 Модуль №4 Обучение модели	9
4.5 Модуль №5 Обнаружение объектов	9
4.6 Модуль №6 Трекинг объектов.....	9
4.7 Модуль №7 Отслеживание порядка появления объектов	10
4.8 Модуль №8 Экспорт модели.....	10

1. Общие сведения

1.1. Обозначение и наименование программы

Наименование ПО - «Фреймворк автоматизации синтеза моделей компьютерного зрения AutoOD».

1.2 Назначение программы

ПО предназначено для решения задач по подготовке наборов данных изображений, обучения моделей компьютерного зрения, а также встраивания различных функций на их основе в прикладные программы. Программа решает следующие прикладные задачи:

1. Балансировка наборов данных изображений по экземплярам объектов классов между обучающей и валидационной выборками.
2. Визуализация аннотаций наборов данных изображений.
3. Генерация конфигурации модели компьютерного зрения для фреймворка MMDetection.
4. Обучение моделей компьютерного зрения для обнаружения объектов.
5. Распознавание объектов на изображениях с помощью обученной или предобученной модели из зоопарка фреймворка MMDetection.
6. Отслеживание (реидентификация) объектов на последовательностях кадров видеозаписи.
7. Анализ порядка появления объектов на кадрах видеозаписи.
8. Экспорт модели компьютерного в формат ONNX.

Программа предназначена для использования разработчиками ПО без значительных компетенции в сфере анализа данных, машинного обучения и компьютерного зрения.

1.3 Языки программирования и дополнительное программное обеспечение

Программное обеспечение написано на языках программирования Python.

Python – это высокоуровневый объектно-ориентированный язык программирования общего назначения. Основной функционал ПО для конфигурации и обучения моделей, а также для полногеномных предсказаний написан на Python.

MMSCV – фреймворк на языке Python, автоматизирующий рутинные задачи разработки и исследования в области компьютерного зрения.

MMDet – фреймворк для разработки приложений с использованием обнаружения объектов на изображениях.

Norfair – библиотека, реализующая функции отслеживания экземпляров объектов на последовательностях кадров видеозаписи. В качестве входных данных используются данные о обнаруженных на текущем и предыдущих кадрах объектах.

- Модули языка Python:
 - openmim
 - mmengine>=0.7.0
 - ninja
 - psutil
 - pytorch
 - torchvision
 - mmdetection
 - norfair>=v2.2.0

2. Функциональное назначение

Программа предназначена для синтеза моделей компьютерного зрения с использованием нейросетевых моделей глубинного обучения разных типов архитектур в целях применения пользователем, не обладающим глубокими навыками машинного обучения и анализа данных и/или не имеющего опыта применения алгоритмов искусственного интеллекта.

Программа реализует следующие функции:

- предварительную обработку наборов данных изображений (Модуль №1);
- механизм поддержки принятия решений для автоматизированного выбора модели компьютерного зрения в зависимости от выделяемых вычислительных ресурсов и параметров набора данных (Модуль №2);
- автоматический подбор параметров обучения модели (Модуль №3);
- обучение модели (Модуль №4);
- обнаружение объектов на изображениях с применением обученной модели (Модуль №5);
- отслеживание объектов на последовательности изображений (Модуль №6);
- отслеживание порядка появления объектов на последовательности изображений (Модуль №7);
- экспорт обученной модели в формат ONNX (Модуль №8).

3. Используемые технические средства

Для установки, запуска и стабильного функционирования программы требуется сервер со следующими минимальными системными требованиями:

- Количество физических ядер процессора (CPU) - 4;
- Оперативная память (RAM) - 32 Гб;
- Дисковое пространство (HDD) – 256 Гб;
- GPU Nvidia с поддержкой CUDA, объём видеопамяти - не менее 8 гигабайт;
- Сервер должен быть подключен к ИБП;
- Операционная система Ubuntu версии 22.XX и выше.

4. Входные данные и выходные данные программы

4.1 Модуль №1 Предобработка набора данных

Предобработка данных включает в себя

- Конвертацию форматов данных.
- Балансировку классов (при делении на тестовую и обучающую выборки).
- Визуализацию аннотаций.
- Анализатор датасета.
- Рекомендация по выбору предпочтительных значений и балансировке

датасета.

Входные данные:

- Конвертация форматов данных:
 - Путь к директории с набором данных, размеченном в формате Pascal VOC или YOLO;
 - Путь к директории для сохранения аннотаций набора данных, сконвертированной в формат COCO JSON.
- Балансировка классов:
 - Путь к директории для сохранения разбитого на train и test набора данных.
- Визуализация:
 - Путь к файлу с аннотациями в формате COCO JSON;
 - Путь к папке для сохранений изображений с нанесёнными аннотациями.
- Анализатор:
 - Путь к файлу с аннотациями набора данных в формате COCO JSON.

Выходные данные:

- Конвертация форматов данных:
 - Файл аннотаций набора данных изображений в формате COCO JSON.
- Балансировка классов:
 - Директории train и test с данными из набора данных, сбалансированными по количеству объектов разных классов на изображениях в каждой из директорий.
- Визуализация:
 - Директория с изображениями с нанесёнными на них аннотациями.
- Анализатор:

- Изображения с графиками, отображающими характеристики датасета;
- Характеристики набора данных в текстовом виде;
- Рекомендации по выбору гиперпараметров модели в текстовом виде.

4.2 Модуль №2 Рекомендация архитектуры

Входные данные:

- Необходимая скорость: низкая, средняя, высокая.
- Необходимая точность: высокая, очень высокая.
- Доступные ресурсы: низкие, средние, высокие.
- Характеристики набора данных:
 - Размеры объектов на изображении: малый, средний, большой;
 - Разрешение изображений: ширина, высота;
 - Количество объектов на изображении: малое, среднее, большое;
 - Разнообразие классов: малое, среднее, большое.

Выходные данные:

- Список с названиями одной или нескольких моделей, подходящих под заданные параметры.

4.3 Модуль №3 Подготовка окружения и конфигурация моделей

Функции:

- • Загрузка предобученной модели и её базовой конфигурации
- • Создание изменённой конфигурации модели с учётом параметров набора данных и локального окружения

1. Загрузка предобученной модели и её базовой конфигурации

Входные данные:

- Имя выбранной модели
- Файл с конфигурацией модели в формате YAML

Выходные данные:

- Подготовленные к работе директории и файлы:
 - checkpoints: директория, содержащая файл с весами модели после предобучения и данные об архитектуре выбранной модели
 - configs: базовая конфигурация выбранной модели
- 2. Создание изменённой конфигурации модели с учётом параметров набора данных и локального окружения

Данные о конфигурации модели, хранящиеся в файле в формате YAML, и параметры обучения вносятся в итоговый файл конфигурации модели.

Входные данные:

- Файл с конфигурацией модели в формате YAML
- Файл с базовой конфигурацией выбранной модели

Выходные данные:

- Изменённый (итоговый) файл конфигурации модели, подготовленный для обучения и инференса

4.4 Модуль №4 Обучение модели

Входные данные:

- Файл с конфигурацией модели в формате YAML
- Файл конфигурации модели, подготовленный для обучения и инференса

Выходные данные:

- Обученная модель в формате PTN

4.5 Модуль №5 Обнаружение объектов

Входные данные:

- Файл конфигурации модели, подготовленный для обучения и инференса
- Файл в формате PTN с весами модели после предобучения и данные об архитектуре выбранной модели

- Путь к изображению для обработки

Выходные данные:

- Предсказания модели

4.6 Модуль №6 Трекинг объектов

Входные данные:

- Параметры трекинга объекта:
 - Максимальное количество кадров, в течение которых объект может отслеживаться без обнаружения;
 - Задержка перед инициализацией отслеживания для нового объекта;
 - Порог для обнаружения объектов;
 - Количество прошлых обнаружений, которые следует учитывать при отслеживании. Трекер будет анализировать не только последнее обнаружение объекта, но и некоторое количество его предыдущих обнаружений для более точного предсказания его текущего и будущего положения;
 - Функция для вычисления расстояния между отслеживаемым объектом и объектами, обнаруженными на новых кадрах.

- Файлы конфигурации модели, подготовленный для обучения и инференса
- Файлы в формате PTN с весами модели после предобучения и данные об архитектуре выбранной модели

- Идентификатор устройства захвата видео/путь к видеофайлу

Выходные данные:

- csv-файл с данными об объектах:
 - номер кадра;
 - идентификатор объекта;
 - координаты bounding box-a.
- Кадры видео с визуализацией информации об отслеживаемых объектах

4.7 Модуль №7 Отслеживание порядка появления объектов

Входные данные:

- Параметры фреймворка Norfair
- Файлы конфигурации модели, подготовленный для обучения и инференса
- Файлы в формате PTN с весами модели после предобучения и данные об архитектуре выбранной модели

- Идентификатор устройства захвата видео/видеофайл
- Список объектов в ожидаемом порядке их появления на видео

Выходные данные:

- csv-файл с данными об объектах:
 - номер кадра;
 - идентификатор объекта;
 - координаты bounding box-a;
 - оценка, ожидалось ли появление объекта в кадре.
- кадры видео с визуализацией информации об отслеживаемых объектах

4.8 Модуль №8 Экспорт модели

Входные данные:

- Файл конфигурации модели, подготовленный для обучения и инференса
- Файл в формате PTN с весами модели
- Путь, по которому будет сохранён файл модели в выбранном для экспорта формате

Входные данные:

- Файл с весами модели в выбранном для экспорта формате.