

Итоговое задание для курса “Профессия ML-инженер”

Цель: Решить одну из предложенных задач используя методы глубокого обучения или классического компьютерного зрения, пройденных на курсе.

Задачи:

- Проанализировать существующие решения для выбранной темы и подобрать релевантный **датасет**;
- Представить **блок-схему** или диаграмму пайплайна обучения и инференса модели(ей), основываясь на найденных решениях и знаниях, полученных на курсе;
- Написать **скрипт для обучения** модели(ей);
- Написать **интерфейс** для работы с обученной моделью(ями);
- Оценить качество модели по выбранным для задачи **метрикам**, сравнить качество на валидационной и тестовой выборках;
- Написать **отчет** (2-4 стр), который будет включать в себя:
(a) обоснование выбора датасета и метрик, (b) предложенное решение, (c) архитектуру и оценку сложности модели(ей), (d) детали обучения, (e) выводы о пригодности вашего подхода к решению выбранной задачи;
- Подготовить финальную **презентацию** проекта.

Форма контроля: защита проектной работы.

Форма работы: индивидуальная.

Предлагаемые темы проектов (*темы, на которые может уйти чуть больше времени и сил):

Проект 1. Система распознавания жестов в реальном времени

- Обнаружение рук в реальном времени с помощью OpenCV
- Модель классификации жестов с использованием PyTorch
- Интерфейс, позволяющий в реальном времени считывать жесты человека или нескольких людей с камеры или видео

Проект 2. Удаление шума из изображений с использованием глубоких автокодировщиков

- Обучение автокодировщика, с подбором различных гиперпараметров
- Модуль моделирования шума (как минимум, 3 различных вида)
- Сравнение модели с классическими методами удаления шума из OpenCV
- Тестирование на стороннем датасете с зашумленными картинками

Проект 3. Подсчет машин на видео

- Обучение или выбор уже готовой нейронной сети (например, YOLO), для детекции машин
- Реализация модуля трекинга объектов на видео, сопоставляющий боксы одних и тех же машин на последовательных кадрах. Видео может быть снятым как из салона машины, так и со статичной камеры
- Подсчет машин как количество завершенных треков (те если машина выехала из кадра)

- Оценка сложности метода и оптимизация для работы на видео (~25-30 FPS)

Проект 4. Распознавание действий на видео с использованием 3D-CNNS или связки CNN + RNN

- Извлечение кадров с помощью OpenCV, предварительной обработки и подачи в модель Pytorch.
- Модель для извлечение признаков действий из изображений и анализ этих признаков, например, с помощью рекуррентной модели
- Пример действий: ходьба, бег, прыжки и тд. Длина видео может составлять около 3-10 сек при 30-60 FPS

Проект 5*. Обнаружение оставленных предметов по записи видео

- Обучение или выбор уже готовой нейронной сети (например, YOLO), для детекции небольших предметов (рюкзаков, пакетов, коробок)
- Реализация модуля трекинга таких объектов
- Реализация модели долгосрочного и краткосрочного фонов
- Обнаружение оставленного предмета как его переход из краткосрочного в долгосрочный фон

Проект 6*. Распознавание рукописного текста с помощью глубокого обучения

- Подготовка изображения (выбор цветового пространства, обнаружение текста, перспектива, и пр) с рукописным текстом с помощью OpenCV
- Обучение глубокой нейронной модели для распознавания текста

- Подсчет и анализ ошибочно распознанных слов
- Сравнение с Tesseract
- Выбор языка (русский или английский) остается на стороне студента

Проект 7. Семантическая сегментация с помощью U-Net на уличных сценах

- Реализация U-Net или DeepLabv3 для пиксельной классификации дорожных сцен (дорог, пешеходов, транспортных средств и пр)
- Оценка, например, на Cityscapes или Pascal VOC.
- Визуализация масок сегментаций, наложенных на исходные изображения, с помощью OpenCV

Проект 8. Классификация дорожных знаков

- Реализация CNN-классификатора на датасете вроде GTSRB для распознавания различных (не менее 30 видов) дорожных знаков
- Оценка качества работы алгоритма на зашумленных или аугментированных картинках

Проект 9. Распознавание эмоций на видео

- Реализация модуля поиска лица на видео
- Обучение модели распознавания не менее 5 эмоций
- Реализация интерфейса, позволяющий считывать эмоции человека с видео или веб-камеры, с выводом результата и вероятностью эмоций

Проект 10. Раскрашивание черно-белых изображений

- Перевести изображение в наиболее подходящее цветовое пространство
- Обучение глубокой модели, например, U-Net, для раскрашивания
- Оценка на стороннем датасете

Проект 11*. Обнаружение аномалий на записях видеонаблюдения

- Извлекать кадры из видео
- Обучить связку автокодировщика для извлечения фичей из кадра и рекуррентной сети для классификации на коротких видео из распределения, где люди ведут себя естественно
- Проверить этот алгоритм на видео, например, с драками, применением оружия и пр

Своя тема проекта. В качестве финального проекта допускается выбрать тему не из предложенного списка. Задачи, однако, для этой темы остаются такими же, а сама тема должна быть связана с глубоким обучением, и (или) компьютерным зрением.