Результаты выполнения тестового задания

**1. Поиск рамок чёрного цвета разной ширины при помощи ИИ**

*Что использовано.* Для выполнения этого задания использованы такие библиотеки Python как OpenCV, SciPy, NumPy и некоторые другие стандартные библиотеки. Примерно за 1 час реализована наиболее простая версия программы, которая определяет на RGB-изображении черные области размера 3\*26 (размер массива – 3\*26\*3 пикселя).

Принцип работы программы – выполнение свертки исходного изображения с окном размера 3\*26\*3, заполненного единицами. По сути, результатом свертки является сумма всех значений пикселей окна. Условие отнесения фрагмента изображения к черному цвету самое простое – сумма значений пикселей должна быть равна 0. Подходящие под данное условие области изображения закрашиваются зеленым цветом и сохраняются в специальную папку.

Программа протестирована на изображениях из папки «Crops/60img/G32V3H0» (60 изображений). Обе папки (исходные и обработанные изображения), а также код программы можно найти в репозитории.

*Перспективные инструменты.* В случае востребованности данной программы возможен такой вариант улучшения программы как экспериментальное определение оптимального порогового значения суммы пикселей по окну. Для этого можно применять размеченные изображения и специальные метрики машинного обучения.

Кроме того, необходимо будет провести оптимизацию кода программы для увеличения ее производительности (например, применять np.sum вместо свертки, если веса окна одинаковые, или переписать код с использованием многопоточности).

**2. Поиск соответствия 10-ти классам (mnist) (способ 1 – от 08.06.2022)**

Одна из типовых задач классификации – определить, какая из рукописных цифр отображена на черно-белом изображении 28\*28 пикселей.

*Что использовано.* Ранее разрабатывался простейший фреймворк для работы с нейронными сетями на языке MATLAB. А именно – с MLP и сверточными нейронными сетями (CNN, СНС). Собственно, модели СНС и применялись мной ранее для классификации цифр из набора MNIST.

Архитектура сети самая простая: по одной карте признаков в каждом слое свертки. Даже на таких простых архитектуре СНС и алгоритмах обучения сети, инициализации весовых коэффициентов наибольшее значение F-меры достигло 70%. Более подробно результаты изложены в статье собственного авторства (pdf-файл в папке 2 задания).

*Перспективные инструменты.* Работа выполнялась в конце 2018 года – когда мои навыки программирования на Python были еще низкими, а MATLAB уже тогда знал очень хорошо. Тогда еще не было соответствующего Toolbox’а в MATLAB (он вышел уже спустя пару месяцев под названием Deep Learning Toolbox, с оптимизированными и усовершенствованными алгоритмами).

И даже после этого, когда уже хорошо освоил Python и ML-библиотеки Keras, PyTorch, Scikit-Learn, результаты обучения даже базовых CNN-моделей на том же самом датасете MNIST уже превышали 99% (точность). А следовательно, если есть готовые фреймворки, то зачем оптимизировать свой? Тем более, что работа была скорее теоретическая и была направлена на более глубокое погружение в архитектуры сверточных сетей.

Однако если бы встал вопрос об улучшении фреймворка, то стоило бы улучшить, в первую очередь, алгоритм обучения СНС (добавить поддержку нескольких карт признаков в слое свертки, переменную скорость обучения, усовершенствованные методы генерации весовых коэффициентов), а также добавить такие features как пакетная нормализация, dropout, предобработка входа и т.д.

**2. Поиск соответствия 10-ти классам (mnist) (способ 2 – от 17.06.2022)**

Отличие от предыдущего способа заключается в применении другого датасета (изображения 32x32 против 28x28), других библиотек (Keras, Matplotlib, NumPy) и несколько измененной стратегии обучения. Для обучения используется та же самая архитектура сети – СНС, 3 слоя свертки/пулинга + MLP с одним скрытым слоем.

Стратегия обучения заключается в использовании GridSearch по различным параметрам сети, таких как кол-во карт признаков, размер фильтра (ядра свертки), тип пулинга (среднее/максимум) и размер скрытого слоя MLP. Длительность обучения – 50 эпох, размер пакета – 200, примерное соотношение обучающей и тестовой выборок – 5:1 (всего 60000 образцов). Мера оценки качества обучения моделей – F-мера (F1 score).

Всего было обучено 16 СНС; обучение произведено на ноутбуке с ОЗУ 8 Гб и 2-ядерным процессором с частотой 3,1 ГГц.

Наивысшее значение F-меры показала сеть с опцией пулинга «max», с максимальным числом ядер свертки среди представленных (4-6-10) и максимальным размером ядер свертки среди представленных (5x5 - 5x5 - 3x3). Оно составило 96,26% на тестовой выборке.

Для улучшения результата необходимо использовать СНС большего размера, ансамбли на их основе, а также применять другие техники машинного обучения как, например, увеличение данных. Не исключается переход и на другие архитектуры НС (примеры [здесь](https://ru.wikipedia.org/wiki/MNIST_(база_данных)#Сопоставление_систем)).

P.S. Во всех указанных случаях необходимо использование более высокопроизводительного «железа» (видеокарты Nvidia) и больше времени на изучение статей, программную реализацию процесса обучения сетей и сам процесс.