

СОЗДАНИЕ СЖАТЫХ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ ПРИЗНАКОВ ОБЪЕКТОВ ПО ОПТИЧЕСКИМ ИЗОБРАЖЕНИЯМ ИХ ПОВЕРХНОСТЕЙ

Сметанин А.А. Университет ИТМО

Першуткин А.Э. Университет ИТМО

Научный руководитель – д.т.н., доцент Духанов А. В.

Университет ИТМО

Исследования выполнены за счет финансирования университета ИТМО в рамках НИР №621311 «Разработка технологии формирования моделей машинного обучения малого объёма в целях распознавания признаков на изображениях с высокой точностью при сверхмалых выборках данных».

Аннотация

В данном исследовании предложен подход к созданию сжатых моделей машинного обучения для распознавания признаков по фото поверхностей объектов на основе выборок малого объема. При потере точности распознавания не более, чем 5%, удается получать модели машинного обучения, которые требуют в 3-4 раза меньше памяти и обеспечивают трех-пятикратное ускорение в соответствующих расчетах 3-5 раз. Экспериментальные исследования проведены в области распознавания заболеваний растительных культур при объемах выборки на один класс менее 50 изображений.

Ключевые слова: машинное обучение, квантизация, классификация признаков

Введение

Современные методы классификации признаков на изображении основаны на применении, прежде всего, сверточных нейронных сетей. Такие модели требуют, как правило ресурсоемкого обучения на большом количестве исходных данных. Более того, эти модели не всегда оптимизированы для вычислений на мобильных устройствах, так же, в силу ресурсоемкости. На данный момент не удалось найти метод, технологию по созданию сжатых моделей машинного обучения, обучаемых на малом наборе данных. Таким образом, возникает необходимость в разработке методов, технологии, обеспечивающих создание сжатых моделей машинного обучения, которые можно применять, в том числе, на мобильных устройствах. Поэтому целью работы является разработка технологии и библиотеки с открытым кодом на языке программирования Python для решения задач подготовки сжатых моделей машинного обучения.

Основная часть

Предлагаемая технология заключается в применении в процессе обучения сиамской модели (три сети-близнеца) с трехчленной функцией ошибки, а также сжатия (статической/динамической) сети-экстрактора признаков.

Предыдущие исследования показывают, что обучение модели с применением Transfer Learning и трехчленной функции ошибки Triplet Loss значительно повысило точность распознавания по сравнению с методами обучения, использующих одну сверточную сеть. Такая высокая точность (средняя величина - 96%) достигалась, в том числе, для набора данных с 20 классами на 1000 изображений (50 или менее изображений на один класс).

Но для того, чтобы модель можно было эффективно использовать на мобильных устройствах (включая отсутствие дискомфорта лага), требуются мероприятия по снижению ее размера и обеспечения расчетов на целых числах. Данные эффекты можно достигнуть путем квантизации. Здесь требуется убедиться не только в снижении размеров модели и ускорении расчетов, но в обеспечении приемлемой точности распознавания, которая не должна снизиться более, чем на 5%.

Чтобы обеспечить выполнение таких требований, авторами работы разработаны методы и программные библиотеки позволяющие выполнять статическую квантизацию сетей следующих архитектур: «MobileNet», «VGGNet», «ResNet», «EfficientNet», а также динамическую квантизацию сети архитектуры «MobileNet».

Экспериментальная часть проведена на основе серий изображений поверхности растительных культур: в среднем по 35 изображений на 12 заболеваний. В большинстве случаев точность снижается не более, чем на 2,5 процента. При этом выбор архитектуры сети влияет на данную точность распознавания разных заболеваний по-разному. Эти определяются будущие исследования.

Выводы.

Предложенная технология создания сжатых моделей машинного обучения позволяет удерживать точность распознавания признаков объектов в пределах пяти процентов по отношению аналогичного показателя для несжатых моделей. Данную технологию предполагается использовать для создания программных средств для мобильных устройств в целях выявления болезней растений на ранних стадиях. Также планируются эксперименты по распознаванию признаков по оптическим изображениям поверхностей объектов иных видов (пленки, территории и др.).

Сметанин А.А. (автор)

Подпись

Першуткин А.Э.(соавтор)

Подпись

Духанов А.В. (научный руководитель)

Подпись