

Предсказание размера набора данных для дообучения модели детекции
Московская Е.А. (Университет ИТМО)
Научный руководитель – сотрудник подразделения, Международная лаборатория
«Компьютерные технологии», Ефимова В.А.
(Университет ИТМО)

Введение. Обнаружение объектов на изображении — актуальная задача, которая находит применение во многих областях как промышленной, так и исследовательской деятельности. Зачастую для этого используется выложенная в открытый доступ модель, предобученная на распространенном наборе данных, например, ImageNet, которая затем дообучается на пользовательском наборе данных, специфичном для решения его задачи. Одним из важных параметров дообучения является размер обучающей выборки. Согласно законам масштабирования нейросетей, чем он больше, тем выше качество работы модели. Однако сбор и разметка набора данных — это, зачастую, затратный процесс, поэтому предпочтительным является наименьший размер набора данных, обеспечивающий желаемое качество работы модели. Расчет данного значения является целью данной работы.

Основная часть. В ходе работы был проведен анализ литературы и найдены следующие решения данной задачи:

1. Законы масштабирования нейросетей
2. Events per Variable rule of thumb
3. Learn-Optimize-Collect paradigm

Однако данные подходы либо не достаточно точны, либо требуют несколько итераций сбора и разметки данных, что может быть неудобно для пользователя.

Мы предлагаем предсказывать размер данных, необходимых для обучения, опираясь на типовое изображение из целевого набора данных. Для этого мы используем нейронную сеть, на вход которой подается вектор признаков типового изображения. Мы используем следующие признаки:

1. Color Histograms — дескриптор цвета, подсчитывающий количество пикселей для каждого уровня интенсивности цвета в данном изображении. Данный дескриптор не учитывает распределение цветов в пространстве, однако он позволяет сравнивать изображения, отталкиваясь от предположения, что изображения со сходным распределением цветов имеют сходные свойства и содержание.
2. Local Binary Patterns — дескриптор текстуры, который для каждого пикселя черно-белого изображения и его окрестности подсчитывает значение, характеризующее его относительную выпуклость.
3. Histogram of Oriented Gradients — дескриптор формы объектов на изображении, основывающийся на подсчете количества направлений градиента в локальных областях изображения.

Выводы. Проведен анализ подходов для оценки размера набора данных для дообучения и разработана нейронная сеть, предсказывающая размер набора данных на основании мета-признаков изображения.