

Курочкин И.И.
Институт проблем передачи информации им. А.А. Харкевича РАН
127051, г. Москва, Большой Каретный переулок, д.19 стр. 1
E-mail: kurochkin@iitp.ru

Классификация большого множества текстурных изображений методами машинного обучения

Задача классификации изображений является достаточно распространенным типом задач. Для ее решения часто используются методы машинного обучения и, в частности, глубокие нейронные сети. Решение задач классификации изображений необходимо в различных областях от медицины и распознавания материалов, до анализа спутниковых снимков. Одним из преимуществ использования глубоких нейронных сетей в задаче классификации является возможность работы с большим количеством классов. Но это накладывает существенные ограничения на формирование датасета – множества изображений, которое используется для обучения глубокой нейронной сети. Для решения прикладных задач очень часто формируются достаточно большие множества изображений – от нескольких тысяч до сотен миллионов изображений.

Популярность использования глубоких нейронных сетей для решения широкого круга задач привело к возникновению большого количества архитектур нейронных сетей, разработанных специально для решения определенного вида прикладных задач. Одним из популярных видов изображений являются текстурные изображения, на которых представлены одна или несколько видов текстур. В отличие от изображений объектов, текстуры могут не иметь четких границ, а их форма может не нести дополнительной информации. Задачи обработки и анализа текстурных изображений часто возникают при обработке данных, полученных с помощью методов микроскопии. Для решения некоторых задач классификации текстур могут использоваться глубокие нейронные сети с небольшим числом слоев свертки. Это связано с отсутствием необходимости выделять высокоуровневые признаки, которые могут присутствовать в больших объектах, но их не будет в изображениях текстур. В качестве примера, можно привести глубокую нейронную сеть T-CNN [1] для классификации текстур.

Далее будет рассматриваться задача классификации изображений текстур, полученных с помощью оптического микроскопа. Изначально были получены порядка 650 исходных изображений размера 2272x1704 точек в градациях серого.

Опишем порядок формирования текстурного датасета для обучения нейронной сети. Пусть есть множество сходных изображений размером 2272x1704, каждое из которых имеет несколько числовых и категориальных признаков. При этом количество классов может варьироваться от 2 до 36, в зависимости от используемых признаков. Пусть глубокая нейронная сеть с архитектурой T-CNN имеет входной слой размерностью 256x256, тогда из каждого изображения 2272x1704 получится $9*7=63$ непересекающихся фрагмента, при 50% наложении фрагментов получится $17*13=221$ фрагмент. Так как ориентация текстуры на изображении неизвестна, то каждый фрагмент нужно представить в виде множества фрагментов повернутых на различные углы: при повороте одного фрагмента изображения на угол от 0° до 345° с шагом 15° получается 24 фрагмента. В результате, датасет для обучения нейронной сети сформированный из 650 исходных изображений будет состоять из 3 447 600 фрагментов изображений.

Обучение даже небольшой глубокой нейронной сети с архитектурой T-CNN на датасете размера порядка 3.5 миллиона изображений является вычислительно сложной задачей. Такой объем вычислений провести за разумное время возможно только при использовании высокопроизводительных кластерных систем или распределенных систем. Для реализации обучения глубокой нейронной сети на распределенной системе нужно реализовать один из подходов к распределенному глубокому обучению [2]: синхронный, асинхронный или децентрализованный. При использовании грид-системы на базе платформы BOINC лучше использовать синхронный или асинхронный подход с параметрическим сервером [3]. В процессе распределенного глубокого обучения на параметрическом сервере будет находиться

актуальная глобальная модель нейронной сети. Процесс обучения будет проходить итеративно, от использования того или иного подхода будет меняться методика обновления глобальной модели с помощью результатов обучения локальных моделей после каждой итерации.

Однако при реализации любого из подходов к распределенному обучению глубокой нейронной сети будет необходимо разделять исходный датасет на части. При этом на вычислительном узле грид-системы локальная модель глубокой нейронной сети может быть обучена на одной итерации только на небольшом количестве (~1000) изображений. Если предположить, что грид-система имеет 200 вычислительных узлов, то для участия всех изображений датасета понадобится порядка 18 итераций. Это возможно в случае последовательного формирования локальных датасетов перед каждой итерацией. При использовании методики случайного формирования локальных датасетов количество итераций увеличивается. Данные оценки позволяют сделать вывод, что общее количество итераций глобальной модели обучения нейронной сети будет не менее 500-1000.

Формирование локальных датасетов может происходить несколькими способами:

- Стохастическое (случайное) формирование локального датасета;
- Разделение датасета на части (статическое) до конца обучения;
- Разделение датасета на части (динамическое) после каждой итерации;
- Гибридное формирование датасета с использованием как случайного выбора, так наличия части изображений из локального датасета прошлой итерации.

Стоит заметить, что при любом способе формирования локального датасета должен соблюдаться баланс классов. Динамическое разделение датасета между узлами предполагает добавление новых данных на каждой итерации. Схему формирования локального датасета при динамическом разделении можно увидеть на рис. 1.

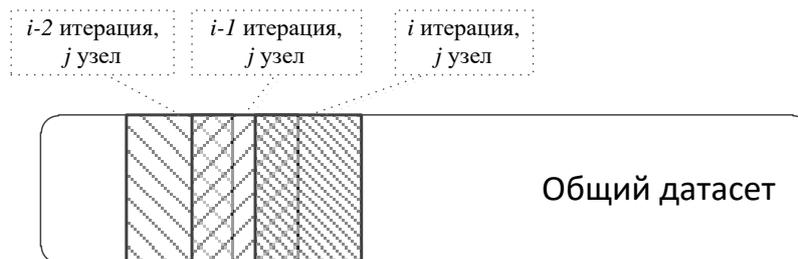


Рис.1. Формирование локального датасета при динамическом разделении.

В случае применения способа гибридного формирования в локальном датасете присутствует как динамическая часть, так и случайная часть (рис. 2).

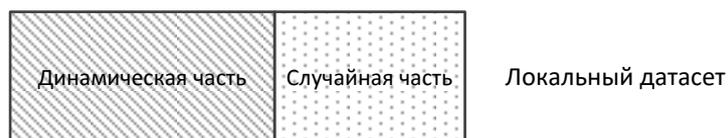


Рис.2. Части локального датасета при гибридном подходе.

Гибридный способ формирования локального датасета позволяет с одной стороны задействовать весь общий датасет, а с другой позволяет не изменять полностью локальный датасет между соседними итерациями обучения. При применении гибридно способа скорость обучения возрастает.

Литература

1. Andrearczyk, Vincent, and Paul F. Whelan "Using filter banks in convolutional neural networks for texture classification." // Pattern Recognition Letters, Vol.84, 2016. pp.63-69.
2. Ben-Nun, Tal, and Torsten Hoefler. "Demystifying parallel and distributed deep learning: An in-depth concurrency analysis." // ACM Computing Surveys (CSUR) Vol.52(4), 2019. pp.1-43.

3. Kurochkin I.I., Kostylev I.S. Solving the problem of texture images classification using synchronous distributed deep learning on desktop grid system // *Communications in Computer and Information Science*, Vol 1331, Springer, Cham, 2020. pp. 647-657.