

Рецензия

магистерской диссертации Патина Михаила Владиславовича
«Автоматизация подготовки обучающего набора для классификации похожих изображений»

Нейронные сети – эффективное средство анализа больших объёмов данных. В диссертации Патина М.В. исследуется возможность автоматизации этапа обучения нейронной сети. Этап обучения обычно долгий, дорогостоящий и трудоёмкий, поэтому задача его автоматизации является актуальной. Описанный в диссертации подход состоит в поиске особых точек изображения, вычислении дескрипторов этих особых точек, построении гистограмм дескрипторов изображения и классификации полученных гистограмм.

В первой главе диссертации приводится математическое описание применяемых методов. Дается определение особых точек и дескрипторов. Подробно описаны методы поиска особых точек и вычисления дескрипторов: ORB, BRISK и AKAZE. По аналогии с алгоритмами анализа текста создаётся словарь; вместо слов используются дескрипторы. Выполняется кластеризация массива дескрипторов обучающей выборки модифицированным методом k-средних. Теперь для каждого изображения можно построить гистограмму совпадений со словарными дескрипторами. Полученные гистограммы подлежат классификации, для чего применяется метод опорных векторов.

Вторая глава диссертации рассказывает о программной реализации методов. Приводится высокоуровневая структура приложения и схема базы данных. Исходный код приложения написан на Java. Задействуются алгоритмы библиотеки OpenCV, специфические для предметной области. База данных спроектирована в СУБД MySQL. Пользовательский интерфейс приложения позволяет выбрать алгоритм поиска особых точек и вычисления дескрипторов, параметры сжатия изображения и размер словаря. Текст проиллюстрирован примерами элементов пользовательского интерфейса.

Третья глава описывает практическое исследование. Автор экспериментировал с изображениями, объединёнными общей темой, например, с фотографиями разных городов, разных гор и т.п. В другом эксперименте исследовались различные фотографии одного объекта – например, Эрмитажа. В первом случае средняя точность классификации не превысила 80% для лучшего из методов (BRISK). Во втором эксперименте – поиске объекта на различных его изображениях – все исследованные методы показали отличные результаты, причём метод BRISK распознал все объекты со стопроцентной точностью.

Можно констатировать, что поставленная в диссертационной работе задача – автоматизация построения обучающей выборки – выполнена не полностью, т.к. не удаётся избежать этапа ручной проверки результатов и коррекции ошибок метода, число которых не ниже 20%. Остаётся неясным, насколько в результате облегчается ручной труд. В то же время достигнуты впечатляющие результаты в задаче поиска известного шаблона. К сожалению, автор рассматривает эти результаты как побочные и не описывает их достаточно подробно.

Не могу согласиться с автором в вопросе оценки влияния размера словаря на точность распознавания. На мой взгляд, флуктуации точности в пределах нескольких процентов не имеют решающего значения, а линейный рост затрат времени на формирование словаря (более 6 часов в одном из экспериментов) будет важен при практическом применении метода.

Считаю, что магистерская диссертация Патина М.В. заслуживает оценки
«Хорошо»

Рецензент,
к.т.н., ст.н.с. ООО «НСН»



Епифанов Н.А.