

Рецензия

магистерской диссертации Доржиева Зоригто Жаргаловича
«Анализ аудиоданных и распознавание событий для систем безопасности»

Диссертация Доржиева З.Ж. посвящена исследованию проблемы распознавания звуковых событий. В рамках данной работы ставится цель разработки и проверки работоспособности алгоритма для классификации аудиоинформации. Современные системы безопасности имеют возможность записывать не только видео, но и аудио формат данных. Аудиосигнал может содержать информацию, которая будет помогать работе различных приложений: умным домам, системам анализа городских звуков, программам мониторинга в области здравоохранения и т.д. Для достижения поставленной цели Автор формулирует следующие задачи: поиск репрезентативного набора данных, извлечение характеристик аудиофайлов, применение и сравнение алгоритмов классификации и, наконец, перенос знаний модели, обученной на большом наборе данных.

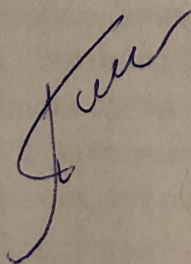
Содержательная часть диссертации состоит из введения, постановки задачи, литературного обзора, четырех глав и заключения. В первой главе описывается процесс подготовки данных, необходимых для обучения и тестирования классификаторов. После перечисления различных датасетов Автор останавливает свой выбор на открытой базе AudioSet от компании Google для предварительного обучения моделей и открытой базе Urban Sound Classification - для переноса знаний и тестирования. Вторая глава посвящена обзору существующих методов вычисления характеристик аудио сигнала и библиотеке pyAudioAnalysis, которая используется в данной работе в качестве модуля извлечения признаков. В третьей главе приводится описание различных классификаторов аудио признаков. В первой части главы Автор показывает предельную точность классических методов классификации на основе k-ближайших соседей и машины опорных векторов. Во второй части описывается нейросетевой подход на базе нейронных архитектур VGG-11 и VGG-16. Для сравнения с VGG Автор моделирует неглубокую сверточную нейронную сеть, состоящую из нескольких слоев. Также особое внимание уделено описанию методов регуляризации нейронных сетей, аугментации данных и процессу переноса обучения между моделями. Последняя глава посвящена тестированию обученных моделей и анализу результатов. В заключении Автор делает выводы относительно оптимизации моделей и их дальнейшего развития.

Замечания по диссертации:

- 1) Раздел, посвященный обзору литературы, включает в себя всего три работы, одна из которых посвящена программной библиотеке. В заключении Автор делает правильные с точки зрения рецензента выводы о том, что для увеличения точности имеет смысл обратить внимание на модели с меньшим количеством параметров. Однако, более полный и подробный анализ публикаций наверняка помог бы Автору сделать такие выводы в самом начале работы, обратив внимание на достаточно хорошие показатели моделей ResNet и SqueezeNet - лидеров на AudoSet. В целом, рецензента смущает выбор VGG нейронных сетей для практических целей в области классификации звуков, поскольку данные архитектуры тяжелее на порядки и менее точны по сравнению со state-of-art моделями даже в своей изначальной области - в области классификации изображений.
- 2) На протяжении работы Автор использует метрику "Точность", однако явно не раскрывает детали ее реализации, в частности, является ли она сбалансированной. Рецензент также обратил внимание, что в публикациях про Google AudioSet в основном применяются метрики mAP (mean Average Precision) и AUC, что также затрудняет сопоставление результатов данной работы с другими опубликованными результатами. В целом, ввиду отсутствия точного протокола тестирования для Urban Sound Classification сравнение результатов, полученных на ее случайных разбиениях и различных фолдах, кажется существенно менее репрезентативным, чем на AudioSet.

Несмотря на данные замечания, работа может быть оценена положительно. Несомненным достоинством диссертации является подробное описание процесса обучения выбранных классификаторов, метода переноса знаний в задаче классификации аудио и высокие показатели точности системы на тестовой базе. Считаю, что магистерская диссертация Доржиева З.Ж. заслуживает оценки «Хорошо».

Рецензент,
к.т.н., руководитель группы,
ООО "ДСП Лабс"



Тимошенко Д.М.