

ДЕТЕКТИРОВАНИЕ АКУСТИЧЕСКИХ СОБЫТИЙ В БЫТОВЫХ УСЛОВИЯХ

Тимофеева Е.П.

(Университет ИТМО)

Научный руководитель – к.т.н. Шуранов Е.В.

(Университет ИТМО)

В настоящее время продолжают активно развиваться технологии анализа речевых сигналов, классификации акустических событий, фильтрации, а также голосовые интерфейсы человеко-машинного взаимодействия, решающие задачи распознавания речи и голосовой биометрии. Для соответствующих систем важную роль играют детекторы акустических событий. На сегодняшний день известно множество методов детектирования речи, однако, разработка новых подходов по-прежнему сохраняет свою актуальность.

Чаще всего для решения задач детектирования применяют сверточные рекуррентные нейронные сети. Так как сверточные нейронные сети используются для распознавания образов, однако в этом случае невозможно учитывать влияние паттернов друг на друга. Для этого придётся анализировать полученные паттерны другой сетью и с другой архитектурой – например, рекуррентной нейронной сетью. Рекуррентные нейронные сети чаще используются для анализа таких данных как текст или аудиофайлы, где структура данных предполагает зависимости данных от значений друг друга. То есть нейронная сеть анализирует и запоминает такие зависимости, обучаясь находить закономерности в них. Объединение этих нейронных сетей даст устойчивость модели и преимущество в поиске закономерностей при обучении. Такая модель эффективна при классификации и при расставлении меток.

В данной работе рассматривается один из методов детектирования, предложенный участниками конкурса DCASE 2019, в основе которого лежит применение рекуррентных сверточных нейронных сетей, применение метода log-mel energies и некоторые оптимизации системы. Задача конкурса заключается в детектировании акустических событий в бытовых условиях с использованием слабо маркированных данных (без отметок времени). Целью систем является предоставление не только класса события, но и границ времени события, учитывая, что в аудиозаписи могут присутствовать несколько событий, а также, что эти события могут пересекаться. Так как есть пересекающиеся события, участниками были предложены некоторые улучшения. В основном это оптимизация постобработки данных.

В дальнейшем для решения задачи детектирования акустических событий будут рассмотрены методы, предложенные участниками конкурса, а также будет проделана работа по поиску улучшений этой системы.