

Детектирование акустических событий

Н.А. Большаков

Университет ИТМО, г. Санкт-Петербург

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

ООО «Группа компаний ЦРТ», г. Санкт-Петербург

Исследования выполнены за счёт стартового финансирования университета ИТМО в рамках НИР № 618278 «Синтез эмоциональной речи на основе генеративных состязательных сетей»

Одним из важнейших этапов при разработке систем детектирования акустических событий является ознакомление с предметной областью и методами используемыми при решении прикладных задач. Естественно, что как и в любой другой научной работе, в каждом случае требуется определённый индивидуальный подход, ввиду разнообразия возможных изначальных условий и анализируемых данных. Однако, изучение методов, уже активно применяемых при детектировании и классификации, позволяет накопить ценный опыт и определённую сравнительную базу, весьма полезные для проведения собственных исследований. В процессе обучения, крайне важно охватить как можно больше необходимых навыков, максимально подготовить специалиста к наиболее вероятным задачам, которые так или иначе возникнут в его научной работе.

Детектирование акустических событий - это направление исследований связанное с изучением методов определения времени, длительности и типа звукового явления. Например, одной из подзадач детектирования можно выделить анализирование аудио файла с целью классификации акустической сцены [1]. Прикладным решениям, связанным с данными исследованиями, находится применение в самых разных областях человеческой жизни - дефектоскопия; исследования гидросферы; вопросы безопасности на производстве, защиты частной и интеллектуальной собственности [2]. Таким образом, с течением времени, появляются всё новые и новые способы решения описанных задач, и одной из относительно недавних, но уже распространённых практик является использование искусственных нейронных сетей, для классификации акустических событий.

Искусственные нейронные сети - принцип построения математической модели, основанный на подражании биологическим нейронным сетям, впервые рассмотренный ещё в сороковые годы 20 века [3]. Наибольшее развитие идея получила с появлением в 2007 году концепции глубокого обучения многослойных нейронных сетей.

В этой статье мы рассматриваем несколько способов построения нейронных сетей, предложенных в 2018 году командой из Констанцкого университета [4], для последующего обучения и классификации акустических событий. В этой работе подробно рассматриваются различные модели, а также предлагаются особенности их реализации на языке программирования Python с использованием открытой нейросетевой библиотеки Keras. В частности, мы детально описываем способ построения нейронной сети, основывающийся на использовании нескольких высокоуровневых, предобученных нейронных сетей.

Данная работа выполнена в рамках НИР «Детектирование акустических событий».

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ И ЛИТЕРАТУРЫ.

1. Статья «Учим компьютер различать звуки: знакомство с конкурсом DCASE и сборка своего аудио классификатора за 30 минут» [Электронный ресурс] : Хабр — коллективный блог об IT - Электрон. Текстовые дан. - URL: <https://habr.com/ru/company/speechpro/blog/437818/> (дата доступа: 27.02.2019 г.).
2. Создание эффективной системы безопасности на объектах промышленного и гражданского назначения. [Электронный ресурс] : ООО «Группа компаний ЦРТ» — официальный сайт - Электрон. Текстовые дан. - URL:

https://www.speechpro.ru/upload/solution/pdf/smart_tracker_frs_sluzhby_bezопасности_3026.pdf (дата доступа: 27.02.2019 г.).

3. Мак-Каллок У. С., Питтс В. Логическое исчисление идей, относящихся к нервной активности Архивная копия от 27 ноября 2007 на Wayback Machine // Автоматы / Под ред. К. Э. Шеннона и Дж. Маккарти. — М.: Изд-во иностр. лит., 1956. — С. 363—384.

(Перевод английской статьи 1943 г.)

4. Technical Report «COMBINING HIGH-LEVEL FEATURES OF RAW AUDIO AND SPECTROGRAMS FOR AUDIO TAGGING» [Электронный ресурс] : D-Case community - Электрон. Текстовые дан. - URL:

http://dcase.community/documents/challenge2018/technical_reports/DCASE2018_Wilhelm_109.pdf (дата доступа: 27.02.2019 г.).