



А.И. Максимов, Д.П. Тимошкин

ИССЛЕДОВАНИЕ МЕТОДОВ ПРЕДОБРАБОТКИ АУДИОЗАПИСЕЙ ДЛЯ ЗАДАЧИ ПРОВЕРКИ ИХ ПОДЛИННОСТИ

(Самарский университет)

Введение

Использование результатов звукозаписи в качестве доказательства в суде на сегодняшний день является обыденной практикой. Соответственно, актуальной является задача установления их подлинности в рамках криминалистической экспертизы.

В данной работе исследуется один из способов определения подлинности фонограмм [1] – основанный на анализе фоновых звуков. Такой метод установления подлинности актуален для случая, при котором злоумышленник подменяет часть фонограммы фрагментом с голосом того же человека, но из другой записи. Подобная задача может быть решена при помощи нейронных сетей [2]. В работе исследуется эффективность применения различных методов предобработки фрагментов исследуемой аудиозаписи перед передачей на вход нейросетевой модели.

Используемые в исследовании методы

Упростим для исследования постановку задачи с определения вставок в фонограмму до сравнения условий записи фрагментов аудиозаписи. Таким образом, в качестве исходных данных имеются пары фрагментов фонограммы без вокальной составляющей.

В ходе исследования определялось евклидово расстояние между векторами признаков, сформированных при помощи предобученной сети DenseNet [3]. Изначально данная модель разрабатывалась для классификации RGB изображений, однако, как показано в работе [4], ее можно использовать и для классификации аудиозаписей. Для получения векторов признаков из сети исключались последние слои, отвечающие за классификацию. Архитектура данной сети представлена на рисунке 1.

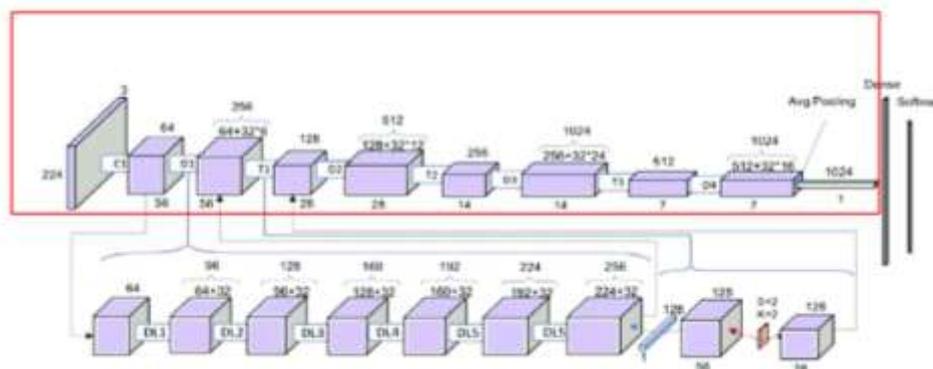


Рис. 1. Архитектура сети DenseNet.

Рамкой выделены слои, использованные в исследовании



Для применения данной модели к обработке аудиозаписей, необходимо произвести предварительную обработку, в ее качестве хорошо зарекомендовало себя построение мел-спектрограммы. В данной работе исследованы два метода предобработки аудиозаписей с построением мел-спектрограмм:

1. Предлагаемый в работе [5] метод с репликацией мел-спектрограммы, полученной при постоянных параметрах оконного преобразования Фурье [6], на все три канала результирующего псевдо-RGB-изображения;
2. Предлагаемый в работе [4] метод с построением мел-спектрограмм с различными параметрами оконного преобразования Фурье для каждого канала результирующего псевдо-изображения.

Примеры предобработанных фрагментов фоновых звуков исследуемыми методами представлены на рисунке 2.

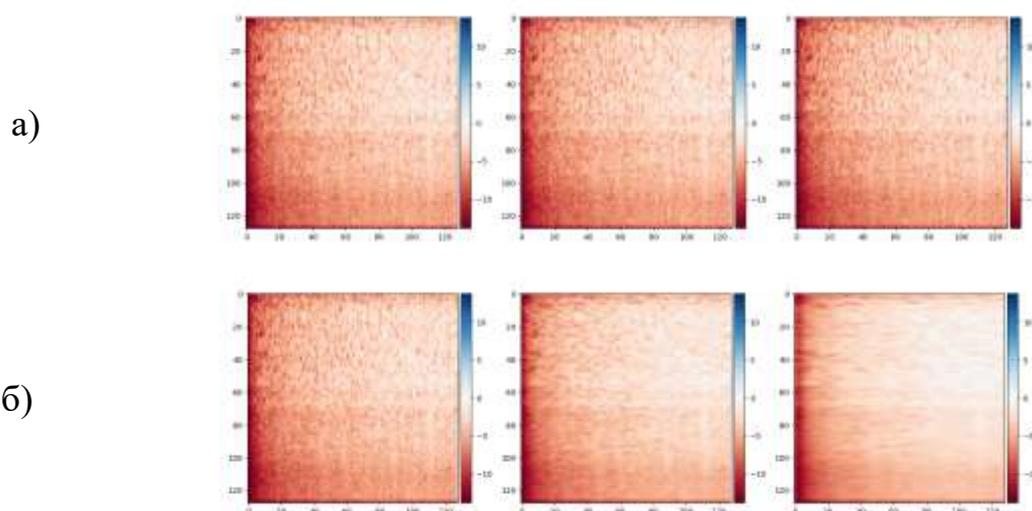


Рис. 2. Примеры предобработанного фрагмента фонограммы а) методом [5], б) методом [4]

В указанных работах разработаны методы для классификации широкого спектра аудиозаписей. Цель данной работы – проверить применимость данных подходов к более узкой и специфической задаче – криминалистической экспертизе фоновых звуков фонограмм.

Исследование методов предобработки аудиофрагментов

В ходе экспериментального исследования фрагменты фонограмм проходили предобработку одним из описанных выше методов, после чего поступали на вход нейронной сети. Полученные в результате векторы признаков объединялись в пары, после чего считалась мера близости векторов в паре. В качестве меры близости использовалось евклидово расстояние. Визуализация полученных результатов показана на рисунке 3.

Как видно на рисунке 3, для обоих методов предобработки фонограмм множество два множества пар являются линейно разделимыми. Это означает, что оба данных метода применимы для задачи выявления вставок в фонограммы. В таблице 1 представлена усредненная мера близости векторов для исследованных методов предобработки и типа пар.



Рис. 3. Результаты расчета мер близости пар векторов признаков при предобработке а) методом [5], б) методом [4]. Зеленым отмечены меры пар, составленных из фрагментов одной аудиозаписи, красным – из фрагментов различных

Таблица 1 – Исследуемые в работе методы обработки аудиозаписей

№	Пары изображений	Среднее расстояние	Модуль разности
1	Метод [5], пары из одинаковых фонограмм	9.3038	16.7444
2	Метод [5], пары из разных фонограмм	26.0482	
3	Метод [4], пары из одинаковых фонограмм	9.1117	17.4782
4	Метод [4], пары из разных фонограмм	26.5901	

Как видно из приведенных в таблице значений, модуль разности между средним значением для одинаковых и среднего значения для разных пар при использовании метода [4] больше.

Заключение

В работе исследована часть задачи определения подлинности фонограмм по фоновым звукам - исследование методов предобработки аудиозаписей для определения их меры близости. В результате экспериментальных исследований выяснилось, что оба метода предобработки подходят для решаемой задачи, однако метод, использующий построение мел-спектрограмм с различными параметрами оконного преобразования Фурье показал наилучший результат.

Литература

1. Каганов А.Ш. Криминалистическая экспертиза звукозаписей / А.Ш. Каганов. — М.: Юрлитинформ, 2005. — 272 с.
2. Лебедева Д.С. Исследование методов диагностики фальсификации фонограмм путем сравнения фрагментов фоновых шумов / Д.С. Лебедева, А.И. Максимов // Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов имени Е.В. Арменского. — 2021. — С. 244-246.



3. Huang, G. Densely Connected Convolutional Networks / G. Huang, Z. Liu, L. van der Maaten, K.Q. Weinberger // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) — 2017. — P. 4700-4708.
4. Palanisamy K. Rethinking CNN Models for Audio Classification / K. Palanisamy, D. Singhania, A. Yao // ArXiv — 2020. — Vol. abs/2007.11154.
5. Guzhov A. ESResNet: Environmental Sound Classification Based on Visual Domain Models / A. Guzhov, F. Raue, J. Hees, A. Dengel // 25th International Conference on Pattern Recognition (ICPR) — 2020. — P. 4933-4940.
6. L. Wyse. Audio Spectrogram Representations for Processing with Convolutional Neural Networks // Proceedings of the First International Workshop on Deep Learning and Music joint with IJCNN. — 2017. — Vol.1. — No.1. — P. 37-41.

А.А. Медведева, В.А. Соболев

ПЕРСОНАЛИЗАЦИЯ ПРЕДЛОЖЕНИЙ НА ОСНОВЕ АБОНЕНТОВ «МЕГАФОН»

(Самарский государственный технический университет)

Цель проекта: разработать модель, которая сможет формировать персональные предложения абонентам «МегаФона».

Задача проекта: Разработка модели, которая принимает файл test.csv из корневой папки и записывает в эту же папку файл answers_test.csv. В этом файле должны находиться четыре столбца: id, vas_id, buy_time и target.

Актуальность проекта

Искусственный интеллект проникает во все сферы бизнеса и индустрий. Теоретические знания в области машинного обучения и нейронных сетей удастся эффективно применять на практике для решения прикладных бизнес-задач. Огромное количество данных о своих клиентах дает телеком-компания «Мегафон» возможность использовать методы искусственного интеллекта для создания дополнительных ценностей и получения конкурентного преимущества

Основные этапы работы

Чтобы решить данную задачу надо сделать правильную оценку и выбор модели. На данный момент есть множество доступных моделей машинного обучения, мы решили попробовать в работе несколько и посмотреть, какой на практике будет лучше.

Для сравнения мы взяли 4 метода:

- Логистическая регрессия
- Случайный лес
- Градиентный бустинг
- Экстремальное усиление градиента

После исследования выяснилось что метод экстремального усиление градиента работает эффективнее, так как точность его больше.