УНИВЕРСИТЕТ ИТМО

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
июль–август 2022 Том 22 № 4 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
Jule–August 2022 Vol. 22 No 4 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК УНИВЕРСИТЕТ ИТМО
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

doi: 10.17586/2226-1494-2022-22-4-708-715
УДК 004.523

# Применение FN-корректора с целью повышения качества классификации аудиособытий

**Александр Михайлович Голубков[1]✉, Евгений Витальевич Шуранов[2]**

[1] Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), Санкт-Петербург, 197022, Российская Федерация
[1,2] ООО «ТЕХКОМПАНИЯ ХУАВЭЙ», Москва, 123007, Российская Федерация
[2] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

[1] kremnikov@gmail.com✉, https://orcid.org/0000-0002-8330-1823
[2] e_v_shuranov@mail.ru, https://orcid.org/0000-0003-0977-5075

**Аннотация**

**Предмет исследования.** Рассмотрена проблема классификации акустических событий, активно применяемых в решениях задач безопасного города, умного дома, IoT устройств, а также для детектирования опасных ситуаций на производстве. Предложено решение повышения точности классификаторов без изменения их структуры и сбора дополнительных данных. Основным источником данных для экспериментов выбран открытый набор данных TUT Urban Acoustic Scenes 2018, Development Dataset. **Метод.** Предложен способ увеличения точности классификации аудиособытий с помощью использования FN-корректора. FN-корректор представляет собой линейный классификатор и работает в два этапа: преобразование пространства признаков в линейно-разделимое пространство и линейное отделение одного класса от другого. В случае применения корректора классы — типы ответов исходного классификатора: положительный (P), отрицательный (N), ложноположительный (FP) и ложноотрицательный (FN). В результате возможно обучить два типа корректоров FP и FN, которые работают как бинарные линейные классификаторы и разделяют ответы на положительные/ложноположительные и отрицательные/ложноотрицательные соответственно. Выполнены эксперименты, где в качестве исходного классификатора использована сверточная нейронная сеть VGGish. Аудиосигнал преобразован в спектрограмму и передан на вход нейронной сети, которая формирует признаковое описание спектрограммы и производит классификацию. **Основные результаты.** В качестве примера демонстрации повышения точности классификации выбраны два «спутанных» класса. С помощью признакового описания аудиозаписей этих классов построен, обучен FN-корректор и подключен к исходному классификатору. Ответ от классификатора, а также признаковое описание передано на вход корректора. Далее корректор переводит пространство признаков в новый базис (в линейно разделимое пространство) и классифицирует ответ классификатора, таким образом «отвечает» на вопрос, ошибается ли исходный классификатор на таком векторе признаков или нет. Если исходный классификатор ошибся, то его ответ изменяется корректором на противоположный. Если нет — ответ остается тем же самым. Результаты экспериментов продемонстрировали снижение уровня спутывания классов и, соответственно, увеличение точности исходного классификатора без изменения его структуры и без сбора дополнительного набора данных. **Практическая значимость.** Полученные результаты могут быть использованы на устройствах IoT, имеющих существенные ограничения по размеру используемых моделей, а также при решении проблем доменной адаптации, актуальной в задачах аудиоаналитики.

**Ключевые слова**

классификация аудособытий, обработка звука, FN-корректор, корректор ложноотрицательных ответов, сверточные сети, аудиоаналитика

**Благодарности**

Работа выполнена в рамках исследований, поддерживаемых СПбГЭТУ «ЛЭТИ» им В.И. Ульянова (Ленина).

**Ссылка для цитирования:** Голубков А.М., Шуранов Е.В. Применение FN-корректора с целью повышения качества классификации аудиособытий // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 4. С. 708–715. doi: 10.17586/2226-1494-2022-22-4-708-715

# Applying the FN-Corrector to improve the quality of audio event classification

**Alexander M. Golubkov[1]✉, Evgeny V. Shuranov[2]**

[1] Saint Petersburg Electrotechnical University "LETI", Saint Petersburg, 197022, Russian Federation
[1,2] Huawei, Moscow, 123007, Russian Federation
[2] ITMO University, Saint Petersburg, 197101, Russian Federation

[1] kremnikov@gmail.com✉, https://orcid.org/0000-0002-8330-1823
[2] e_v_shuranov@mail.ru, https://orcid.org/0000-0003-0977-5075

**Abstract**

The paper deals with the problem of acoustic events classification which is actively applied to the problems of a safe city, smart home, IoT devices, and for the detection of industrial accidence. A solution to improve the accuracy of classifiers without changing their structure and collecting additional data is proposed. The main data source for the experiments was the TUT Urban Acoustic Scenes 2018, Development Dataset. The paper presents the way to increase the accuracy of audio event classification by using the FN-corrector. The FN-corrector is a linear two-stage classifier performing the transformation of the feature space into a linearly separable space and the linear separation of one class from another. If a corrector is applied, the responses of the original classifier generate four classes: positive (P), negative (N), false positive (FP), and false negative (FN). As a result, it becomes possible to train two types of correctors: the FP-corrector separating positive and false positive classifier responses, and the FN-corrector separating negative and false negative classifier responses. In the experiments, the VGGish convolutional neural network was used as the initial classifier. The audio signal is converted into a spectrogram and is fed to the input of the neural network which forms the spectrogram feature description and performs a classification. As an example, two "confused" classes are selected to demonstrate the increase in classification accuracy. Using the feature description of audio recordings of these classes, an FN-corrector was built, trained and connected to the original classifier. The response from the classifier, as well as the feature description, has been passed to the corrector input. Next, the corrector translated the feature space into a new basis (into a linearly separable space) and classified the classifier answer responding to the question whether the original classifier makes a mistake on such a feature vector or not. If the original classifier made a mistake, then his answer is changed by the corrector to the opposite, otherwise the answer remains the same. The results of the experiments demonstrated a decrease in the level of class confusion and, accordingly, an increase in the accuracy of the original classifier without changing its structure and without collecting an additional data set. The results obtained can be used on IoT devices that have significant limitations on the size of the models used, as well as in solving the problems of domain adaptation which is relevant in audio analytics.

**Keywords**

acoustic event detection, audio processing, FN-corrector, false negative corrector, DSP, CNN, convolutional neural network, audio analytics

**Aknowledgements**

The work was carried out as part of research supported by LETI.

**For citation:** Golubkov A.M., Shuranov E.V. Applying the FN-corrector to improve the quality of audio event classification. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 4, pp. 708–715 (in Russian). doi: 10.17586/2226-1494-2022-22-4-708-715

## Введение

Решение задачи классификации акустических событий заключается в определении события, которое происходит на анализируемой аудиозаписи. Для решения данной задачи применяются нейронные сети разных архитектур с использованием разнообразных признаковых представлений. При этом применяются записи сигнала от набора амплитуд до модификаций MFCC-коэффициентов (Mel-Frequency Cepstral Coefficients, мел-частотные кепстральные коэффициенты) и предобученных признаков (OpenL3 [1] и i-vectors (вектора представления голоса человека)), в случае если требуется, кроме классификации, решить сопутствующие задачи сегментации или локализации [2].

Актуальной задачей является улучшение точности существующих алгоритмов. Регулярные конкурсы по улучшению точности при различных условиях и ограничениях активно поддерживаются как академическими, так и индустриальными исследовательскими командами[1]. Цель настоящей работы — получение способа повышения точности классификации на стадии постпроцессинга, что позволит применить данный способ для широкого класса решений. Предложено использовать FN-корректор (False Negative корректор, т. е. корректор, исправляющий ложноотрицательные ошибки) [3, 4]. Модель корректора основана на использовании каскадной редукции размерности признаков и позволяет за линейное время скорректировать ошибочный ответ классификатора.

## Способ применения метода каскадной редукции к задаче классификации аудиособытий

Метод каскадной редукции пространства признаков заключается в последовательном уменьшении размер-

[1] Detection and Classification of Acoustic Scenes and Events. [Электронный ресурс]. URL: https://dcase.community (дата обращения: 24.06.2022).

ности и преобразовании пространства признаковых векторов. В этом случае независимо и случайно выбранный вектор одного класса отделяется от независимо и случайно выбранного вектора другого класса линейным дискриминантом (например, линейным дискриминантом Фишера). В основе метода представлена модель системы, способная исправить (скорректировать) ответы существующего классификатора. Принцип работы системы заключается в переразметке исходного набора данных с разметки «объект, класс» на разметку «признаковое описание объекта, тип ответа классификатора». При этом способ построения векторов признаков никак не фиксируется и, следовательно, качество работы системы напрямую зависит от метода получения признаковых векторов. Результат будет точнее, если: больше информации об объекте сохранено в его векторе признаков, и более выраженной кластерной структурой обладают вектора в пространстве признаков.

В работе представлен способ применения системы для решения задачи улучшения точности классификации аудиособытий. При этом выбраны следующие условия: в качестве объекта классификации — мел-спектрограмма аудиозаписи; в качестве способа получения вектора признаков – предобученная сверточная нейронная сеть VGGish; в качестве исходного классификатора — последовательно подключенные к сети VGGish полносвязный слой и слой Softmax, выходом которого является распределение вероятностей принадлежности спектрограммы каждому из классов.

Предлагаемый способ улучшения точности исходного классификатора предполагает следующие шаги.

Шаг 1. Выделение классов, на которых исходный классификатор допускает ошибки, и определение типа этих ошибок с помощью тренировочного набора данных.

Шаг 2. Построение спектрограммы и вектора признаков каждой аудиозаписи из тренировочного набора данных.

Шаг 3. Выбор типа ошибки, которую предполагается исправить (FN или FP) и дальнейшая переразметка тренировочного набора данных.

Шаг 4. Преобразование пространства признаков методом каскадной редукции таким образом, чтобы отделить те вектора признаков, на которых классификатор допускает ошибки (FN ошибки), от тех, на которых не допускает (т. е. правильные отрицательные ответы, N).

Шаг 5. Обучение линейного дискриминанта Фишера для проведения бинарной классификации в построенном на Шаге 4 пространстве.

Шаг 6. Подключение полученной системы коррекции к выходу исходного классификатора.

Шаг 7. Проведение процесса исправления ответа исходного классификатора: если система коррекции отнесла вектор признаков к классу ошибочных ответов, то исходный ответ корректора исправляется на противоположный. Если к классу правильных, то ответ остается тем же самым.

Использование предложенного способа исправления ответов исходного классификатора, в частности — FN-корректора и сверточной нейронной сети в качестве способа получения векторов признаков, показало значимый прирост точности в задаче классификации акустических событий.

## Практическая реализация способа увеличения точности классификации аудиособытий с помощью использования FN-корректора

Выберем набор данных TUT Urban Acoustic Scenes 2018[1], содержащий 10 классов акустических событий: *аэропорт, торговый центр, станция метро, пешеходный переход, городская площадь, городская улица, поездка в трамвае, поездка в автобусе, поездка в метро и городской парк*. Весь набор состоит из 8640 записей акустических событий, каждая из которых принадлежит одному из классов (закрытая задача). Каждый класс имеет 864 аудиозаписей. Суммарная длительность записей составляет 24 часа. Задача эксперимента — сопоставить каждую аудиозапись из набора соответствующему классу. Каждая запись может и должна принадлежать только одному классу.

**Подготовка аудиозаписей.** Для ускорения обучения нейронных сети и повышения ее точности аудиозапись разбита на отрезки по 100 мс, каждому из которых присвоен класс аудиозаписи. Каждая аудиозапись имеет два канала (с правого и левого микрофонов) с частотой дискретизации 16 кГц и разрядностью 16 бит. В настоящей работе использован метод аналого-цифровых преобразователей — PCM 16 (Pulse Code Modulation), а также стандарт квантования сигнала по уровню, в данном случае уровень сигнала кодируется 16 битным значением. Общая схема подготовки аудиозаписей представлена на рис. 1.

**Построение признакового описания аудиозаписей.** Перед расчетом признакового описания стереозапись делится на два канала, при этом каждый обрабатывается отдельно. Для улучшения качества классификации из моноканала выделяются еще два канала — гармонический и перкуссивный, используя для этого алгоритм HPSS (Harmonic-Percussive Sound Separation) [5]. Таким образом, двухканальная аудиозапись преобразована в четыре моноканальные записи.

В качестве способа представления аудиозаписей выбраны мел-спектрограммы со следующими параметрами: размер окна преобразования Фурье 100 мс; длина наложения 20 мс; количество банков мел-частотных фильтров 128.

**Подготовка и обучение нейронной сети.** Выберем сверточную нейронную сеть с 4-канальным входом на основе архитектуры VGGish. Каждый канал сформируем путем построения следующих спектрограмм: мел-спектрограммы левого и правого аудиоканалов, а также применение фильтра-гармоники и фильтра-перкуссии к усредненной по каналам мел-спектрограмме. Далее ко всем каналам применим набор фильтров-сверток, составив таким образом карты признаков, которые «разворачиваются» в одномерный вектор и подаются на вход полносвязному выходному слою. На выходе

[1] TUT Urban Acoustic Scenes 2018, Development Dataset [Электронный ресурс]. Режим доступа: https://zenodo.org/record/1228142 (дата обращения: 11.07.2022).

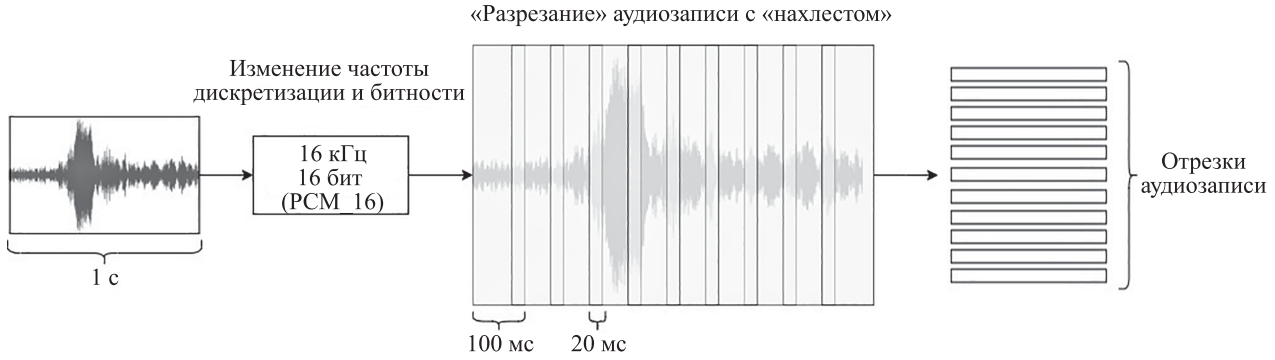


*Рис. 1*. Подготовка аудиозаписей

*Fig. 1*. Audio record preprocessing

нейронной сети установим слой Softmax, настроенный на 10 классов. Таким образом, ответом обученной нейронной сети будет распределение вероятности принадлежности аудиозаписи каждому из классов. Общая архитектура используемой сети представлена на рис. 2.

Обучение нейронной сети выполним методом обратного распространения ошибки с использованием стохастического градиентного спуска, а также момента Нестерова [6] для ускорения обучения и предотвращения переобучения. Шаг обучения составил 0,01, размер батча — 128. Результаты работы обученной нейронной сети на тестовом наборе данных представлены на рис. 3 и в табл. 1.

Из полученных данных видно, что классы «*поездка в трамвае*» и «*поездка в автобусе*» распознаются недостаточно хорошо, так как «спутаются» друг с другом. Решить данную проблему предложено с помощью подключения корректоров к тем классам, на которых ошибается модель.

*Таблица 1*. Качество классификации обученной модели для каждого класса

*Table 1.* Model precision and recall

| Класс | Точность | Полнота |
|---|---|---|
| Аэропорт | 0,8624 | 0,8704 |
| Торговый центр | 0,8876 | 0,9502 |
| Станция метро | 0,9197 | 0,9803 |
| Пешеходный переход | 0,9624 | 0,6798 |
| Городская площадь | 0,7879 | 0,8299 |
| Городская улица | 0,6450 | 0,7697 |
| Поездка в трамвае | 0,7063 | 0,5706 |
| Поездка в автобусе | 0,6272 | 0,7303 |
| Поездка в метро | 0,9140 | 0,9097 |
| Городской парк | 0,9963 | 0,9306 |
| Средняя невзвешенная точность (UAP): **0,8309**<br>Средняя невзвешенная полнота (UAR): **0,8221** | | |

### Применение корректора ошибок

Для устранения проблемы «спутывания» классов применим тип корректора — FN, который исправит ложноотрицательные срабатывания классификатора. Чтобы построить такой корректор, в первую очередь необходимо преобразовать обучающий набор данных. Для этого отберем все примеры в наборе, которые соответствуют необходимому классу. Рассмотрим случай «спутывания» классов «*поездка в трамвае*» и «*поездка в автобусе*», при этом классификатор имеет большое количество ложноотрицательных (FN) срабатываний на классе «*поездка в трамвае*». Для подготовки данных необходимо выполнить следующие шаги.

Шаг 1. Отобрать из тестового набора данных все примеры, соответствующие классу «*поездка в трамвае*».

Шаг 2. Применить обученный классификатор (в данному случае — нейронная сеть) для классификации примеров из Шага 1.

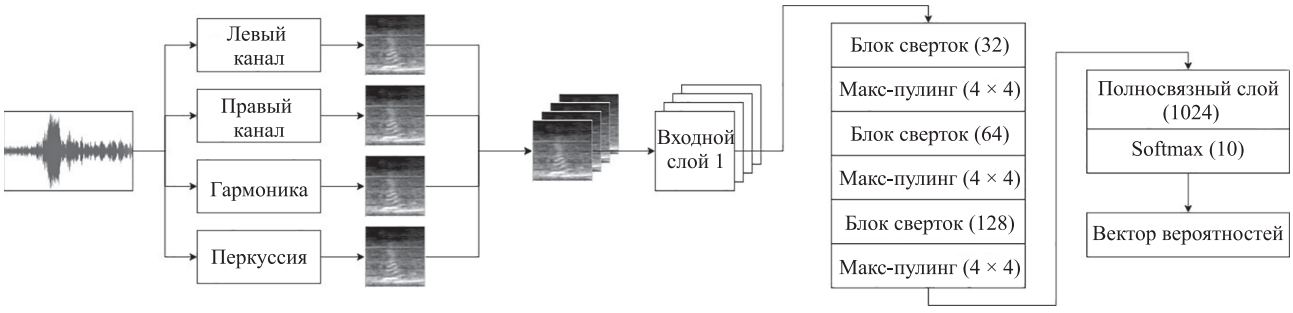


*Рис. 2*. Общая архитектура нейронной сети

*Fig. 2*. Audio classification pipeline

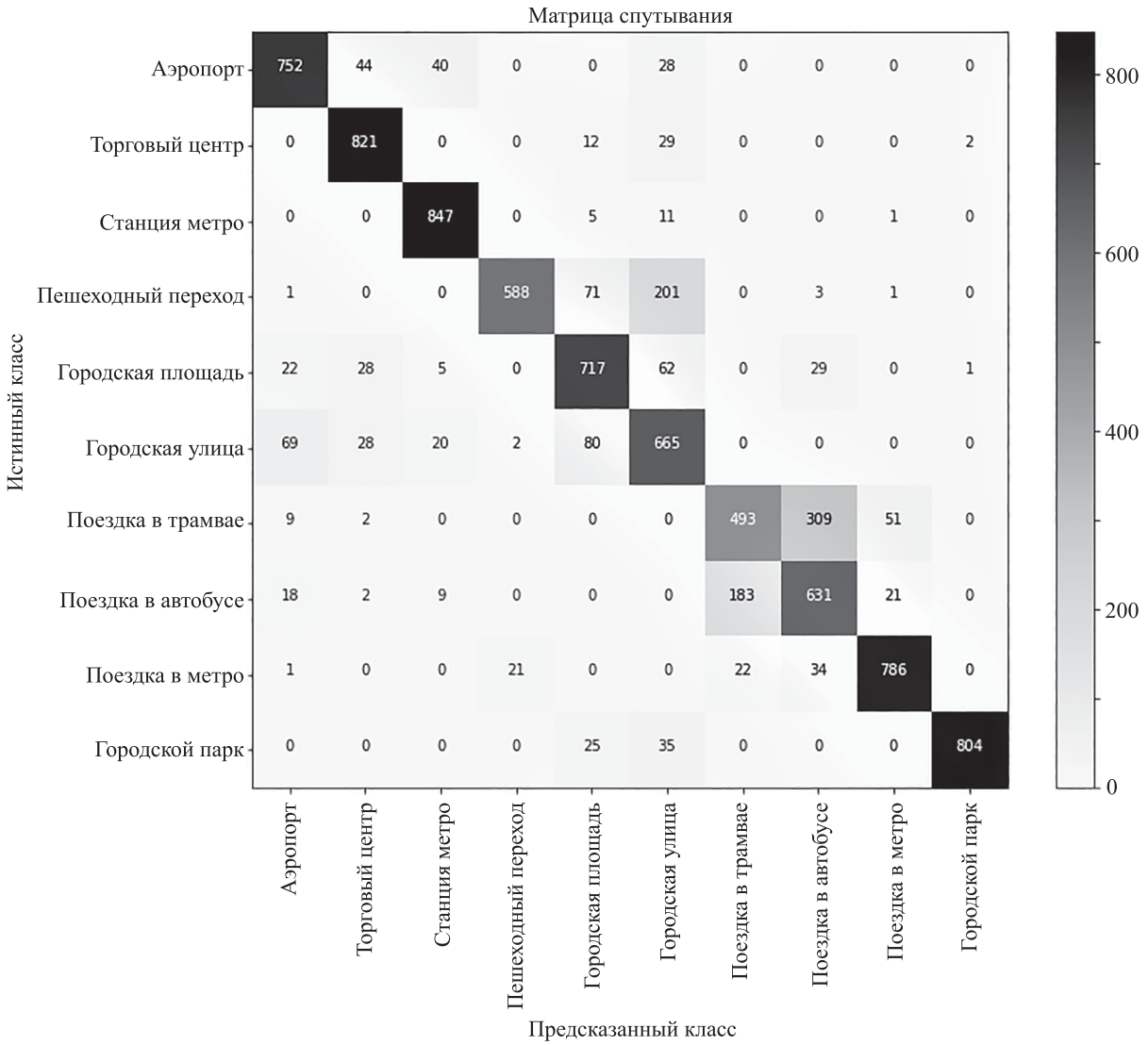


*Рис. 3*. Результаты классификации обученной модели в виде матрицы спутывания

*Fig. 3*. Classification confusion matrix

Шаг 3. На основе ответов классификатора составить набор данных вида {*признаковое описание примера, ответ*}, где *ответом* является 1, если ответ правильный, и 0, если ответ неправильный, т. е. классификатор «спутал» данный класс с другим.

Схема преобразования набора данных для корректора, рассматриваемого в эксперименте, приведена на рис. 4.

Обучим FN-корректор, как описано в работе [7]. В результате построим систему исправления ответа на классе «*поездка в трамвае*» (рис. 5).

После применения FN-корректора к выходу классификатора получим результаты, представленные на рис. 6 и в табл. 2.

В результате применения FN-корректора количество правильно распознанных объектов класса *«поездка в трамвае»* улучшилось с 493 (рис. 3) до 750 (рис. 6). Так как классификатор не был дополнительно изменен и обучен, то полученный результат говорит о том, что FN-корректор успешно решил поставленную задачу. По итогам эксперимента (табл. 2) видно увеличение точ-

*Таблица 2*. Качество классификации обученной модели для каждого класса после применения FN-корректора

*Table 2.* Classification quality of the trained model for each class after application of the corrector

| Класс | Точность | Полнота |
|---|---|---|
| Аэропорт | 0,8664 | 0,8704 |
| Торговый центр | 0,8876 | 0,9502 |
| Станция метро | 0,9197 | 0,9803 |
| Пешеходный переход | 0,9624 | 0,6798 |
| Городская площадь | 0,7879 | 0,8299 |
| Городская улица | 0,6450 | 0,7697 |
| Поездка в трамвае | 0,7853 | 0,7042 |
| Поездка в автобусе | 0,6614 | 0,7303 |
| Поездка в метро | 0,9140 | 0,9097 |
| Городской парк | 0,9963 | 0,9306 |
| Среднее | **0,8426** | **0,8355** |

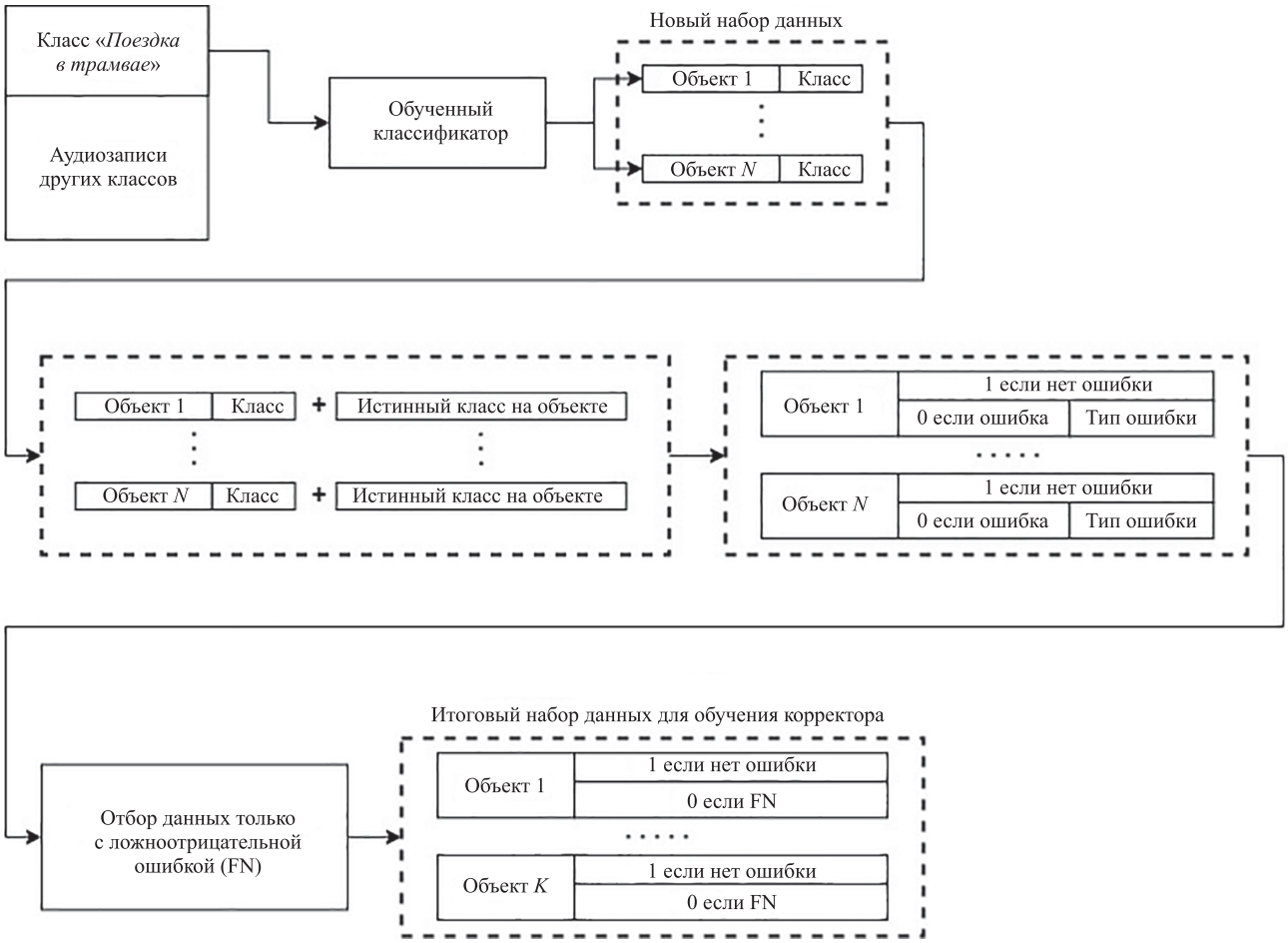


*Рис. 4*. Схема преобразования набора данных для построения FN-корректора

*Fig. 4*. Data preprocessing to train FN-corrector

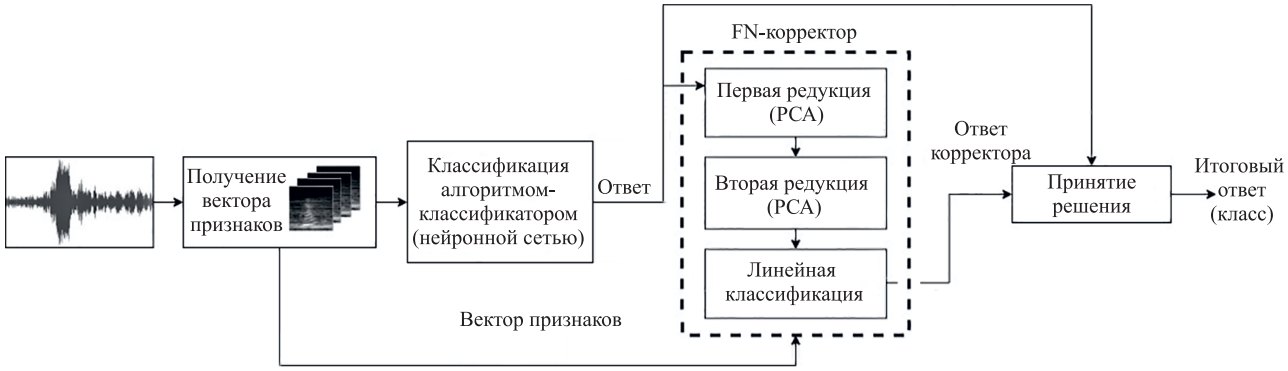


*Рис. 5*. Схема системы классификатор-корректор FN в применении к классу «*поездка в трамвае*»

*Fig. 5*. Classification-correction pipeline

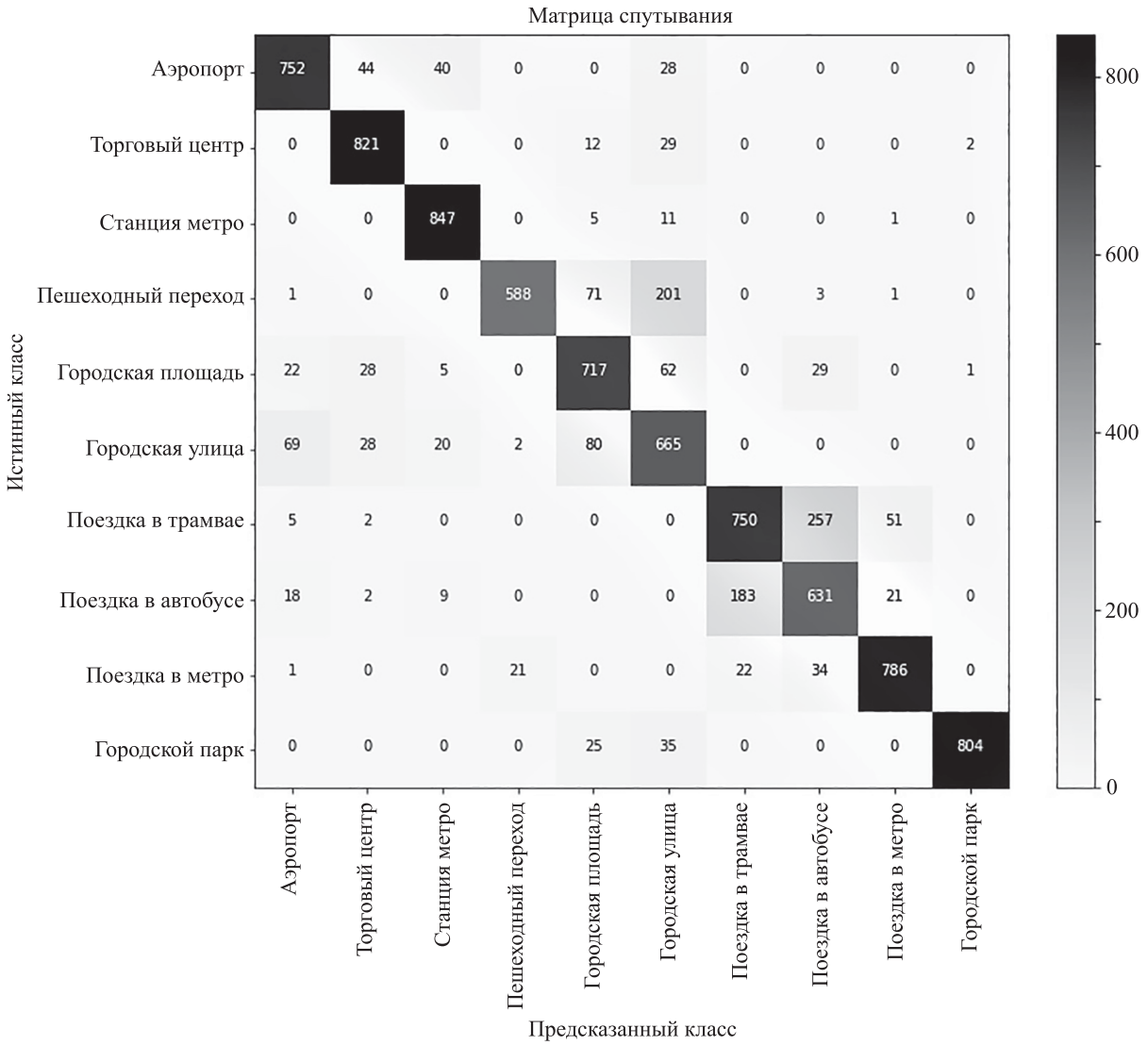


*Рис. 6*. Матрица смешивания после применения FN корректора

*Fig. 6*. Classification-correction of the model confusion matrix

ности классификатора до 0,8426 и полноты до 0,8355 (значимость результатов подтверждается проверкой методом кросс-валидации по 4 группам).

## Заключение

В работе представлен и описан способ использования метода каскадной редукции пространства признаков при решении задачи классификации акустических событий с помощью построения системы коррекции ответов классификатора. Продемонстрировано статистически значимое увеличение точности классификатора при использовании описанного способа для решения задачи классификации акустических событий. Выделен класс, который классификатор распознает недостаточно хорошо («спутывает» его с другим классом), и на данном классе был обучен и применен FN-корректор. В результате эксперимента показано улучшение точности классификатора с 0,8308 до 0,8426 и полноты с 0,8221 до 0,8355 (значимость результатов подтверждается проверкой методом кросс-валидации по 4 группам). Преимущество описанного способа — возможность применения в качестве постобработки и использование совместно со многими существующими архитектурами классификаторов. Другое потенциальное преимущество — возможность создания модели небольшого размера, использующей данный способ, что важно для задач обработки звука на устройствах, например IoT. В дальнейших работах планируется исследовать возможности задач обработки звука на устройствах IoT.

### Литература

1. Grollmisch S., Cano E., Kehling C., Taenzer M. Analyzing the potential of pre-trained embeddings for audio classification tasks // Proc. of the 28th European Signal Processing Conference (EUSIPCO). 2021. P. 790–794. https://doi.org/10.23919/Eusipco47968.2020.9287743
2. Matveev Y.N., Shuranov E.V., Avdeeva A.S., Shchemelinin V.L., Krylova E.V. Acoustic data based automatic object detection system // Proc. of the 2nd International Conference on Control in Technical Systems (CTS). 2017. P. 301–303. https://doi.org/10.1109/CTSYS.2017.8109551
3. Голубков А.М. Бинарная классификация изображений на примере задачи распознавания лиц // Известия СПбГЭТУ «ЛЭТИ». 2018. № 7. С. 26–30.
4. Голубков А.М., Клионский Д.М. Применение метода каскадной редукции к решению задачи распознавания лиц // Известия СПбГЭТУ «ЛЭТИ». 2019. № 8. С. 47–53.
5. Ono N., Miyamoto K., Le Roux J., Kameoka H., Sagayama S. Separation of a monaural audio signal into harmonic/percussive components by complementary diffusion on spectrogram // Proc. of the 16th European Signal Processing Conference (EUSIPCO). 2008. P. 1–4.
6. Sutskever I., Martens J., Dahl G., Hinton G. On the importance of initialization and momentum in deep learning // Proc. of the 30th International Conference on Machine Learning (ICML). 2013. P. 2176–2184.
7. Gorban A., Golubkov A.M., Grechuk B., Mirkes E., Tyukin I.Y. Correction of AI systems by linear discriminants: probabilistic foundations // Information Sciences. 2018. V. 466. P. 303–322. https://doi.org/10.1016/j.ins.2018.07.040



### References

1. Grollmisch S., Cano E., Kehling C., Taenzer M. Analyzing the potential of pre-trained embeddings for audio classification tasks. *Proc. of the 28th European Signal Processing Conference (EUSIPCO)*, 2021, pp. 790–794. https://doi.org/10.23919/Eusipco47968.2020.9287743
2. Matveev Y.N., Shuranov E.V., Avdeeva A.S., Shchemelinin V.L., Krylova E.V. Acoustic data based automatic object detection system. *Proc. of the 2nd International Conference on Control in Technical Systems (CTS)*, 2017, pp. 301–303. https://doi.org/10.1109/CTSYS.2017.8109551
3. Golubkov A.M. Face recognition using images binary classification methods. *Proceedings of Saint Petersburg Electrotechnical University Journal*, 2018, no. 7, pp. 26–30. (in Russian)
4. Golubkov A.M., Klionskii D.M. Cascade reduction method applied to face recognition probleb. *Proceedings of Saint Petersburg Electrotechnical University Journal*, 2019, no. 8, pp. 47–53. (in Russian)
5. Ono N., Miyamoto K., Le Roux J., Kameoka H., Sagayama S. Separation of a monaural audio signal into harmonic/percussive components by complementary diffusion on spectrogram. *Proc. of the 16th European Signal Processing Conference (EUSIPCO)*, 2008, pp. 1–4.
6. Sutskever I., Martens J., Dahl G., Hinton G. On the importance of initialization and momentum in deep learning. *Proc. of the 30th International Conference on Machine Learning (ICML)*, 2013, pp. 2176–2184.
7. Gorban A., Golubkov A.M., Grechuk B., Mirkes E., Tyukin I.Y. Correction of AI systems by linear discriminants: probabilistic foundations. *Information Sciences*, 2018, vol. 466, pp. 303–322. https://doi.org/10.1016/j.ins.2018.07.040



### Авторы

**Голубков Александр Михайлович** — кандидат технических наук, ассистент, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), Санкт-Петербург, 197022, Российская Федерация; лидер команды аудио аналитики, ООО «ТЕХКОМПАНИЯ ХУАВЭЙ», Москва, 123007, Российская Федерация, SC 57190975154, https://orcid.org/0000-0002-8330-1823, kremnikov@gmail.com

**Шуранов Евгений Витальевич** — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация; руководитель лаборатории, ООО «ТЕХКОМПАНИЯ ХУАВЭЙ», Москва, 123007, Российская Федерация, SC 57190970283, https://orcid.org/0000-0003-0977-5075, e_v_shuranov@mail.ru



### Authors

**Alexander M. Golubkov** — PhD, Assistant, Saint Petersburg Electrotechnical University "LETI", Saint Petersburg, 197022, Russian Federation; Senior Machine Learning Engineer, Huawei, Moscow, 123007, Russian Federation, SC 57190975154, https://orcid.org/0000-0002-8330-1823, kremnikov@gmail.com

**Evgeny V. Shuranov** — PhD, Associate Professor, ITMO University, 197101, Russian Federation; Head of Laboratory, Huawei, Moscow, 123007, Russian Federation, SC 57190970283, https://orcid.org/0000-0003-0977-5075, e_v_shuranov@mail.ru



*Статья поступила в редакцию 17.04.2022*
*Одобрена после рецензирования 24.06.2022*
*Принята к печати 30.07.2022*

*Received 17.04.2022*
*Approved after reviewing 24.06.2022*
*Accepted 30.07.2022*