Рубрика 2. НАУЧНЫЕ И ПРАКТИЧЕСКИЕ РАЗРАБОТКИ
Направление – Математическое моделирование, численные методы и комплексы программ

УДК [UDC] 629.1.06
DOI 10.17816/transsyst20239195-107

© А.А. Лисов, А.З. Кулганатов, С.А. Панишев
Южно-Уральский государственный университет
(Челябинск, Россия)

АКУСТИЧЕСКОЕ ОБНАРУЖЕНИЕ ТРАНСПОРТНЫХ СРЕДСТВ АВАРИЙНЫХ СЛУЖБ С ИСПОЛЬЗОВАНИЕМ СВЕРХТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Обоснование: Сирена – это особый сигнал, подаваемый транспортными средствами службы экстренной помощи, такими как: пожарные машины, полицейские машины и машины скорой помощи для предупреждения водителей или пешеходов на дороге. Однако водители иногда могут не услышать звуки сирены из-за звукоизоляции современного автомобиля, шума городского трафика или причине собственной невнимательности. Эта проблема может привести к задержке в предоставлении помощи экстренных служб или даже к дорожно-транспортным происшествиям.

Цель: разработка акустического метода обнаружения присутствия автомобилей экстренных служб на дороге посредством применения сверточных нейронных сетей**.**

Материалы и методы: Алгоритм работы основан на преобразовании звука из внешней среды в его спектрограмму, для анализа методом машинного обучения – сверточными нейронными сетями. В качестве датасета звуков сирены и городского трафика использовался открытый набор данных (Emergency Vehicle Siren Sounds) из источников, доступных на интернет-сайтах, таких как Google и Youtube, сохраненных в аудиоформате “.wav”. Код разрабатывался на платформе Google.Colab при помощи облачного хранилища.

Результаты: Проведенные эксперименты показали, что предлагаемый метод и архитектура нейросети позволяют достичь средней эффективности определения типа звука с точностью 93,3 % и скоростью 0,0004±5 % секунды.

Заключение: Использование разработанной технологии распознавания сигналов экстренных служб в условиях городского трафика позволит повысить безопасность дорожного движения и увеличить шансы на предотвращение опасной ситуации. Также данная система может являться дополнительным помощником для слабослышащих людей во время вождения и повседневной жизни для своевременного оповещения о наличии поблизости экстренных служб.

Ключевые слова: машинное обучение, сверточные нейронные сети, распознавание сигналов аварийных служб.

Rubric 2. SCIENTIFIC AND PRACTICAL DEVELOPMENTS
Field – Mathematical modeling, numerical methods and software packages

© A.A. Lisov, A.Z. Kulganatov, S.A. Panishev
South Ural State University
(Chelyabinsk, Russia)

Received: 24.02.2023 **Revised: 10.03.2023** **Accepted: 30.03.2023**
Поступил: 24.02.2023 **Одобрена: 10.03.2023** **Принята: 30.03.2023**

USING CONVOLUTIONAL NEURAL NETWORKS FOR ACOUSTIC-BASED EMERGENCY VEHICLE DETECTION

Background: A siren is a special signal given by emergency vehicles such as fire trucks, police cars and ambulances to warn drivers or pedestrians on the road. However, drivers sometimes may not hear the siren due to the sound insulation of a modern car, the noise of city traffic, or their own inattention. This problem can lead to a delay in the provision of emergency services or even to traffic accidents.

Aim: develop an acoustic method for detecting the presence of emergency vehicles on the road through the use of convolutional neural networks.

Materials and Methods: The algorithm of work is based on the conversion of sound from the external environment into its spectrogram, for analysis by a convolutional neural network. An open dataset ("Emergency Vehicle Siren Sounds") from sources available on Internet sites such as Google and Youtube, saved in the ".wav" audio format, was used as a dataset for siren sounds and city traffic. The code was developed on the Google.Colab platform using cloud storage.

Results: The conducted experiments showed that the proposed method and model of the neural network make it possible to achieve an average efficiency of determining the type of sound with an accuracy of 93.3 % and a speed recognition of 0.0004±5 % of a second.

Conclusion: The use of the developed technology for recognizing siren sounds in city noize will improve traffic safety and increase the chances of preventing a dangerous situation. Also, this system can be an additional assistant for hearing-impaired people while driving and everyday life for timely notification of the presence of emergency services nearby.

Key words: machine learning, convolutional neural networks, emergency vehicle signal recognition.

ВВЕДЕНИЕ

Автомобили экстренной помощи играют важную роль в ситуации, угрожающей жизни, своевременное реагирование является важным аспектом данной работы. Городские пробки уносят более 20 % жизней пациентов в машине скорой помощи, а при очень тяжелом состоянии пациента процент летального исхода увеличивается [1]. В данной ситуации многие водители могут не пропустить машину скорой помощи из-за шума городского трафика, невнимательности, громкой музыки в машине или внешней звукоизоляции автомобиля. Для решения данной проблемы мы предлагаем внедрить интеллектуальную систему распознавания сигналов экстренных служб при помощи методов машинного обучения, в частности сверточных нейронных сетей.

Чтобы построить модель для распознавания звуков экстренных служб, первым делом нужно решить, в каком виде будут представлены и использованы данные. Можно строить модели, используя необработанную форму звуковой волны [2, 3], либо использовать двумерное представление звука, например в виде спектрограмм [4–6]. Спектрограммы становятся все более популярными в последнее время, потому что они хорошо работают со сверточными нейронными сетями (CNN) [4, 7]. Хоть модели CNN были

Received: 24.02.2023 **Revised: 10.03.2023** **Accepted: 30.03.2023**
Поступил: 24.02.2023 **Одобрена: 10.03.2023** **Принята: 30.03.2023**

созданы для анализа естественных изображений, двумерные спектрограммы также могут быть использованы для обучения. Спектрограммы звуков содержат различные, но повторяющиеся паттерны. Поэтому в данной работе и было принято решение преобразования звука в спектрограмму, которая будет анализироваться специально обученной сверточной нейронной сетью.

Для решения данной задачи некоторые авторы создали ядра свертки (kernels), которые перемещаются только в одном направлении для сбора временных данных [8]. Другие предлагали рекуррентные нейронные сети RNN [9–13] или их комбинацию с CNN [6, 14, 15] для улучшения последовательного понимания данных. В 2014 году исследование [16] показало, что мы можем рассматривать эти спектрограммы как изображения и использовать стандартную архитектуру CNN, такую как AlexNet [17], предварительно обученную на ImageNet [18] для задачи классификации звука.

МАТЕРИАЛЫ И МЕТОДЫ

В данной работе проводились исследования в области машинного обучения для решения задачи обнаружения транспортных средств аварийных служб в условиях городского трафика. Основой разработанной программы для решения поставленной задачи является сверточная нейронная сеть, которая работает с открытым датасетом “Emergency Vehicle Siren Sounds”, который можно найти на “Kaggle” [19]. Этот набор данных был полностью создан вручную путем извлечения звука из источников, доступных на интернет-сайтах, таких как Google и Youtube, и сохранен в аудиоформате “.wav”. Датасет содержит звуки сирен автомобилей скорой помощи, шума городского трафика и пожарных машин. Каждая категория содержит 200 звуковых файлов и 200 изображений спектрограмм (которые не будут использованы в данной работе) для каждого звукового файла. Также следует отметить, что этот набор данных был специально создан для проекта машинного обучения, этот набор данных не связан с какой-либо организацией и не подлежит внешнему лицензированию, скачивание бесплатное, изменение и распространение этого набора данных полностью свободное. Исходная выборка изображений была разделена на 2 класса “test” и “train”, разделение выборки произведено в соотношении 1:10.

Код был разработан на платформе “Google.Collab”, на языке программирования “python”, основной библиотекой для машинного обучения является “TensorFlow” и в особенности ее модуль “Keras”. Код программы можно найти на GitHub одного из авторов по имени пользователя “AnLiMan” по названию репозитория “CNN-for-audio-recognition” [20]. Код является полностью доступным к редактированию и

Received: 24.02.2023 Revised: 10.03.2023 Accepted: 30.03.2023
Поступил: 24.02.2023 Одобрена: 10.03.2023 Принята: 30.03.2023

коммерческому использованию и свободному распространению. Модель сверточной нейросети – последовательная (Sequential), содержащая в себе несколько слоев свертки для обработки изображений, визуализация ее архитектуры представлена на Рис. 1.

В работе использовались следующие методы: эксперимент и математическое моделирование. Основными методами процесса обучения нейронной сети, использованными в данной работе, являются алгоритм обратного распространения ошибки и использование матриц свертки для обнаружения закономерностей на изображениях. Ошибка распознавания рассчитывалась с использованием метода расчета стандартного отклонения.

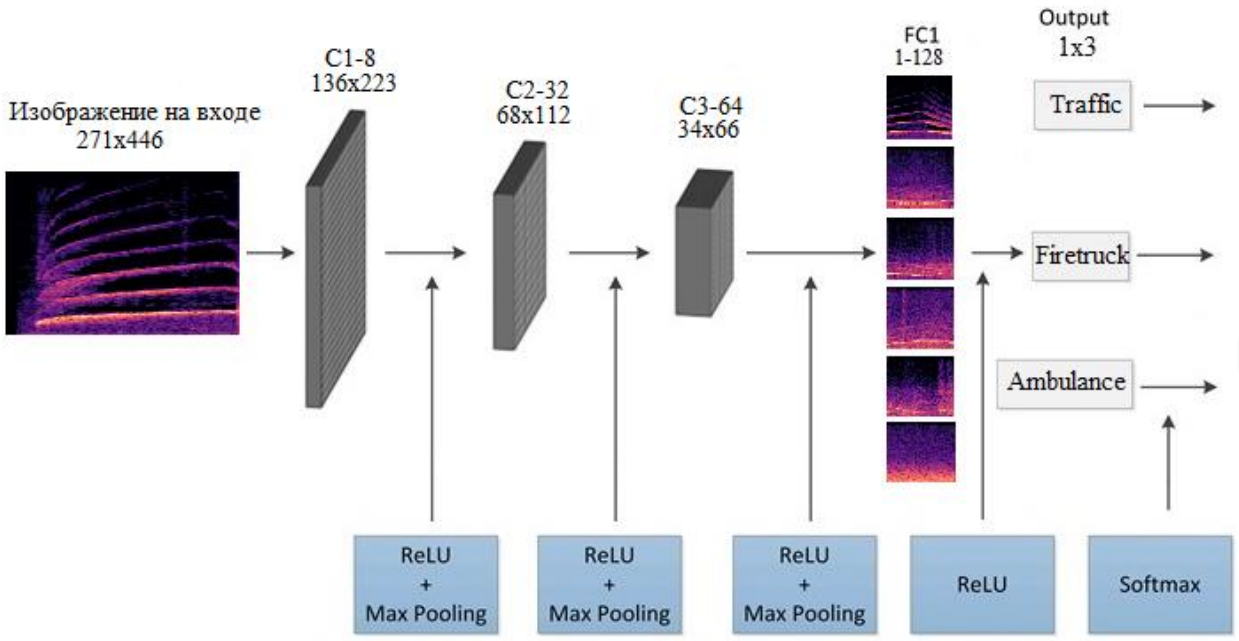

Рис. 1. Архитектура CNN для распознавания сирены аварийных служб

Источник: составлено авторами

Для преобразования звуков в спектрограммы использовалась библиотека “Librosa” – это библиотека Python для анализа музыки и аудио. Она предоставляет функции, необходимые для создания систем поиска информации о музыке и аудио. В соответствии с разработанной методикой и идей проекта – звук с микрофона передается в записывающее устройство, которое делит его на фрагменты по 3 секунды и отправляет на обработку сверточной нейросети. Пример, генерируемых спектрограмм скорой машины (ambulance) показана на Рис. 2, пожарной машины (firetruck) на Рис. 3 и городского шума (traffic) на Рис. 4. Как видно из иллюстраций спектрограммы звуков сирены имеют четкие паттерны и определенный частотный диапазон, в то время как городской трафик просто равномерно занимает всю доступную слышимую область. Однако

Received: 24.02.2023 Revised: 10.03.2023 Accepted: 30.03.2023
Поступил: 24.02.2023 Одобрена: 10.03.2023 Принята: 30.03.2023

для работы сети требуется убрать дополнительную информацию о сэмпле, которая также генерируется библиотекой “Librosa” до уровня, показанного на Рис. 5.

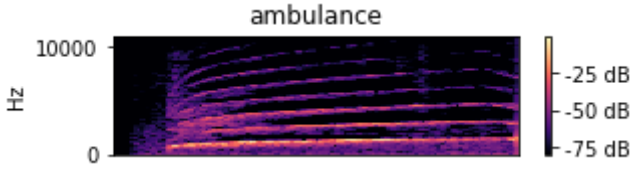

Рис. 2. Спектрограмма скорой машины (ambulance)

Источник: составлено авторами

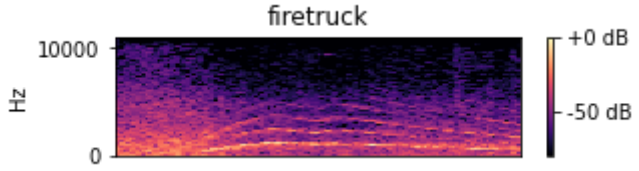

Рис. 3. Спектрограмма пожарной машины (firetruck)

Источник: составлено авторами

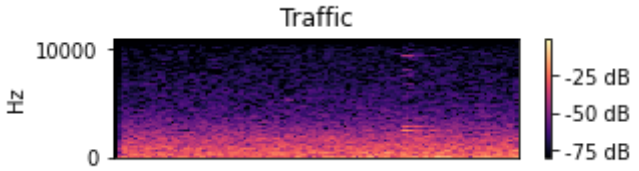

Рис. 4. Спектрограмма городского шума (traffic)

Источник: составлено авторами

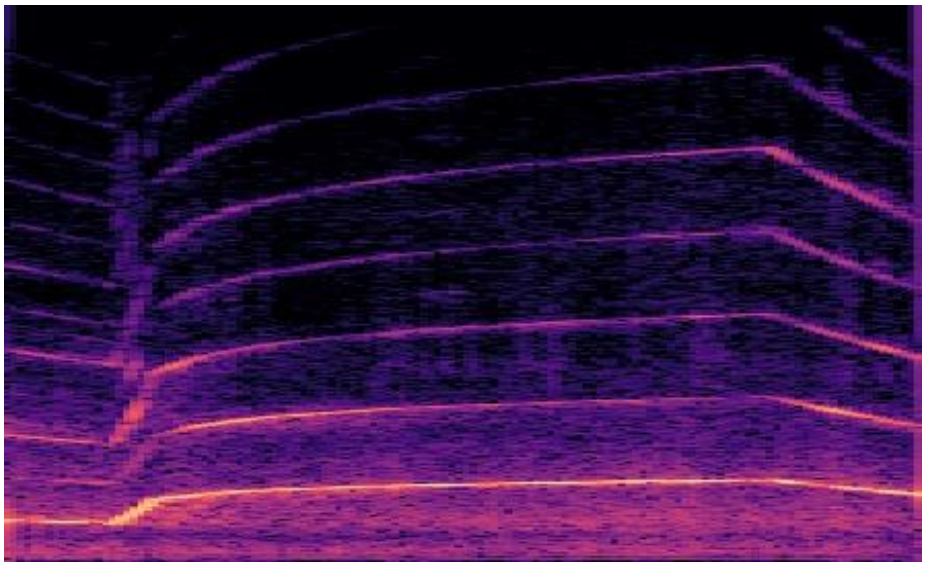

Рис. 5. «Очищенная» спектрограмма звуковой дорожки

Источник: составлено авторами

Received: 24.02.2023 **Revised: 10.03.2023** **Accepted: 30.03.2023**
Поступил: 24.02.2023 **Одобрена: 10.03.2023** **Принята: 30.03.2023**

Создание модели CNN.

Сверточная нейронная сеть (ConvNet / CNN) – это алгоритм глубокого обучения, который может принимать входное изображение, назначать важность различным аспектам или объектам на изображении и иметь возможность отличать одино от другого. CNN может успешно фиксировать пространственные и временные зависимости в изображении с помощью применения соответствующих фильтров.

При разработке сверточной нейросети использовалась библиотека от Google – “Tensorflow” и в особенности ее модуль “Keras”, модель сверточной нейросети – последовательная (Sequential), содержащая в себе несколько слоев свертки для обработки изображений, код для создания архитектуры модели представлен ниже, визуализация архитектуры представлена на Рис. 1.

```
model = tf.keras.models.Sequential([
    tf.keras.layers.Conv2D(8, (2,2), activation='relu', input_shape=(271, 446, 3)),
    tf.keras.layers.MaxPooling2D(2, 2),
    tf.keras.layers.Conv2D(32, (2,2), activation='relu'),
    tf.keras.layers.MaxPooling2D(2,2),
    tf.keras.layers.Conv2D(64, (2,2), activation='relu'),
    tf.keras.layers.MaxPooling2D(2,2),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(128, activation='relu'),
    tf.keras.layers.Dropout(0.2),
    tf.keras.layers.Dense(3, activation='softmax')
])
```

Алгоритм работы.

Алгоритм обучения сверточной нейросети показан на Рис. 6. Процесс обучения начинается с импорта библиотек машинного обучения, работы со звуком, массивами, архивами, изображениями и облачными хранилищами. После импорта звуков они с помощью библиотеки “Librosa” преобразуются в спектрограммы, показанные ранее на Рис. 2–4 до уровня изображений, показанных на Рис. 5. Далее идет определение архитектуры сети, которая была подробно описана ранее. Основной процесс обучения занимает порядка 12 минут. В случае, если точность распознавания превысит 93 %, либо если закончится количество эпох обучения, то процесс обучения завершится. После модель и ее веса сохраняются на облачном хранилище, чтобы не повторять процесс обучения снова в будущем, а просто загрузить готовую модель для работы.

Received: 24.02.2023 Revised: 10.03.2023 Accepted: 30.03.2023
Поступил: 24.02.2023 Одобрена: 10.03.2023 Принята: 30.03.2023

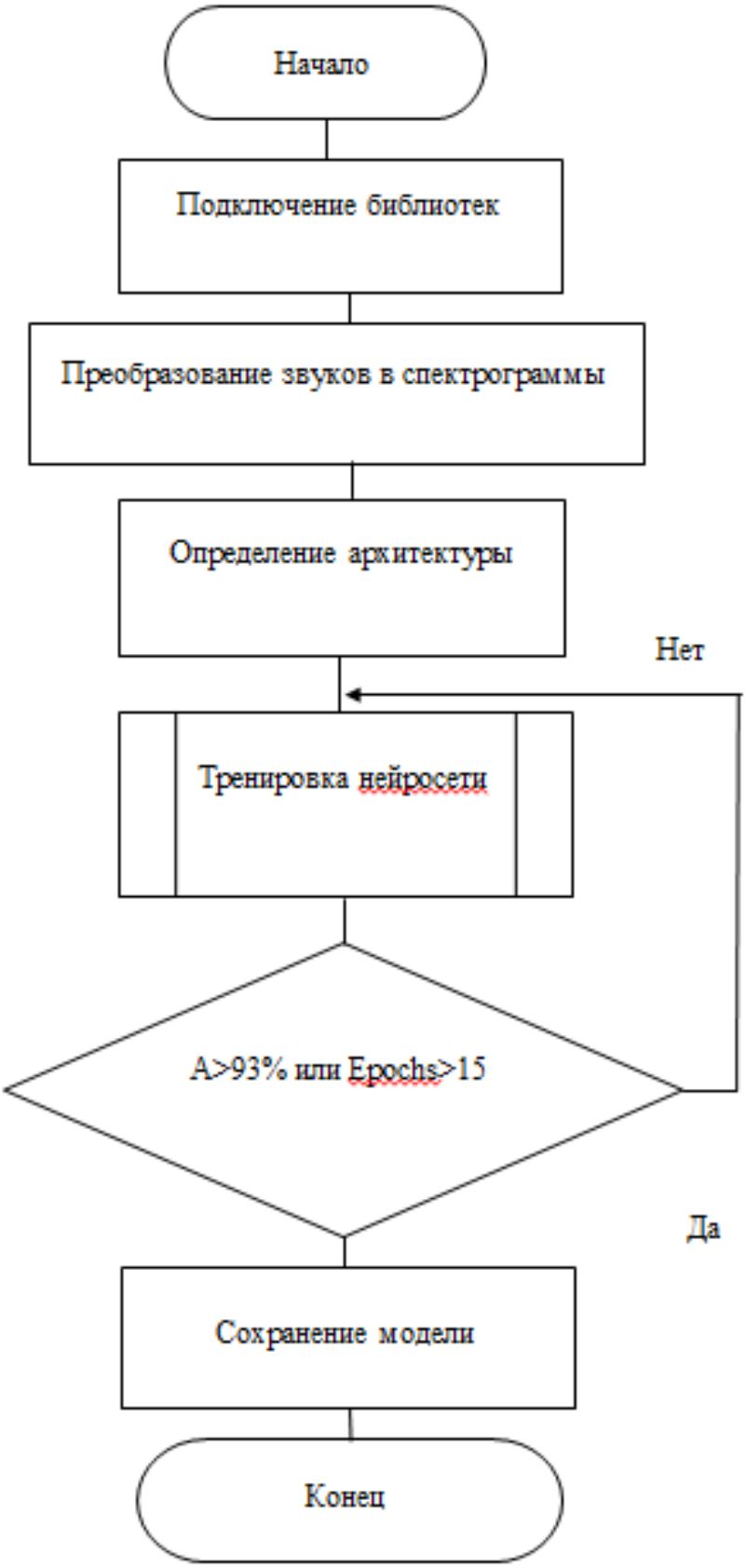

Рис. 6. Алгоритм обучения сверточной нейронной сети

Источник: составлено авторами

Received: 24.02.2023 Revised: 10.03.2023 Accepted: 30.03.2023
Поступил: 24.02.2023 Одобрена: 10.03.2023 Принята: 30.03.2023

Резюме сверточной нейронной сети приведено в Табл. 1. Из данной таблицы можно увидеть, что конечная архитектура нейросети содержит примерно 14,6 млн параметров, что является, довольно, небольшим значением.

Таблица. Параметры модели нейросети

Тип слоя	Размер на выходе	Число параметров
conv2d (Conv2D)	(None, 270, 445, 8)	104
max_pooling2d (MaxPooling2D)	(None, 135, 222, 8)	0
conv2d_1 (Conv2D)	(None, 134, 221, 32)	1056
max_pooling2d_1 (MaxPooling 2D)	(None, 67, 110, 32)	0
conv2d_2 (Conv2D)	(None, 66, 109, 64)	8256
max_pooling2d_2 (MaxPooling 2D)	(None, 33, 54, 64)	0
flatten (Flatten)	(None, 114048)	0
dense (Dense)	(None, 128)	14 598 272
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 3)	387
Всего параметров		14 608 075

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Как уже было сказано ранее, примерное время обучения составило 12 мин., график данного процесса показан на Рис. 7. Для последней эпохи обучения были получены следующие результаты: loss: 0,1844 – accuracy: 0,9212 – val_loss: 0,2081 – val_accuracy: 0,9333. Здесь loss – «потери», разница между полученным значением предсказания и реальным, accuracy – точность распознавания для текущей эпохи. Главные показатели это “val_loss” и “val_accuracy”, отображающие итоговые результаты обучения сети.

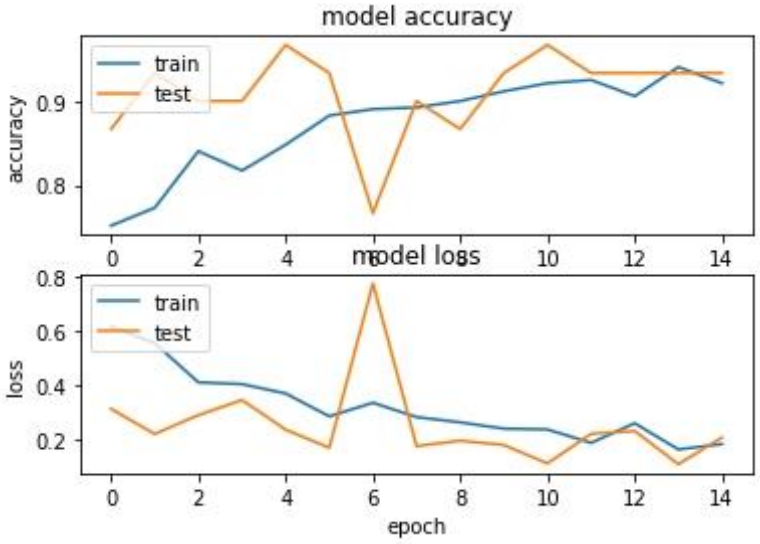

Рис. 7. График процесса обучения

Источник: составлено авторами

Received: 24.02.2023 Revised: 10.03.2023 Accepted: 30.03.2023
Поступил: 24.02.2023 Одобрена: 10.03.2023 Принята: 30.03.2023

Далее была проведена проверка точности распознавания нейросети. Ниже на Рис. 8 приведен результат распознавания 16-ти случайных изображений из тестовой выборки. Зеленый цвет означает, что изображение было распознано правильно, за скобками указано предсказанное значение, а в скобках реальное. Как видно из Рис. 8, точность в 93,3 % позволяет достичь достаточно высоких результатов. Алгоритм работы можно несколько изменить, если в качестве обучающего набора использовать аудиофайлы короче 3-х секунд и усреднять итоговый прогноз. Однако в рамках данного исследования такой эксперимент не проводился.

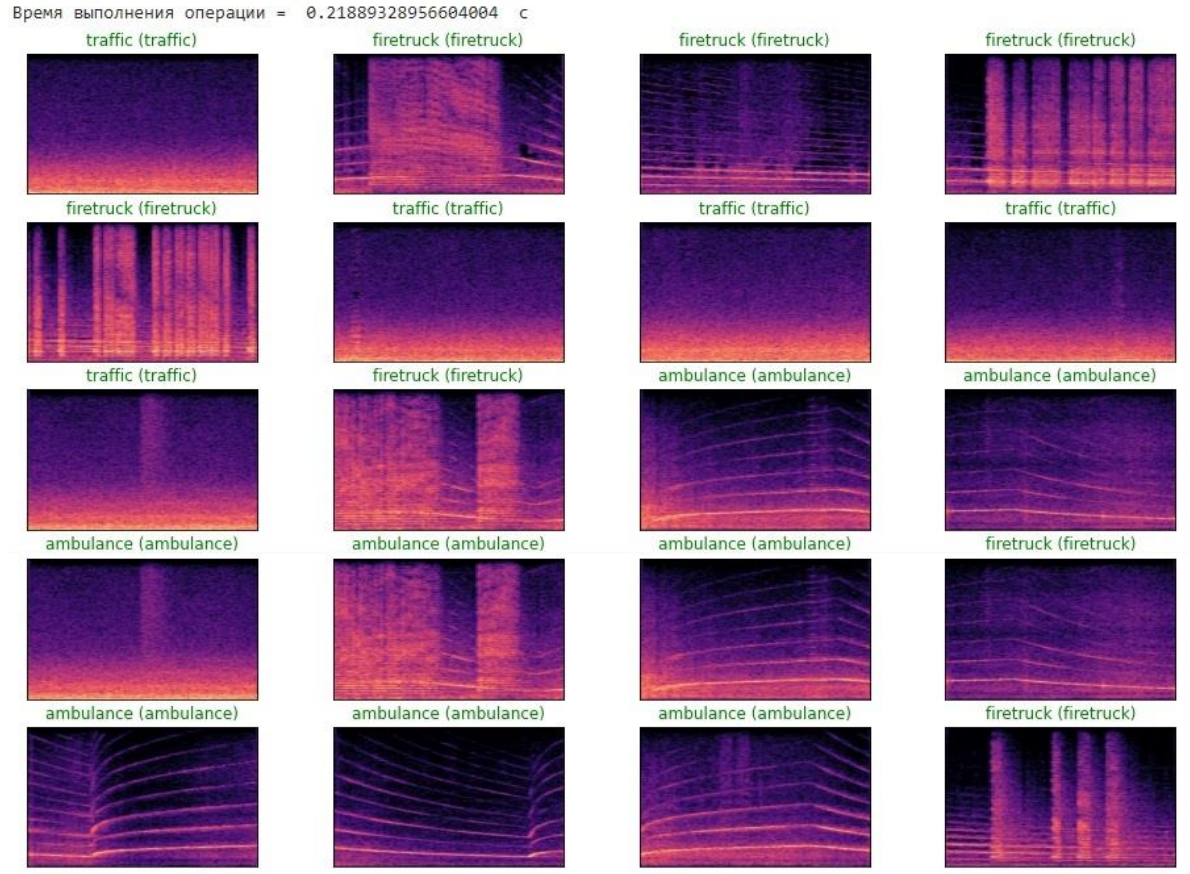

Рис. 8. Проверка точности распознавания на 16-ти случайных спектрограммах

Источник: составлено авторами

Также был проведен тест скорости распознавания и вывода одного изображения (Рис. 9) результат составил примерно 0,0205 секунды. Однако алгоритм работы можно усовершенствовать, если убрать вывод изображения, в котором нет необходимости человеку, в этом случае скорость распознавания составит – 0,0004±5 % секунды.

Received: 24.02.2023 **Revised: 10.03.2023** **Accepted: 30.03.2023**
Поступил: 24.02.2023 **Одобрена: 10.03.2023** **Принята: 30.03.2023**

Рис. 9. Проверка отдельного изображения из тестовой выборки

Источник: составлено авторами

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Для увеличения скорости работы нейронной сети мы предлагаем убрать компонент визуализации результата предсказания, что позволит добиться скорости распознавания – 0,0004±5 % секунд для всех изображений из тестовой выборки.

Следует отметить важное ограничение данной работы – достаточно сложно добиться высокой точности распознавания без использования дополнительных методов фильтрации и обработки изображений. Поэтому для повышения производительности сети необходимо применять алгоритмы предварительной обработки необработанных изображений, например, фильтром Гаусса или аналогичным. Также следует провести постобработку результатов прогнозирования, путем усреднения или перепроверки. Для улучшения качества предсказания также можно повысить порог срабатывания (threshold) нейрона, но в этом случае потребуется увеличить размер датасета.

ЗАКЛЮЧЕНИЕ

Использование разработанной технологии распознавания сигналов экстренных служб в условиях городского трафика позволит повысить безопасность дорожного движения и увеличить шансы на предотвращение опасной ситуации.

Также данная система может являться дополнительным помощником для слабослышащих людей во время вождения и повседневной жизни для своевременного оповещения о наличии поблизости экстренных служб.

Разработанная модель сверточной нейронной сети может быть легко интегрирована в интеллектуальную систему помощи водителю или для системы автопилота беспилотного автомобиля.

Результаты исследования показали эффективность работы сети на уровне 93,3 % и скорость распознавания примерно равной 0,4 м/с при использовании интерпретируемого языка python. Данное время можно

Received: 24.02.2023 **Revised: 10.03.2023** **Accepted: 30.03.2023**
Поступил: 24.02.2023 **Одобрена: 10.03.2023** **Принята: 30.03.2023**

уменьшить, если экспортировать модель на компилируемый язык программирования, например, C#.

Датасет звуков сирены, который использовался для эксперимента может быть успешно заменен соответствующими звуками страны, в которой планируется использование данной системы без критического изменения кода разработанной программы.

РЕКОМЕНДАЦИЯ К ПУБЛИКАЦИИ

Старший научный сотрудник Южно-Уральского Государственного университета, доктор технических наук, профессор Александр Григорьевич Возмилов рекомендует данную статью к публикации.

Авторы заявляют, что:

1. у них нет конфликта интересов;
2. настоящая статья не содержит каких-либо исследований с участием людей в качестве объектов исследований.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК / References

1. Kanzaria HK, Probst MA, Hsia RY. Emergency department death rates dropped by nearly 50 percent, 1997–2011. *Health Affairs*. 2016 Jul 1;35(7):1303-8. doi:10.1377/hlthaff.2015.1394

2. Lee J, Park J, Kim KL, Nam J. Sample-level deep convolutional neural networks for music auto-tagging using raw waveforms. *arXiv preprint* arXiv:1703.01789. 2017 Mar 6. doi: 10.48550/arXiv.1703.01789

3. Zhu Z, Engel JH, Hannun A. Learning multiscale features directly from waveforms. arXiv preprint arXiv:1603.09509. 2016 Mar 31. doi: 10.48550/arXiv.1603.09509

4. Choi K, Fazekas G, Sandler M. Automatic tagging using deep convolutional neural networks. *arXiv preprint* arXiv:1606.00298. 2016 Jun 1. doi: 10.48550/arXiv.1606.00298

5. Nasrullah Z, Zhao Y. Music artist classification with convolutional recurrent neural networks. *In2019 International Joint Conference on Neural Networks (IJCNN)* 2019 Jul 14 (pp. 1-8). IEEE. doi: 10.1109/IJCNN.2019.8851988

6. Wang Z, Muknahallipatna S, Fan M, et al. Music classification using an improved crnn with multi-directional spatial dependencies in both time and frequency dimensions. *In2019 International Joint Conference on Neural Networks (IJCNN)* 2019 Jul 14 (pp. 1-8). IEEE. doi: 10.1109/IJCNN.2019.8852128

7. Dieleman S, Brakel P, Schrauwen B. Audio-based music classification with a pretrained convolutional network. *In12th International Society for Music Information Retrieval Conference (ISMIR-2011)* 2011 (pp. 669-674). University of Miami.

8. Chen MT, Li BJ, Chi TS. CNN based two-stage multi-resolution end-to-end model for singing melody extraction. *InICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 2019 May 12 (pp. 1005-1009). IEEE. doi: 10.1109/ICASSP.2019.8683630

9. Phan H, Koch P, Katzberg F, et al. Audio scene classification with deep recurrent

Received: 24.02.2023 **Revised: 10.03.2023** **Accepted: 30.03.2023**
Поступил: 24.02.2023 **Одобрена: 10.03.2023** **Принята: 30.03.2023**

neural networks. *arXiv preprint arXiv*:1703.04770. 2017 Mar 14. doi: 10.48550/arXiv.1703.04770

10. Gimeno P, Viñals I, Ortega A, et al. Multiclass audio segmentation based on recurrent neural networks for broadcast domain data. *EURASIP Journal on Audio, Speech, and Music Processing*. 2020 Dec;2020:1-9.

11. Dai J, Liang S, Xue W, et al. Long short-term memory recurrent neural network based segment features for music genre classification. *In2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP)* 2016 Oct 17 (pp. 1-5). IEEE. doi: 10.1109/ISCSLP.2016.7918369

12. Zhang Z, Xu S, Zhang S, et al. Attention based convolutional recurrent neural network for environmental sound classification. *Neurocomputing*. 2021 Sep 17;453:896-903. doi: 10.1016/j.neucom.2020.08.069

13. Wang H, Zou Y, Chong D, Wang W. Environmental sound classification with parallel temporal-spectral attention. *arXiv preprint* arXiv:1912.06808. 2019 Dec 14. doi: 10.48550/arXiv.1912.06808

14. Sang J, Park S, Lee J. Convolutional recurrent neural networks for urban sound classification using raw waveforms. *In2018 26th European Signal Processing Conference (EUSIPCO)* 2018 Sep 3 (pp. 2444-2448). IEEE. doi: 10.23919/EUSIPCO.2018.8553247

15. Choi K, Fazekas G, Sandler M, Cho K. Convolutional recurrent neural networks for music classification. *In2017 IEEE International conference on acoustics, speech and signal processing (ICASSP)* 2017 Mar 5 (pp. 2392-2396). IEEE. doi: 10.1109/ICASSP.2017.7952585

16. Gwardys G, Grzywczak D. Deep image features in music information retrieval. *International Journal of Electronics and Telecommunications*. 2014;60:321-6. doi: 10.2478/eletel-2014-0042

17. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*. 2017 May 24;60(6):84-90. doi: 10.1145/3065386

18. Deng J, Dong W, Socher R, et al. Imagenet: A large-scale hierarchical image database. *In2009 IEEE conference on computer vision and pattern recognition* 2009 Jun 20 (pp. 248-255). IEEE. doi: 10.1109/CVPR.2009.5206848

19. Emergency Vehicle Siren Sounds [Internet]. Kaggle [cited 2023 February 23]. Available from: https://www.kaggle.com/vishnu0399/emergency-vehicle-siren-sounds

20. CNN for audio recognition. GitHub [cited 2023 February 23]. Available from: https://github.com/AnLiMan/CNN-for-audio-recognition

Сведения об авторах:
Лисов Андрей Анатольевич, аспирант;
eLibrary SPIN: 1956-3662; ORCID: 0000-0001-7282-8470;
E-mail: lisov.andrey2013@yandex.ru

Кулганатов Аскар Зайдакбаевич, аспирант;
eLibrary SPIN: 7607-9723; ORCID: 0000-0002-7576-7949;
E-mail: kulganatov97@gmail.com

Панишев Сергей Алексеевич, аспирант;
eLibrary SPIN: 2676-5207; ORCID: 0000-0003-2753-2341;
E-mail: panishef.serega@mail.ru

Received: 24.02.2023 **Revised: 10.03.2023** **Accepted: 30.03.2023**
Поступил: 24.02.2023 **Одобрена: 10.03.2023** **Принята: 30.03.2023**

Information about the authors:
Andrey A. Lisov, postgraduate student;
eLibrary SPIN 1956-3662; ORCID: 0000-0001-7282-8470;
E-mail: lisov.andrey2013@yandex.ru

Askar Z. Kulganatov, postgraduate student;
eLibrary SPIN 7607-9723; ORCID: 0000-0002-7576-7949;
E-mail: kulganatov97@gmail.com

Sergei A. Panishev, postgraduate student;
eLibrary SPIN 2676-5207; ORCID: 0000-0003-2753-2341;
E-mail: panishef.serega@mail.ru

Цитировать:
Лисов А.А., Кулганатов А.З., Панишев С.А. Акустическое обнаружение транспортных средств аварийных служб с использованием сверхточных нейронных сетей // Инновационные транспортные системы и технологии. – 2023. – Т. 9. – № 1. – С. 95–107. doi: 10.17816/transsyst20239195-107

To cite this article:
Lisov AA, Kulganatov AZ, Panishev SA. Using convolutional neural networks for acoustic-based emergency vehicle detection. *Modern Transportation Systems and Technologies*. 2023;9(1):95-107. doi: 10.17816/transsyst20239195-107

Received: 24.02.2023 **Revised: 10.03.2023** **Accepted: 30.03.2023**
Поступил: 24.02.2023 **Одобрена: 10.03.2023** **Принята: 30.03.2023**