



тия выполнена попытка извлечения данных из картинок. В качестве сообщения использована простая фраза: «Привет».

После попыток извлечения информации из сжатых изображений получились следующие результаты.

Метод LSB: «дВр\_С\$=...».

Метод PVD: «т2;Sf\*6æ3...».

Метод ДКП: «Привет».

### Итоги исследований

Исходя из полученных результатов наиболее надежным можно считать метод ДКП, так как он обладает устойчивостью к компрессии. Для домашнего использования фаворитом является метод LSB из-за объемов скрываемой информации и простоты реализации.

### Литература

1. Методы компьютерной стеганографии [Электронный ресурс]. – URL: <http://www.vsavm.by/knigi/kniga3/1740.html>.
2. Сейеди, С.А. Сравнение методов стеганографии в изображениях [Текст] / С.А. Сейеди, Р.Х. Садыхов // Информатика. – 2013. – №1. – С. 66-75.
3. Конахович, Г.Ф. Компьютерная стеганография. Теория и практика [Текст] / Г.Ф. Конахович, А.Ю. Пузыренко. – М.: МК-Пресс, 2006. – 288 с.
4. Белим, С.В. Стеганоанализ алгоритма Коха-Жао [Текст] / С.В. Белим, Д.Э. Вильховский // Математические структуры и моделирование. – 2018. – Вып. 4. – С. 113-119.

Р.Ф. Сагитов, И.А. Лёзин

## ИССЛЕДОВАНИЕ СПОСОБОВ ОРГАНИЗАЦИИ АВТОМАТИЗИРОВАННЫХ СИСТЕМ ЗВУКОВОЙ ФИЛЬТРАЦИИ

(Самарский университет)

### Задача распознавания

Фильтрация аудиопотока представляет собой отделение в смешанном потоке одних источников звука от других на базе некоторого множества критериев и преобразование в выходной поток тех звуков, сохранение которых требует исходная задача. Иными словами, в основе стоит задача классификации.

Для нейронной сети фактические критерии отображаются уже множеством признаков, именно обучившись определять признаки, сеть сможет выполнить корректную классификацию [1], в конечном итоге – операции по определению требуемых источников звука.

Главными элементами всех форм представления данных выступают признаки, которые являются наблюдаемыми свойствами объекта. Векторы имеют плоскую и простую структуру и обычно, в большинстве приложений для ма-



шинного обучения, представляют собой объединение данных. В программе, использующей машинное обучение, вектор применяют для измерения свойств данных, таких как цвет, плотность, сила звука, близость к чему-либо, то есть всего того, что можно описать с помощью последовательности чисел – по одному числу для каждого оцениваемого свойства [2].

В качестве примера предположим, что стоит задача фильтрации голосовой информации и что системе нужно выводить звуковую дорожку только с человеческим голосом, без посторонних шумов.

Это говорит в первую очередь о том, что задача классификации становится бинарной – есть отдельные спектральные данные, подлежащие сохранению, когда как все остальные должны быть убраны со спектра.

Спектрограмма преобразованного интервала аудиопотока представляет собой не что иное, как матрицу чисел, явно указывающую на все имеющиеся звуки в интервале; при этом, звуки разного происхождения будут приносить разные состояния на эту матрицу.

### **Обзор классов глубоких нейронных сетей**

Разнообразие структур глубоких сетей довольно широкое. Однако в исследовательских целях следует остановиться на небольшом количестве самых успешно зарекомендовавших себя моделей.

Обзор следует начать со сверточных нейронных сетей (CNN).

Особенностью архитектуры CNN является использование ограниченной матрицы весов для операции “свёртки”. Суть заключается в том, что матрица выполняет движение по своему слою, при этом после каждого сдвига запрашивается нейрон на аналогичной позиции уже в следующем слое. Матрица является “ядром” свёртки, или “фильтром”, а операция свёртки выполняется между тензорами (из двух тензоров на входе получается один на выходе) [3]. Другой важной чертой CNN являются её строительные блоки. Во-первых, это сверточные слои; такой слой принимает на вход тензор, применяет к нему определенное число фильтров, задаёт смещение, запускает функции активации. Сводящие слои идут сразу после сверточных, образуя стек, они как правило требуются для сокращения размерности тензоров. Помимо них в сети участвуют полносвязные слои и слои активации.

В огромном пространстве состояний спектрограммы, нахождение нелинейных зависимостей и признаков в принципе может оказаться неподъемной задачей для данной модели, признаки могут формироваться очень долгое (по меркам обучения сети) время, и, возможно, сети в конечном итоге так и не удастся корректно определить их.

В поисках предотвращения подобных потенциальных проблем можно рассмотреть сети долгой краткосрочной памяти (LSTM). Её формула с вентилями забывания  $c_0 = 0$  и  $h_0 = 0$  (где  $\circ$  – произведение Адамара) выглядит следующим образом [4]:

$$\begin{aligned}f_t &= \sigma_g(W_f x_t + U_f h_{t-1} + b_f), \\i_t &= \sigma_g(W_i x_t + U_i h_{t-1} + b_i),\end{aligned}$$



$$\begin{aligned}o_t &= \sigma_g(W_o x_t + U_o h_{t-1} + b_o), \\c_t &= f_t \circ c_{t-1} + i_t \circ \sigma_c(W_c x_t + U_c h_{t-1} + b_c), \\h_t &= o_t \circ \sigma_h(c_t).\end{aligned}$$

где  $x_t$  – входной вектор,  $h_t$  – выходной вектор,  $c_t$  – вектор состояний,  $W$ ,  $U$  и  $b$  – матрицы параметров и вектор,  $f_t$ ,  $i_t$ ,  $o_t$  – векторы вентилях,  $f_t$  – вектор вентиля забывания, вес запоминания старой информации,  $i_t$  – вектор входного вентиля, вес получения новой информации,  $o_t$  – вектор выходного вентиля, кандидат на выход,  $\sigma_g$  – на основе сигмоиды,  $\sigma_c$  – на основе гиперболического тангенса.

LSTM – модель организации ячеек памяти продолжительного действия, которые способны хранить несколько наборов весов с предыдущих шагов обучения, что в свою очередь способно оказать положительный эффект при обучении сети. LSTM имеет три основных вида узлов: входной гейт (input gate), забывающий гейт (forget gate) и выходной гейт (output gate), а также рекуррентная ячейка со скрытым состоянием [1].

Наличие подобного блока в исследованных моделях явно указывает на то, что полученные путем оконных преобразований Фурье (FFT) идущие друг за другом спектрограммы не должны быть обособлены друг от друга.

В конечном итоге, в качестве исследовательской работы выбрана модель LSTM-сети, поскольку именно такая модель способна гарантированно обеспечить процесс корректного обучения под поставленную задачу.

### Модель автоматизированной информационной системы

Исходя из рассмотренных здесь и ранее материалов, можно сформировать схему организации автоматизированной системы.

Автоматизированная система реального времени по фильтрации цифрового аудиосигнала включает набор блоков, отвечающих за работу по захвату аудиосигнала с записывающего устройства, чтению аудиофайлов, преобразования фреймов в спектральную форму оконными преобразованиями Фурье, а также подсистему формирования отчетов. В ядре системы находится блок глубокой нейронной сети, выполняющий непосредственную фильтрацию от всех шумов и внешних источников звука. На рисунке 1 демонстрируется обобщенная модель создаваемой системы.

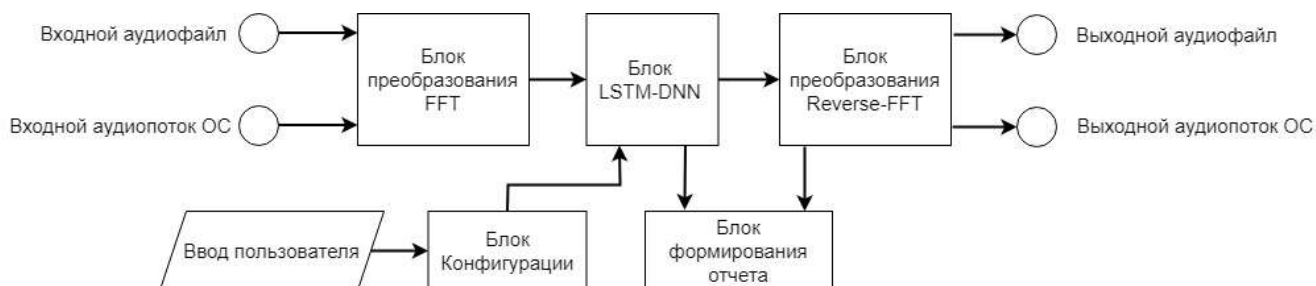


Рис. 1. Обобщенная схема организации демонстрационной автоматизированной системы



## Изучение программных комплексов для организации автоматизированной системы

Реализация потребует следующие программные инструменты:

- Библиотека(-и) организации и обучения глубоких нейронных сетей, в т.ч. сверточных сетей;
- Фреймворк для организации программной оболочки захвата и аудиопотока операционной системы и вывода результата обработки;
- Средства разработки интерфейса системы;
- Средства для формирования графиков аудио обработки и отчета системы.

В таблице 1 указаны сравнительные характеристики исследованных библиотек и фреймворков для работы с моделями глубоких нейронных сетей [5].

Таблица 1 – Сравнительная характеристика библиотек

| Название       | Платформа разработки                   | Поддержка архитектур нейронных сетей | Язык(-и) разработки | Доступный язык интерфейса                        | Дост. Исп. GPU CUDA |
|----------------|--|--------------------------------------|---------------------|--|---------------------|
| Deeplearning4J | Linux, Mac OS X, Windows               | RNN, CNN                             | Java                | Java, Scala, Clojure, Python (Keras)             | +                   |
| Keras          | Linux, Mac OS X, Windows               | RNN, CNN                             | Python              | Python   | +                   |
| TensorFlow     | Linux, Mac OS X, Windows               | RNN, CNN                             | C++, Python         | Python, C/C++, Java, Go                          | +                   |
| Torch          | Linux, Mac OS X, Windows, Android, iOS | RNN, CNN                             | C, Lua              | Lua, LuaJIT, C, библиотека утилит для C++/OpenCL | +                   |

Как видно из сравнения, все эти библиотеки являются полнофункциональными средствами для формирования, обучения и тестирования требуемых архитектур нейронных сетей. Также можно заметить, что все они поддерживают использование GPU ядер CUDA.

### Литература

1. Интеллектуальные системы. Курс лекций. [Текст] / О.П. Солдатова – Самара.: 2019. – 384 с.: ил (дата обращения 05.10.2021).



2. Шакла, Н. Машинное обучение & Tensor Flow [Текст] / Н. Шакла – СПб.: Питер, 2019. – 336 с.: ил (дата обращения 23.09.2021).
3. Микелуччи, У. Прикладное глубокое обучение. Подход к пониманию глубоких нейронных сетей на основе метода кейсов [Текст] / У. Микелуччи – СПб.: БХВ-Петербург, 2020. – 368с.: ил (дата обращения 20.08.2021).
4. Бринк, Х. Ричардс, Дж. Феверолф, М. Машинное обучение [Текст] / Х. Бринк, Дж. Ричардс, М. Феверолф – СПб.: Питер, 2017. – 336 с.: ил (дата обращения 20.10.2021).
5. Сравнительный Анализ Математических Моделей Финансовых Временных Рядов [Текст] / А.А. Садовский, О.П. Солдатова – Самара.: 2021. – 6 с.: ил (дата обращения 07.03.2022).
6. Deep Learning for Java. [Электронный ресурс] // Сайт <https://deeplearning4j.org>. URL: <https://deeplearning4j.org>. Просмотрено: 02.06.2021.

Я.О. Сайгак, И.В. Лёзина

## РАСПОЗНАВАНИЕ ЛАТИНСКИХ БУКВ МНОГОСЛОЙНЫМ ПЕРСЕПТРОНОМ

(Самарский университет)

За недолгое время, прошедшее с начала XXI века, информационные технологии шагнули далеко вперед. Повсеместное распространение электронно-вычислительных машин, с одной стороны, внесло огромный вклад в решение широкого спектра задач прикладного и теоретического характера, но также породило новые, неизвестные или же ранее не рассматриваемые задачи. Одной из актуальных проблем современных информационных технологий в настоящее время остается распознавание образов, в частности графически представленных символов [1].

Таким образом, выбранная тема остается актуальной даже в контексте современных информационных систем. А при дальнейшем развитии такие технологии смогут обеспечивать более качественные и интерактивные модели интерфейсов «человек - электронно-вычислительная машина» [2].

Персептрон – простейший вид нейронной сети. В основе лежит математическая модель восприятия информации мозгом, состоящая из сенсоров, ассоциативных и реагирующих элементов.

Архитектура многослойного персептрона основывается на модели нейронной сети, предложенной Фрэнком Розенблаттом. В её состав входят три основных элемента:

- 1) Входной слой;
- 2) Скрытый слой;
- 3) Выходной слой [3].