



А.Д. Божимов, О.П. Солдатова

РАЗРАБОТКА АРХИТЕКТУРЫ НЕЙРОННОЙ СЕТИ ДЛЯ ЗАДАЧИ ФИЛЬТРАЦИИ ШУМА В ЗВУКОВОМ ПОТОКЕ

(Самарский университет)

Взаимодействие в современном обществе сложно представить без цифровых средств коммуникации. В условиях удаленной работы и обучения особенно актуальными становятся системы фильтрации звукового потока, обеспечивающие четкое звучание голоса и подавление отвлекающего шума. Наличие подобных систем способно значительно улучшить работу средств автоматического распознавания речи, снизить объем передаваемых данных и наконец – обеспечить беспроблемное общение между людьми. Помимо голосовой связи технологии фильтрации звука активно используются в музыкальной индустрии, кинопроизводстве и медицине [1].

Задача фильтрации шума состоит в том, чтобы убрать из исходного сигнала все звуки, не являющиеся человеческой речью: помехи, фоновые и посторонние шумы. С помехами и первым типом шумов довольно просто справиться линейными методами, однако, второй тип шумов содержит в себе неоднородные, проявляющиеся без известной закономерности наборы звуков [2].

Развитие технологий машинного обучения позволило решить эту проблему. Фильтрация звука на их основе использует так называемый «доверительный фильтр». Доверительный фильтр обеспечивает минимально возможный доверительный интервал для каждого исходного звукового кадра из входного звукового потока. Преимуществом данного подхода является самостоятельная «настройка» в каждом конкретном случае, что обеспечивает наилучший возможный результат для каждого участка потока.

Базой алгоритма является метод наименьших квадратов, основанный на аппроксимации последовательности исходных данных линейной комбинацией некоторых функций. Сумма квадратов отклонений исходных данных от аппроксимированных значений при этом минимизируется [3]. В качестве аппроксимирующих функций обычно используются полиномиальные функции.

Одним из вариантов применения нейронных сетей является аппроксимация сложных нелинейных функций. Нейронная сеть с любой нелинейностью остается универсальным аппроксиматором, что доказывается обобщенной аппроксимационной теоремой [4]. Из этого следует актуальность разработки алгоритмов фильтрации звукового потока на базе нейронной сети.

Для реализации была выбрана нейронная сеть Кохонена, способная обеспечить обучение без учителя, когда задаются только сами объекты x_i и выделяются кластеры, соревнующиеся за право присоединить объект x [5].

Обучение модели нейронной сети требует некоторой подготовки входных данных, а также определения конфигурации. Структура всей подсистемы обучения представлена на рисунке 1.

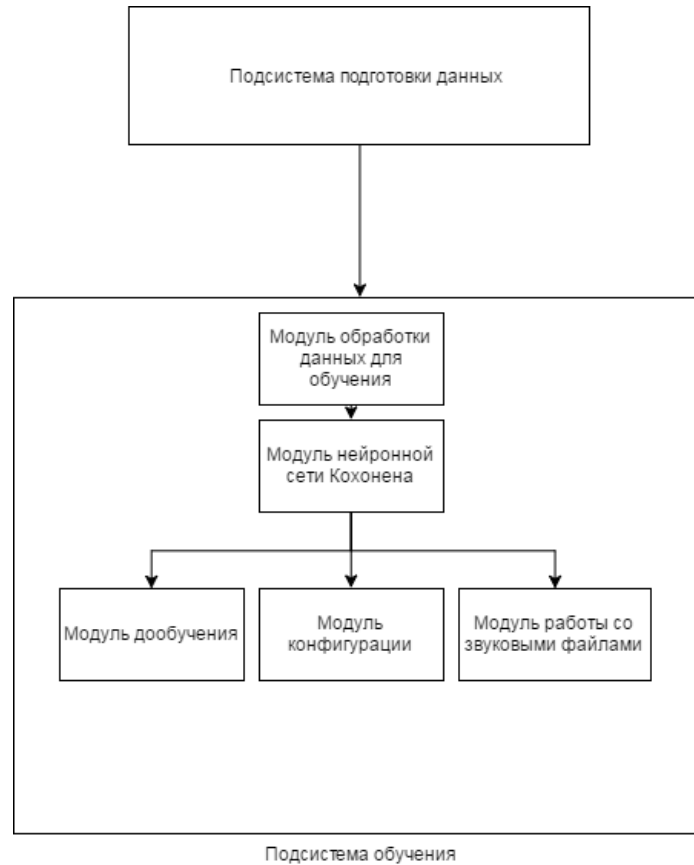


Рис. 1. Структура подсистемы обучения

Топология модели нейронной сети представлена рисунке 2.

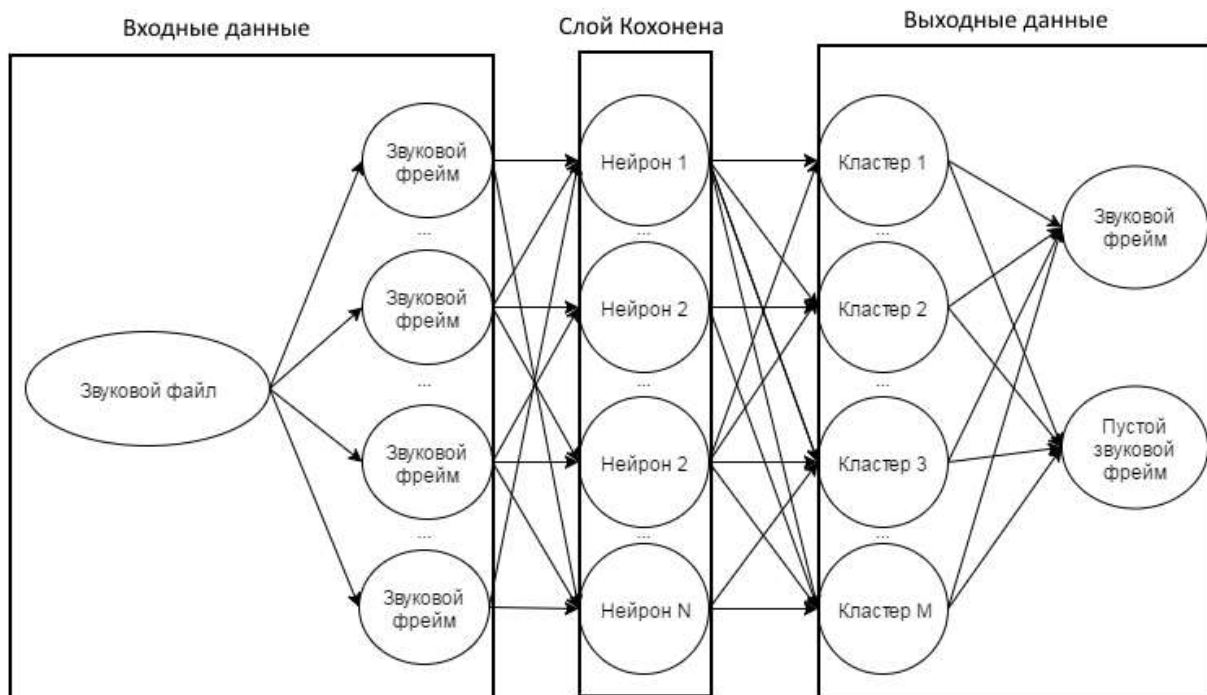


Рис. 2. Топология нейронной сети

На рисунке 2 используются следующие обозначения:

N – число нейронов в слое Кохонена,

M – число кластеров.



В тестовом примере сеть является однослойной и состоит из нейронного слоя Кохонена. Результатом её работы является кластеризованный набор звуковых кадров, а выходными данными – номера кластеров, к которым относятся конкретные кадры.

Для каждого нейрона входным сигналом является нормализованная часть звукового кадра, представленная вещественным числом от 0 до 1. Веса нейрона представлены вектором вещественных чисел w_i , соответствующим по размерности количеству признаков для звукового кадра. Корректировка весов в процессе обучения выполняется при помощи Евклидовой метрики, а функцией активации служит SoftMax [6].

В рамках данной работы была разработана программная система на языке C#, обеспечивающая высокий уровень интеграции с работой операционной системы Windows, что в перспективе позволит выполнять процесс очистки звукового потока «на лету» [7].

Обучающая выборка представлена подготовленным набором записей речи и шума в формате MP3 с частотой дискретизации 44100 Гц.

Результат одной из тестируемых записей в виде спектрограммы представлен на рисунке 3. Как видно, исходный файл (левая половина) в значительной степени зашумлен, однако, результирующий поток показывает четкие границы звуков.

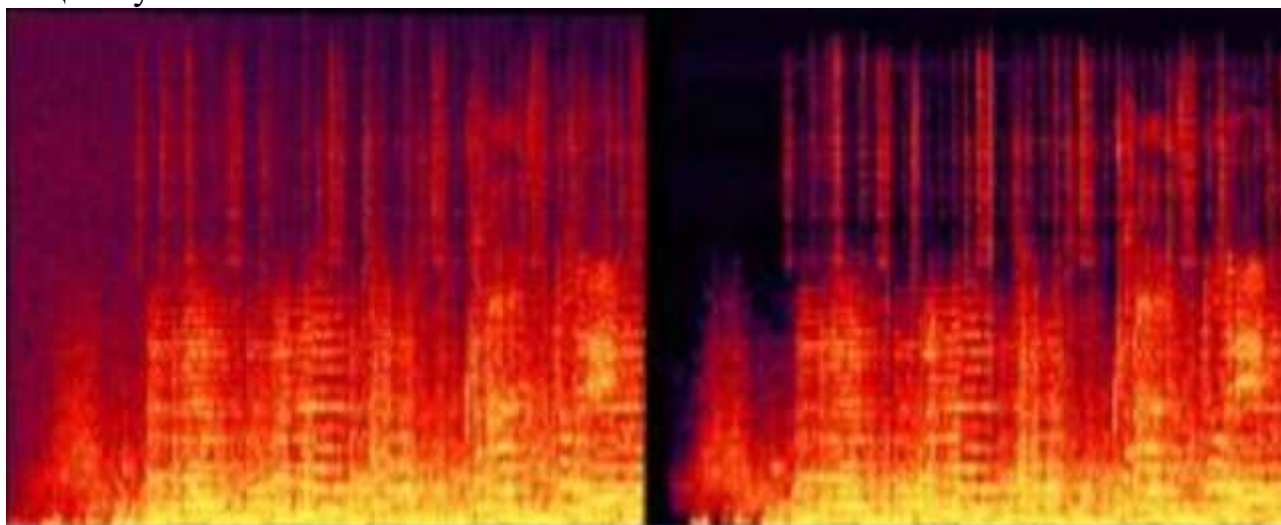


Рис. 3. Результат работы модели

Анализ полученных результатов позволил подтвердить предположение об актуальности использования модели нейронной сети для фильтрации звукового потока, а также доказал эффективность использования нейронной сети Кохонена для решения поставленной задачи.

Литература

1. Noise Suppression for Medical Devices [Электронный ресурс]. URL: <https://www.murata.com/en-global/products/emc/emifil/overview/medical> (дата обращения: 25.03.2022).



2. Обзор методов улучшения речи и шумоподавления [Электронный ресурс]. URL: https://habr.com/ru/company/ru_mts/blog/584308/ (дата обращения: 28.03.2022)

3. Метод наименьших квадратов и основы теории обработки наблюдений [Текст]/ Линник Юрий – М.: Государственное издательство физико-математической литературы”. 1962. – 352 с.

4. Аппроксимационная теорема Вейерштрасса [Электронный ресурс]. URL: https://math.fandom.com/ru/wiki/Аппроксимационная_теорема_Вейерштрасса/ (дата обращения: 12.03.2022)

5. Нейронные сети Кохонена [Электронный ресурс]. URL: <https://neuronus.com/theory/nn/955-nejronnye-seti-kokhonena.html> (дата обращения: 02.04.2022)

6. Софтмакс (Softmax) - Машинное обучение доступно [Электронный ресурс]. URL: <https://www.helenkapatsa.ru/softmaks/> (дата обращения: 02.03.2022)

7. Everything you need to know about C# | Pluralsight [Электронный ресурс]. URL: <https://www.pluralsight.com/blog/software-development/everything-you-need-to-know-about-c-/> (дата обращения: 15.03.2022)

В.Д. Гижевская, Д.С. Оплачко

РАЗРАБОТКА КЛАССИФИКАТОРА ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ

(Самарский университет)

Введение

Человек около 90% информации об окружающем мире получает благодаря зрению. В сфере компьютерных технологий в качестве источника информации может выступать текст, аудио-/видеофайл или изображение. Очень часто, а в основном в сети Интернет, возникает необходимость найти то или иное изображение. В последние годы фотография стала популярным увлечением среди людей благодаря доступу к аппаратному обеспечению, как минимум, практически в каждый телефон встроена камера. Фотографии имеют свойство накапливаться и с течением времени может усложниться поиск по каталогу снимков.

Классификация изображений – одна из наиболее важных задач компьютерного зрения [1]. Проще говоря, она состоит в том, чтобы дать вам изображение, содержащее определённый объект, что позволит вам определить, что этот объект содержит на этом изображении.

Например, по фотографии кошки мы можем легко идентифицировать её как кошку. Это потому, что наш мозг много узнал об этом. Если компьютеру разрешено распознавать это изображение, оно будет совершенно другим, потому что компьютер и то, что мы видим, совершенно разные [2].