ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ

«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет компьютерных наук

Департамент программной инженерии

СОГЛАСОВАНО

Старший преподаватель факультета компьютерных наук департамента программной инженерии \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Д. В. Пантюхин

«\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2016 г.

УТВЕРЖДАЮ

Академический руководитель образовательной программы   
«Программная инженерия»

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ В. В. Шилов

«\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2016 г.

Программа нейросетевой шумоочистки звукового сигнала.

Program for Neural Network Sound Noise Reduction.

|  |  |
| --- | --- |
| Подп. и дата |  |
| Инв. №  дубл. |  |
| Взам. инв. № |  |
| Подп. и дата |  |
| Инв. № подл. |  |

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

ЛИСТ УТВЕРЖДЕНИЯ

RU.17701729.503200-01 81 01-1-ЛУ

Выполнил

Студент группы бПИ141 НИУ ВШЭ

\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Н. Д. Мелентьев

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2016 г.

2016

УТВЕРЖДЕНО

RU.17701729.503200-01 81 01-1-ЛУ

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ «ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет Компьютерных наук

Департамент Программной Инженерии

Программа нейросетевой шумоочистки звукового сигнала,

Program for Neural Network Sound Noise Reduction

|  |  |
| --- | --- |
| Подп. и дата |  |
| Инв. №  дубл. |  |
| Взам. инв. № |  |
| Подп. и дата |  |
| Инв. № подл. |  |

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

ТИТУЛЬНЫЙ ЛИСТ

RU.17701729.503200-01 81 01-1

Листов 13

2016

СОДЕРЖАНИЕ

1. Введение............................................................................................................. 3
   1. Наименование программы.......................................................................... 3
   2. Основания для разработки......................................................................... 3
2. Назначение разработки..................................................................................... 4
   1. Функциональное назначение..................................................................... 4
   2. Эксплуатационное назначение.................................................................. 4
3. Технические характеристики........................................................................... 5
   1. Постановка задачи...................................................................................... 6
   2. Описание алгоритмов................................................................................. 7
   3. Описание и обоснование выбора состава технических и программных средств.......................................................................................................... 7
      1. Состав технических и программных средств.................................. 7
      2. Обоснование выбора технических и программных средств….......9
4. Ожидаемые технико-экономические показатели........................................... 9
   1. Предполагаемая потребность..................................................................... 9
   2. Экономические преимущества разработки по сравнению с отечественными и зарубежными образцами или аналогами................ 10
5. Источники, использованные при разработке............................................... 12

1. ВВЕДЕНИЕ

1.1 Наименование программы

Наименование программы – «Программа нейросетевой шумоочистки звукового сигнала»

Условное наименование программы – «Noiser»

1.2 Основание для разработки

Разработка ведется на основании приказа Национального исследовательского университета "Высшая школа экономики" № 6.18.1-02/1112-19 от 11.12.2015

2. НАЗНАЧЕНИЕ РАЗРАБОТКИ

2.1. Функциональное назначение

Очистить звук от шума, наложенного на исходный аудио сигнал с помощью нейросетей и различных алгоритмов распространения ошибки.

2.2. Эксплуатационное назначение

Предоставить будущим исследователям в данной области инструмент, позволяющий составлять и тренировать различные виды нейросетей, проверять их на тестовых сигналах и сохранять успешные экземпляры.

На просторах интернета довольно легко найти примеры использования различных вариаций нейросетей для решения большого круга задач, в том числе и для распознавания или очистки изображений.

Однако, с аудио все немного сложнее. Можно найти статьи с теоретическим описанием структуры, которая, если ее построить, может и очистить, а может и не очистить, звуковой сигнал. Работающих программ в близко обозримом виртуальном пространстве так же не нашлось.

3. ТЕХНИЧЕСКИЕ ХАРАКТЕРИСТИКИ

3.1. Постановка задачи

3.1.1 Подготовка данных

Необходимо подобрать веса нейросети таким образом, чтобы входные данные, после прохода через нее, были максимально приближены к тем, которые нам нужны.

Сначала происходит тренировка нейросети на уже известных данных. Допустим, у нас есть звуковая дорожка с «чистым» звуком длиной в 3 секунды. В работе используется частота записи ~20 kHz. Примерный объем выборки для тренировки ~ 60 000 единиц. Далее необходимо добавить «шум»[13]. Происходит это по формуле

noisy\_Speech = speech + noise \* fraction

, где fraction – выбранная пользователем доля шума в аудио.

В зависимости от структуры нейросети создаются входные и выходные данные. Например, если структура сети {20, 18, 9, 1}, это значит, что с помощью 20 входных единиц предсказывается 1 выходная. Поэтому из speech мы берем 20 единиц, из noisy\_Speech – 1. 18 и 9 – количество внутренних нейронов нейросети (hidden neurons), которые помогают в обучении.

Структура нейросети подбирается экспериментально для наилучшей очистки звука, о чем будет сказано позднее.

Каждый нейрон, кроме входных (у нас это 20-ти нейроновая прослойка), обладает матрицей весов и функцией активации.

Функция активации – сигмоида с параметром alpha, который указывает пользователь.

Матрица весов – соединяет нейрон со всеми нейронами следующего уровня нейросети. Для одного нейрона получается вектор весов, для уровня нейросети – матрица весов.

3.1.2 Обучение нейросети

В нейросеть поступают входные данные (в нашем случае 20 единиц), далее с помощью матрицы весов и функций активации на каждом уровне высчитывается конечный результат.

Получается некая разница между полученным результатом yj и тем, который у нас записан в выходных данных dj. Поэтому необходимо изменить веса нейросети таким образом, чтобы сократить эту разницу.

Это можно сделать *методом обратного распространения ошибки.*

Стоит отметить, что в работе использовались сторонние библиотеки для построения сети и ее обучения (Accord, Aforge), где реализация распространения ошибки скрыта от пользователя, но на просторах интернета можно с легкостью найти алгоритмы и попробовать их запрограммировать, что я и сделал на ранних этапах работы.

Например, общий вид работы Backpropagation Neural Network[10]:

Веса изменяются по следующей формуле

, wij – значение из матрицы весов для нейрона i, соединённого с нейроном j,

– скорость обучения нейросети, указывается пользователем,

– высчитывается следующим образом:

, для последнего уровня (выходного) нейросети

, для других. Xj – выход текущего нейрона, \* wjk – взвешенная сумма ошибок с верхнего уровня нейросети.

Xi – входной элемент I к текущему нейрону.

Таким образом, итерационно, достигается наименьшее отклонение от тренировочных данных.

Далее необходимо проверить сеть на тестовой выборке, и пробовать другие вариации для достижения лучших результатов.

3.2. Описание алгоритмов

В работе были испробованы различные структуры сети, и различные алгоритмы обучения сетия, такие как BackPropagaion, Levenberg-Marquadt, ParalelResilientBakpPropagaion [11]. Как уже было сказано выше, их реализация скрыта в библиотеках, поэтому их описание не прилагается.

3.3 Описание и обоснование выбора состава технических и программных средств

3.3.1 Состав технических и программных средств

Для работы программы необходим следующий состав технических средств:

* персональный компьютер, оснащенный 32-разрядным (x86) или 64-разрядным (x64) процессором Pentium с тактовой частотой 400 MГц и выше или аналогичный процессор (рекомендуется Pentium с тактовой частотой 1 ГГц и выше или аналогичный процессор);
* рекомендуется 256 МБ оперативной памяти или больше;
* не менее 1,5 ГБ свободного места на жестком диске;
* Монитор и видеоадаптер Super VGA с минимальным разрешением 800 X 600;
* совместимое указывающее устройство;
* клавиатура;
* звуковая плата;
* мышь.

Для работы программы необходим следующий состав программных средств:

* операционная система Microsoft Windows XP SP3 или более поздняя версия;
* установленный Microsoft .NET Framework 2.0, требующий Windows Installer 3.1 или более поздняя версия.

3.3.2 Обоснование выбора технических и программных средств

Требование Microsoft .Net Framework 2.0 или более поздних версий обусловлено тем, что используемые в программе пространства имен System.Collections.Generic впервые появились именно в версии 2.0.

Все остальные требования – минимальные, которые должен иметь любой компьютер в наше время.

Для комфортной работы с большими нейросетями требуется более серьезное программное обеспечение.

4. ОЖИДАЕМЫЕ ТЕХНИКО-ЭКОНОМИЧЕСКИЕ ПОКАЗАТЕЛИ

4.1. Предполагаемая потребность

Шумоподавление является серьезным вопросом для машинного обучения, поэтому он будет открыт еще долгое время. Распознавание речи, синтез речи улучшения качества мобильной связи и многое другое опирается на предположение о том, что сигнал должен быть не зашумлен. Поэтому такой инструмент необходим для практической проверки теоретических выкладок. С необходимыми добавлениями и улучшениями он может стать незаменимым помощником исследователя в данной области.

4.2. Экономические преимущества разработки по сравнению с отечественными и зарубежными образцами или аналогами

Не найдено аналогов программы на просторах интернета. Многие исследователи вынуждены с нуля создавать программы для проверки своих гипотез, что замедляет и затрудняет сам процесс.

5. ИСТОЧНИКИ, ИСПОЛЬЗУЕМЫЕ ПРИ РАЗРАБОТКЕ

1. ГОСТ 19.101-77 Виды программ и программных документов. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.

2. ГОСТ 19.102-77 Стадии разработки. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.

3. ГОСТ 19.103-77 Обозначения программ и программных документов. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.

4. ГОСТ 19.104-78 Основные надписи. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.

5. ГОСТ 19.105-78 Общие требования к программным документам. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.

6. ГОСТ 19.106-78 Требования к программным документам, выполненным печатным способом. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.

7. ГОСТ 19.404-79 Пояснительная записка. Требования к содержанию и оформлению. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.

8. ГОСТ 19.603-78 Общие правила внесения изменений. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.

9. ГОСТ 19.604-78 Правила внесения изменений в программные документы, выполненные печатным способом. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.

10. An analysis of noise reduction using back-propagation neural networks. . [Электронный ресурс]//URL: http://www.dtic.mil/dtic/tr/fulltext/u2/a203057.pdf (Дата обращения: 15.03.2016, режим доступа: свободный).

11Accord documentation. [Электронный ресурс]//URL: <http://accord-framework.net/> (Дата обращения: 25.04.2015, режим доступа: свободный).

12. Aforge documentation. [Электронный ресурс]//URL: <http://www.aforgenet.com/framework/documentation.html> (Дата обращения: 12.01.2016, режим доступа: свободный).

13. Noise database. [Электронный ресурс]//URL: <http://spib.linse.ufsc.br/noise.html> (Дата обращения: 28.10.2015, режим доступа: свободный).

14. Removing noise from speech signals using different approaches of artificial neural networks. [Электронный ресурс]//URL: <http://www.mecs-press.org/ijitcs/ijitcs-v7-n7/IJITCS-V7-N7-2.pdf> (Дата обращения: 20.02.2016, режим доступа: свободный).

15. Recurrent Neural Networks for Noise Reduction in Robust ASR. [Электронный ресурс]//URL: <http://www1.icsi.berkeley.edu/~vinyals/Files/rnn_denoise_2012.pdf> (Дата обращения: 21.02.2016, режим доступа: свободный).

ПРИЛОЖЕНИЕ 1

ТЕРМИНОЛОГИЯ

Список терминов, в которых следует разбираться для работы с программой:

SampleSize – структура нейросети, в нашем примере {20, 18, 9, 1}

Alpha – коэффициент сигмоидной функции

Iterations – количество итераций в обучении нейросети

Learning rate – скорость обучения нейросети

Delimiter – на что поделить входные данные, нужно для нормализации.

Shuffle – как часто менять наложенный шум

Save Net – как часто сохранять нейросеть при обучении

Speech – чистый звук

Noise – шум, наложенный на звук

Save session – сохранить некоторые параметры программы

ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Изм . |  | Номера листов (страниц) | | | Всего листов  (страниц)  в  документе | №  документа | Входящий  №  сопроводительного  документа и дата | Подпись | Дата |
| Изменен-ных | Заменен-ных | Новых | Аннулиро-ванных |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |