

Министерство образования Республики Беларусь

Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.42:534.2-047.27

Петровец
Владислав Николаевич

Алгоритмы и программное средство распознавания звуков
окружающей среды

АВТОРЕФЕРАТ

на соискание степени магистра

по специальности 1-40 80 04 – Информатика и технологии
программирования

Научный руководитель
Батура М.П.
доктор технических наук,
профессор

Минск 2024

ВВЕДЕНИЕ

Звуковая окружающая среда – это комплексный набор звуков, которые окружают нас в повседневной жизни и предоставляют ценную информацию о нашем окружающем мире. Исследования в области распознавания звуков окружающей среды призваны расширить наши знания о том, какие звуковые события происходят в окружающей нас среде, и как эти события могут влиять на нашу жизнь. Путем анализа звуковых сигналов и выявления их особенностей и закономерностей исследователи стремятся создать инструменты и методы, которые позволят автоматизировать процесс распознавания звуковых явлений. В данной работе представлен алгоритм и программное средство, нацеленные на распознавание звуков окружающей среды с особым вниманием к голосам птиц.

Выбор звуков птиц в качестве объекта исследования обусловлен их многогранной значимостью. Прежде всего, птицы являются неотъемлемой частью экосистемы, играя ключевую роль в ее балансе и поддержании биоразнообразия. Их голоса являются важным индикатором состояния окружающей среды, отражая изменения, связанные с климатом и экологическими факторами. Однако, помимо функциональной значимости, разнообразие звуков и уникальные голосовые характеристики различных видов птиц делают их звуки объектом важного исследования. Вариативность мелодий и многообразие звуковых шаблонов предоставляют огромный массив данных, который может быть использован для создания высокоэффективных алгоритмов распознавания. Это в свою очередь позволяет не только отслеживать популяции птиц, но и получать ценные сведения о состоянии окружающей среды в целом, делая выбор звуков птиц в качестве объекта исследования оправданным и перспективным.

Распознавание звуков птиц имеет свои особенности и проблемы. Среди особенностей можно выделить высокую разнообразность птичьих звуков, уникальные голосовые характеристики различных видов, а также вариативность мелодий и звуковых шаблонов, что предоставляет богатый материал для анализа и классификации. Основные проблемы включают наличие антропогенных шумов и других звуковых помех, сложность в различении звуков близкородственных видов, а также необходимость большого количества помеченных данных для обучения моделей. Эти факторы могут существенно влиять на точность и надежность распознавания птичьих звуков, требуя применения передовых методов машинного обучения и тщательной обработки данных.

Предложенное программное средство и алгоритм распознавания звуков окружающей среды могут иметь практическое применение в различных

областях, включая экологический мониторинг, изучение биоразнообразия и научные исследования.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цели и задачи исследования

Целью данной работы является разработка эффективного алгоритма и программного средства для автоматического распознавания звуков окружающей среды с уклоном на звуки птиц.

Для достижения поставленной цели необходимо решить следующие задачи:

- 1 Выбрать область распознавания звуков окружающей среды.
- 2 Провести обзор существующих методов распознавания звуков и техник машинного обучения.
- 3 Собрать и подготовить набор данных для обучения моделей.
- 4 Разработать и реализовать архитектуру нейронной сети для распознавания звуков.
- 5 Обучить модель на подготовленных данных и провести эксперименты для оценки ее производительности.
- 6 Разработать веб-приложение для интеграции обученной модели.
- 7 Интегрировать обученную модель в мобильное приложение.

Объектом исследования является процесс распознавания звуков окружающей среды с использованием модели глубокого обучения.

Предметом исследования является разработка и анализ эффективности модели распознавания звуков на основе архитектуры моделей глубокого обучения.

Модель на основе архитектуры моделей глубокого обучения способна эффективно и точно распознавать звуки окружающей среды, что делает ее привлекательной для применения в различных областях, таких как мониторинг экологической обстановки, поддержка городских служб и разработка приложений для анализа звукового ландшафта.

Эта гипотеза предполагает, что модель обладает достаточной обобщающей способностью для успешного распознавания звуков окружающей среды в различных условиях, что делает ее полезной и применимой в практических задачах.

Личный вклад соискателя

Соискателем выполнены все изложенные исследования и разработки. Постановка задач и обсуждение результатов проводились совместно с научным руководителем и сотрудниками кафедры Информатики Белорусского государственного университета информатики и радиоэлектроники. Соавторы опубликованных работ принимали участие в

обсуждении промежуточных и конечных результатов. Обработка, интерпретация данных, а также выводы сделаны автором самостоятельно.

Апробация результатов диссертации

Результаты исследований апробированы на практике и были доложены на следующих научно-практических конференциях:

- 59-я научная конференция аспирантов, магистрантов и студентов БГУИР;
- 60-я научная конференция аспирантов, магистрантов и студентов БГУИР.

Публикации результатов диссертации

По теме диссертации опубликовано 3 печатных работы в сборниках трудов и материалов конференций.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, трех глав, заключения, списка использованных источников, списка публикаций автора и приложений.

Первая глава представляет собой обзор предметной области, включающий в себя анализ существующих методов и подходов к распознаванию звуков.

Вторая глава посвящена методологии, в которой формулируются постановка задачи, описываются выбранные методы и подходы, а также подробно описывается выбранный набор данных.

В третьей главе рассматривается реализация предложенной модели, включая архитектуру нейронной сети, процесс обучения, использование техник предобработки данных и реализация веб-приложение для интеграции обученной модели.

Общий объем работы составляет 78 страниц.

Проверка на уникальность

Проведена экспертиза диссертации Петровца Владислава Николаевича «Алгоритмы и программное средство распознавания звуков окружающей среды» на корректность использования заимствованных материалов с применением сетевого ресурса «Антиплагиат» (адрес доступа: <https://antiplagiat.ru>) в on-line режиме 12.06.2024 г. В результате проверки

установлена корректность использования заимствованных материалов (оригинальность диссертационной работы составляет 86,44 %).

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Введение включает в себя вступительные разделы работы. В ней обосновывается выбор темы исследования, описываются мотивации исследователя, а также выдвигаются цели и задачи работы. Введение представляет контекст и значимость выбранной темы, объясняет, почему она является актуальной и важной для изучения. Кроме того, устанавливаются цели исследования, а также формулируются конкретные задачи, направленные на достижение этих целей.

Первая глава представляет собой обзор исследований в области распознавания звуковых сигналов, начиная с исторического развития этой области. В ней рассматриваются проблемы и вызовы, с которыми сталкиваются исследователи при распознавании звуков, а также описываются традиционные методы мониторинга птиц, включая визуальные наблюдения, использование сетей для отлова и маркировки, и методы аудиозаписи. Особое внимание уделено технологическим возможностям для автоматизации мониторинга, таким как использование машинного обучения и акустических методов. Примеры существующих технологий и программных решений иллюстрируют текущие достижения и потенциал в области автоматизированного распознавания звуков. Глава служит основой для понимания контекста и современных тенденций в этой научной области.

Вторая глава посвящена постановке задачи распознавания звуков птиц и анализу современных подходов к ее решению. Сначала формулируются основные цели и задачи распознавания звуков птиц. Затем подробно рассматриваются различные методы и модели, используемые в этой области, включая скрытые марковские модели, сверточные нейронные сети, а также рекуррентные нейронные сети и сети долгосрочной и краткосрочной памяти. Особое внимание уделяется спектральному анализу как важному инструменту для обработки и анализа звуковых данных. Глава предоставляет всесторонний обзор текущих методик и их применимости к задаче распознавания птичьих звуков.

Третья глава посвящена практическим аспектам создания и использования системы распознавания звуков птиц. В начале главы рассматриваются процессы создания набора данных, включая поиск и выбор источников данных, их компоновку и предобработку. Далее приводится описание модели сверточной нейронной сети, используемой для распознавания звуков. Особое внимание уделяется разработке бэкенда для интеграции обученной модели, а также созданию мобильного приложения,

позволяющего использовать обученную модель для распознавания звуков птиц в реальном времени. Глава предоставляет комплексный обзор всех этапов, начиная от подготовки данных до конечной реализации программного обеспечения.

ЗАКЛЮЧЕНИЕ

Выявление достижений и проблем, возникших в процессе работы над проектом по распознаванию звуков птиц, поможет лучше понять и оценить результаты исследования.

Достижения:

1 Разработка рабочей модели. Разработана рабочая модель распознавания звуков птиц на основе сверточной нейронной сети, что позволяет автоматизировать процесс распознавания звуков птиц по аудиозаписи.

2 Успешное обучение модели на большом наборе данных. Удалось обучить модель на большом наборе данных с точностью 81,9%.

3 Интеграция обученной модели в мобильное-приложение. Модель была успешно интегрирована в мобильное приложение.

Проблемы:

1 Ограниченность набора данных. Одной из проблем является ограниченность набора данных звуков птиц. Недостаточное количество обучающих данных для некоторых видов птиц может привести к низкой точности распознавания.

2 Несовершенство набора данных. Наличие различных антропогенных шумов, таких как шум от транспорта, строительные работы и другие человеческие активности, которые могут загрязнять аудиозаписи. Эти шумы затрудняют точное распознавание птичьих пений, создавая ложные срабатывания и снижая общую точность модели. Кроме того, записи могут содержать пения других видов птиц, что также усложняет процесс идентификации и может приводить к ошибкам в классификации.

3 Требования к вычислительным ресурсам. Использование сверточной нейронной модели может потребовать значительных вычислительных ресурсов, особенно при работе с большими объемами данных или в реальном времени.

Идентификация достижений и проблем помогает определить ключевые аспекты, которые следует учитывать при дальнейшем развитии и улучшении модели распознавания звуков птиц.

Для будущих исследователей в области распознавания звуков птиц, следует учитывать следующие рекомендации:

1 Расширение набора данных. Сбор качественного и разнообразного набора данных является ключевым для успешного обучения модели. Будущим исследователям следует уделить достаточное внимание сбору данных, включая аудиозаписи других видов птиц, не затронутых в рамках данного исследования.

2 Исследование прочих архитектур. Помимо сверточных нейронных сетей, в данной работе были изучены прочие подходы для разработки средства классификации птиц по их звукам.

3 Учет несовершенности набора данных. Процесс распознавания звуков птиц может быть значительно улучшено путем избавления исходного набора данных от шумов.

С учетом этих рекомендаций будущие исследователи могут эффективно развивать область распознавания звуков птиц.

Подведение итогов исследования по распознаванию звуков птиц включает следующие ключевые моменты:

1 Разработка модели. Была разработана модель на основе сверточной нейронной сети для распознавания звуков птиц. Эта модель обладает способностью работать на мобильных устройствах и встраиваемых системах благодаря своей легкой архитектуре.

2 Сбор данных. Набор данных был собран и структурирован для 50 видов птиц, имея по 3000 аудиозаписей удовлетворительного качества для каждого вида.

3 Предобработка данных. Каждая аудиозапись была сконвертирована в двумерный массив mel-коэффициентов размером 24 на 512 значений.

4 Обучение модели. Модель была обучена на предобработанном наборе данных. Обучение включало в себя настройку гиперпараметров модели, выбор оптимизатора и функции потерь, а также оптимизацию процесса обучения.

5 Оценка результатов. Результаты работы модели были оценены с использованием метрик качества, таких как точность и функция потерь. Точность модели составляет 81,9% на тестовом наборе данных, а функция потерь составила 10,2.

6 Дальнейшие направления исследований. Были выделены направления дальнейших исследований и улучшений, включая улучшение качества классификации, оптимизацию скорости работы модели, расширение набора данных видов птиц.

Итоги исследования подчеркивают значимость и потенциал модели распознавания звуков птиц в различных областях, связанных с орнитологией. Дальнейшее развитие и улучшение этой модели может привести к созданию новых инновационных приложений и сервисов, а также к расширению возможностей анализа фауны птиц.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

1-А. Петровец В. Н. Использование алгоритма MFCC в задаче распознавания музыкальных композиций / В. Н. Петровец, А. А. Раловец, Д. В. Деменковец // Компьютерные системы и сети: сборник статей 59-й научной конференции аспирантов, магистрантов и студентов, Минск, 17–21 апреля 2023 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2023. – С. 412–413.

2-А. Петровец В. Н. Использование wavelet-преобразования для решения задачи распознавания птиц / В. Н. Петровец, П. Л. Мискевич, Д. В. Деменковец // Компьютерные системы и сети: сборник статей 60-й научной конференции аспирантов, магистрантов и студентов, Минск, март 2024 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2023.

3-А. Мискевич, П. Л. Развитие и оценка нейросетевых методов верификации собственноручной подписи / П. Л. Мискевич, В. Н. Петровец, А. А. Раловец // Компьютерные системы и сети : сборник статей 59-й научной конференции аспирантов, магистрантов и студентов, Минск, 17–21 апреля 2023 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2023. – С. 365–369.