

**Заключение диссертационного совета Д 01.024.04 на базе  
ГОСУДАРСТВЕННОГО ОБРАЗОВАТЕЛЬНОГО УЧРЕЖДЕНИЯ ВЫСШЕГО  
ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ  
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ» и ГОСУДАРСТВЕННОГО  
ОБРАЗОВАТЕЛЬНОГО УЧРЕЖДЕНИЯ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО  
ОБРАЗОВАНИЯ «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ»  
Министерства образования и науки Донецкой Народной Республики  
по диссертации на соискание ученой степени кандидата технических наук  
аттестационное дело № \_\_\_\_\_  
решение диссертационного совета Д 01.024.04 от 12.10.2021 г. протокол №19/21**

**О ПРИСУЖДЕНИИ  
Пикалёву Ярославу Сергеевичу  
ученой степени кандидата технических наук**

Диссертация «Совершенствование методов и программных средств распознавания слитной русской речи» по специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки) принята к защите «22» июня 2021 г. диссертационным советом Д 01.024.04 (протокол № 14/21) на базе ГОУВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ» и ГОУВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ», 283001, г. Донецк, ул. Артема, 58, корп. 1, ауд. 203 Тел./факс: 380(62) 304-30-55, e-mail: [uchensovnet@donntu.org](mailto:uchensovnet@donntu.org) (приказ о создании диссертационного совета № 802 от 20.09.2018 г., приказы об изменении состава диссертационного совета № 1743 от 09.12.2019 г. и №1550 от 08.12.2020 г.).

Соискатель, Пикалёв Ярослав Сергеевич, 1993 года рождения, в 2015 году окончил ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ по специальности «Автоматизированное управление технологическими процессами» с присвоением степени специалиста. Работает научным сотрудником отдела распознавания речевых образов ГОСУДАРСТВЕННОГО УЧРЕЖДЕНИЯ «ИНСТИТУТ ПРОБЛЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА».

Диссертация выполнена в ГОСУДАРСТВЕННОМ УЧРЕЖДЕНИИ «ИНСТИТУТ ПРОБЛЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА».

Научный руководитель: Ермоленко Татьяна Владимировна, кандидат технических наук, доцент, доцент кафедры компьютерных технологий ГОУ ВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ».

**Официальные оппоненты:**

1. ХАРЛАМОВ АЛЕКСАНДР АЛЕКСАНДРОВИЧ, доктор технических наук, старший научный сотрудник ФГБУН «Институт высшей нервной деятельности и нейрофизиологии Российской Академии Наук», г. Москва.

2. БУРЛАЕВА ЕКАТЕРИНА ИГОРЕВНА, кандидат технических наук, специалист 1 категории отдела сетевых сервисов дирекции по информационным технологиям ГУП ДНР «ЭНЕРГИЯ ДОНБАССА», г. Донецк.

Официальные оппоненты дали положительные отзывы на диссертацию.

**Ведущая организация** – Государственное учреждение «Институт прикладной математики и механики» (ГУ ИПММ), г. Донецк, в своем положительном заключении, подписанном директором института, доктором физико-математических наук, старшим научным сотрудником Судаковым С.Н., указала, что работа посвящена актуальной теме, имеет научную новизну, практическое значение и реализована на практике. Обоснованность научных выводов и рекомендаций автора не вызывает сомнений. Научные выводы и рекомендации достаточно полно изложены в опубликованных статьях. Представленная диссертационная работа отвечает требованиям п.2.2 «Положения о присуждении ученых степеней», предъявляемым к кандидатским диссертациям и требованиям паспорта специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки), в частности: п.5 «Разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений и обработки информации»; п.12 «Визуализация, трансформация и анализ информации на основе компьютерных методов обработки информации», а ее автор – Пикалёв Ярослав Сергеевич заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки).

Выбор официальных оппонентов и ведущей организации обосновывается их компетентностью в области научно-практических исследований по созданию средств и систем интеллектуализации компьютерных интерфейсов, в частности, по разработке аппаратно-программного комплекса распознавания русской речи, а также наличием публикаций в соответствующей сфере исследований.

Соискатель имеет 17 опубликованных научных работ, 7 из них в рецензируемых научных изданиях: в том числе 5 – в рецензируемых научных журналах и изданиях, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата и доктора наук в Донецкой Народной Республике и Российской Федерации, 10 – по материалам научных конференций.

#### **Наиболее значимые работы по теме диссертации:**

1. Пикалёв, Я. С. Разработка системы автоматического распознавания слитной русскоязычной речи на основе дискриминативного обучения // Информатика и кибернетика. – Донецк: ДонНТУ. – 2018. – №3(13). – С. 61-68.
2. Пикалёв, Я.С. Технология повышения робастности акустической модели в задаче распознавания речи / Я.С. Пикалёв, Т.В. Ермоленко // Известия ЮФУ. Технические науки. – Ростов-на-Дону: ЮФУ. – 2019. – № 7 (209). – С. 45-57.
3. Пикалёв, Я.С. Система автоматической генерации транскрипций русскоязычных слов-исключений на основе глубокого обучения / Я.С. Пикалёв, Т.В. Ермоленко // Проблемы искусственного интеллекта. – Донецк: ГУ ИПИИ. – 2019. – № 4(15). – С. 35-51.

4. Пикалёв, Я.С. Адаптация нейросетевой модели ALBERT для задачи языкового моделирования / Я.С. Пикалёв, Т.В. Ермоленко // Проблемы искусственного интеллекта. – Донецк: ГУ ИПИИ. – 2020. – № 3(18). – С. 111-122.

5. Пикалёв, Я.С. Обзор архитектур систем интеллектуальной обработки естественно-языковых текстов / Я.С. Пикалёв // Проблемы искусственного интеллекта. – Донецк: ГУ ИПИИ. – 2020. – № 4(19). – С. 45-68.

На автореферат поступило 6 отзывов. Все отзывы положительные. В них отражены актуальность исследования, дана оценка основным результатам, указаны замечания, а также сделаны положительные заключения о соответствии работы требованиям, предъявляемым к диссертациям на соискание ученой степени кандидата технических наук.

В отзывах содержатся следующие замечания:

1. **Кипяткова Ирина Сергеевна**, кандидат технических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей, старший научный сотрудник ФГБУН «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН), г. Санкт-Петербург.

1.1. В автореферате автором не раскрыты технологии обучения языковой модели, которая составляет важную часть системы распознавания слитной речи, ограничившись лишь упоминанием о триграммах и базовой архитектуре LSTM, хотя современный подход к построению языковых моделей использует ВРЕ-словарь и сложные модификации архитектуры Transformer.

2. **Линд Юлия Борисовна**, кандидат физико-математических наук по специальности 05.13.18 – Математическое моделирование, численные методы и комплексы программ, ученый секретарь ООО «РН-БашНИПИнефть», г. Уфа.

2.1. В автореферате недостаточно полно описана нейросетевая модель синтаксического анализатора, в частности, остается неясным, что подается ей на вход и что получается на выходе.

2.2. Автором разработана архитектура сети для классификации ошибок в тексте, что необходимо для процесса нормализации текста и получения качественной языковой модели. Однако автор не описал, на какие классы данная модель делит ошибки.

3. **Конопацкий Евгений Викторович**, доктор технических наук по специальности 05.01.01 – Инженерная геометрия и компьютерная графика, доцент, профессор кафедры «Специализированные информационные технологии и системы» ГОУ ВПО «ДОНБАССКАЯ НАЦИОНАЛЬНАЯ АКАДЕМИЯ СТРОИТЕЛЬСТВА И АРХИТЕКТУРЫ», г. Макеевка.

3.1. В разделе методология и методы исследования заявлены методы математической статистики для оценки эффективности разработанных моделей. При этом классические методы математической статистики, которые используются для оценки эффективности математических моделей, в автореферате отсутствуют, как и сами математические модели. Возможно, речь идет о показателе пословной ошибки распознавания применительно к акустическим моделям, но тогда в описании следовало бы это пояснить.

3.2. В автореферате отсутствует информация о том, какие конкретно результаты исследований были внедрены в Институте проблем искусственного интеллекта. Не указана также степень внедрения результатов научных исследований автора.

3.3. Эффективное использование нейросети во многом зависит от ее обучения. Отсюда возникает вопрос, какому из жанров русского языка автор обучал свою нейросеть: официально-деловому, научному, публицистическому, художественному или всем вместе?

4. **Левин Виталий Ильич**, доктор технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки), профессор, ведущий научный сотрудник ФГБОУ ВО «Пензенский государственный технологический университет» (ПензГТУ), г. Пенза.

4.1. Для исправления текстовых ошибок автором используется языковая модель ELMo, которая генерирует векторное представление для слова на основе его контекста. Модель улучшена при помощи техник teacher forcing и градиентного отсечения, за счет чего предложенная архитектура работает быстрее, чем стандартно используемые в NLP-задачах LSTM. Однако при разработке системы распознавания на декодер поступают вероятности, генерируемые стандартной языковой моделью на базе LSTM. Неясным остается выбор автора в пользу стандартной архитектуры LSTM при декодировании, а не улучшенной модели на базе ELMo.

5. **Кузнецов Олег Петрович**, доктор технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки), профессор, главный научный сотрудник ФГБУН «Институт проблем управления им. В.А. Трапезникова Российской Академии Наук», г. Москва.

5.1. В автореферате диссертации не уделено внимание статистическим языковым моделям.

5.2. В автореферате описана модель, используемая для синтаксического анализа, на вход которой подается вектор размерности 64, однако не указано, как формируется этот вектор.

6. **Трунаев Андрей Михайлович**, кандидат технических наук по специальности 05.13.06 – Автоматизация и управление технологическими процессами и производствами (по отраслям) (технические науки), старший преподаватель кафедры «АТС и ВТ» ГОУ ВПО «ДОНЕЦКИЙ ИНСТИТУТ ЖЕЛЕЗНОДОРОЖНОГО ТРАНСПОРТА», г. Донецк.

6.1. Для адаптации акустической модели под диктора используются  $i$ -вектора, которые содержат суммарную информацию о голосе диктора, помехах и искажении сигнала в канале связи, однако, нет количественных данных о допустимом уровне помех в речевом сигнале. При невысоком отношении сигнал-шум качество распознавания сильно падает. Было бы целесообразным при обучении акустической модели отдельно учесть влияние помех.

6.2. Для увеличения объема обучающих данных и повышения робастности акустической модели предложена техника аугментации речевых данных путем наложения шумов и голоса другого диктора. Речевой шум существенно снижает

точность распознавания. Следовало бы провести дополнительные исследования влияния уровня речевого шума на качество акустической модели.

**Диссертационный совет отмечает, что на основании выполненных соискателем исследований по специальности 05.13.01:**

– установлено, что предложенный процедурно-декларативный подход к генерации транскрипций слов русского языка, использующий словари транскрипций и глубокие нейросети, а также ряд дополнительных правил, учитывающих особенности фонетики русского языка, позволяет получать транскрипцию как для простых слов, так и для слов-исключений и иноязычных и заимствованных слов с учетом орфоэпических норм;

– показано, что компьютерная реализация алгоритмов получения акустической модели и распознавания фонем в единой системе автоматического распознавания слитной русской речи, основанных на методах глубокого обучения, позволяет распознавать слитную речь с точностью, не уступающей современным аналогам, в несколько превосходит их по скорости распознавания и требует для своего обучения значительно меньше ресурсов.

**Теоретическая значимость** результатов диссертационной работы заключается в экспериментальном обосновании использования модифицированных и новых нейросетевых моделей в задачах транскрибирования, параметризации речевого сигнала и распознавания фонем:

1) усовершенствован нейросетевой метод автоматического определения позиции ударения в слове путем увеличения количества слоёв, использования методов градиентного отсечения и *teacher forcing* для оптимизации параметра скорости обучения;

2) усовершенствована нейросетевая модель генерации практических транскрипций англоязычных слов и слов-исключений на базе архитектуры Transformer за счет применения механизма обучения с подкреплением метода *beam-search* для выбора наиболее вероятной последовательности символов;

3) разработан новый метод извлечения высокоуровневых информативных акустических признаков речевого сигнала, комбинирующий архитектуры глубоких сетей с «узким горлом» и сверточной сети на базе ResNet-50, что позволило обучать акустическую модель на признаках, устойчивых к шумам канала записи, темпу речи и смене диктора;

4) получили дальнейшее развитие методы нейросетевой классификации фонем за счет объединения архитектур глубоких сетей – сети с временными задержками и двунаправленной сети с долгой кратковременной памятью, а также использования механизма внимания в последнем скрытом слое, что позволило избежать явной сегментации и учитывать контекст из двух направлений.

**Практическая значимость** результатов исследований состоит в возможности использования разработанных алгоритмов и нейросетевых моделей в таких областях как: автоматическая обработка естественно-языковых текстов, синтез речи, разработка систем распознавания речи.

В работе получены компьютерные средства решения задач нормализации текстов на русском языке и создания аннотированных речевых баз, генерации

транскрипций слов, а также построения акустических моделей и классификации фонем, позволяющие использовать их создания широкого спектра систем на базе речевых технологий.

Практическая ценность исследований подтверждается внедрением в интеллектуальную систему, разработанную в в ГУ «ИПИИ» при выполнении фундаментальной научно-исследовательской работы «Исследование и разработка методов семантического анализа и интерпретации потоков данных интеллектуальными системами», государственный регистрационный номер работы 0118D000003 (справка №347/01-01 от 01.12.2020).

**Оценка достоверности результатов исследования** обеспечивается корректным использованием технологий глубокого обучения для разработанных средств и проверкой точности предложенных моделей и алгоритмов на речевых корпусах, которые были записаны в естественных условиях и являются представительными с точки зрения таких параметров, как характеристики канала, наличие фоновых шумов различной природы, длительность речевого сигнала, количество дикторов и их пол.

**Личный вклад соискателя** состоит в формулировании цели, задач исследований, основных научных положений и выводов, разработке вычислительных алгоритмов, нейросетевых моделей и комплекса программ для их компьютерной реализации. Основные научные результаты диссертации, которые включают вычислительные алгоритмы генерации транскрипций русских слов, получения акустических моделей и распознавания последовательностей фонем, а также комплекс программ, сформированный на основе этих алгоритмов, устойчиво функционирующий по отношению к темпу речи, акустической среде и междикторской вариативности, позволяющий с высокой точностью выполнять распознавание русской речи, получены автором лично.

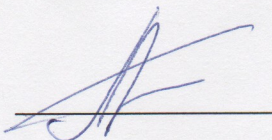
На основании вышеизложенного представленная диссертационная работа Пикалёва Ярослава Сергеевича «Совершенствование методов и программных средств распознавания слитной русской речи» является завершённой научно-исследовательской работой, в которой дано решение важной научно-технической задачи повышения эффективности дикторонезависимой системы автоматического распознавания слитной русской речи, учитывающей её особенности и адаптирующейся под любую предметную область, работающей с быстродействием и точностью, достаточными для практических задач, и требующей для своего обучения объём данных более, чем в 500 раз меньший, чем существующие аналоги. Работа отвечает требованиям п. 2.2 «Положения о присуждении ученых степеней», предъявляемым к диссертациям на соискание ученой степени кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки).

На заседании от «12» октября 2021 г. диссертационный совет принял решение: присудить Пикалёву Я.С. ученую степень кандидата технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (по отраслям) (технические науки).

При проведении тайного голосования диссертационный совет в количестве

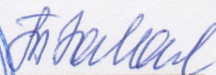
15 человек, из них 3 докторов наук, участвовавших в заседании, из 20 человек, входящих в состав совета, проголосовали: за 13, против 2, недействительных бюллетеней нет.

Председатель диссертационного совета  
Д 01.024.04 при ГОУВПО «ДОННТУ»  
и ГОУ ВПО «ДОННУ»,  
д-р техн. наук, профессор



В.Н. Павлыш

Ученый секретарь диссертационного  
совета Д 01.024.04  
канд. техн. наук, доцент



Т.В. Завадская