

ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ

әл-Фараби атындағы Қазақ ұлттық университетінің

Х А Б А Р Л А Р Ы

ИЗВЕСТИЯ

НАЦИОНАЛЬНОЙ АКАДЕМИИ
НАУК РЕСПУБЛИКИ КАЗАХСТАН
Қазақстан Республикасының
Ғылым Академиясының
Әл-Фараби атындағы
Қазақ ұлттық университетінің

NEWS

OF THE ACADEMY OF SCIENCES
OF THE REPUBLIC OF
KAZAKHSTAN
al-Farabi Kazakh National University

PHYSICO-MATHEMATICAL SERIES

1 (345)

JANUARY – MARCH 2023

PUBLISHED SINCE JANUARY 1963

PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

БАС РЕДАКТОР:

МУТАНОВ Ғалымқайыр Мұтанұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институтының бас директорының м.а. (Алматы, Қазақстан), **Н=5**

РЕДАКЦИЯ АЛҚАСЫ:

КАЛИМОЛДАЕВ Мақсат Нұрәліұлы (бас редактордың орынбасары), физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институты бас директорының кеңесшісі, зертхана меңгерушісі (Алматы, Қазақстан), **Н=7**

МАМЫРБАЕВ Өркен Жұмажанұлы (ғалым хатшы), Ақпараттық жүйелер саласындағы техника ғылымдарының (PhD) докторы, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институты директорының ғылым жөніндегі орынбасары (Алматы, Қазақстан), **Н=5**

БАЙГУНЧЕКОВ Жұмаділ Жанабайұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Кибернетика және ақпараттық технологиялар институты, қолданбалы механика және инженерлік графика кафедрасы, Сәтбаев университеті (Алматы, Қазақстан), **Н=3**

ВОЙЧИК Вальдемар, техника ғылымдарының докторы (физ-мат), Люблин технологиялық университетінің профессоры (Люблин, Польша), **Н=23**

СМОЛАРЖ Анджей, Люблин политехникалық университетінің электроника факультетінің доценті (Люблин, Польша), **Н=17**

ӘМІРҒАЛИЕВ Еділхан Несіпханұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Жасанды интеллект және робототехника зертханасының меңгерушісі (Алматы, Қазақстан), **Н=12**

КИЛАН Әлімхан, техника ғылымдарының докторы, профессор (ғылым докторы (Жапония), ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институтының бас ғылыми қызметкері (Алматы, Қазақстан), **Н=6**

ХАЙРОВА Нина, техника ғылымдарының докторы, профессор, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институтының бас ғылыми қызметкері (Алматы, Қазақстан), **Н=4**

ОТМАН Мохаммед, PhD, Информатика, коммуникациялық технологиялар және желілер кафедрасының профессоры, Путра университеті (Селангор, Малайзия), **Н=23**

НЫСАНБАЕВА Сауле Еркебұланқызы, техника ғылымдарының докторы, доцент, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институтының аға ғылыми қызметкері (Алматы, Қазақстан), **Н=3**

БИЯШЕВ Рустам Гакашевич, техника ғылымдарының докторы, профессор, Информатика және басқару мәселелері институты директорының орынбасары, Ақпараттық қауіпсіздік зертханасының меңгерушісі (Қазақстан), **Н=3**

КАПАЛОВА Нұрсұлтан Алдажарқызы, техника ғылымдарының кандидаты, ҚР БҒМ ҚҰО ақпараттық және есептеу технологиялар институтының киберқауіпсіздік зертханасының меңгерушісі (Алматы, Қазақстан), **Н=3**

КОВАЛЕВ Александр Михайлович, физика-математика ғылымдарының докторы, Украина Ұлттық Ғылым академиясының академигі, Қолданбалы математика және механика институты (Донецк, Украина), **Н=5**

МИХАЛЕВИЧ Александр Александрович, техника ғылымдарының докторы, профессор, Беларусь Ұлттық Ғылым академиясының академигі (Минск, Беларусь), **Н=2**

ТИГИНЯНУ Ион Михайлович, физика-математика ғылымдарының докторы, академик, Молдова Ғылым академиясының президенті, Молдова техникалық университеті (Кишинев, Молдова), **Н=42**

«ҚР ҰҒА Хабарлары. Информатика сериясы».

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Меншіктеуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы қ.). Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігінің Ақпарат комитетінде 14.02.2018 ж. берілген **№ 16906-Ж** мерзімдік басылым тіркеуіне қойылу туралы куәлік.

Тақырыптық бағыты: *физика-математикалық сериясы*.

Қазіргі уақытта: *«ақпараттық технологиялар» бағыты бойынша ҚР БҒМ БҒСБҚ ұсынған журналдар тізіміне енді.*

Мерзімділігі: *жылына 4 рет.*

Тиражы: *300 дана.*

Редакцияның мекен-жайы: *050010, Алматы қ., Шевченко көш., 28, 218 бөл., тел.: 272-64-39*

<http://www.physico-mathematical.kz/index.php/en/>

© Қазақстан Республикасының Ұлттық ғылым академиясы, 2023
Типографияның мекен-жайы: «Аруна» ЖК, Алматы қ., Мұратбаев көш., 75.

Главный редактор:

МУТАНОВ Галимкаир Мутанович, доктор технических наук, профессор, академик НАН РК, и.о. генерального директора «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), **Н=5**

Редакционная коллегия:

КАЛИМОЛДАЕВ Максат Нурадилович, (заместитель главного редактора), доктор физико-математических наук, профессор, академик НАН РК, советник генерального директора «Института информационных и вычислительных технологий» КН МНВО РК, заведующий лабораторией (Алматы, Казахстан), **Н=7**

МАМЫРБАЕВ Оркен Жумажанович, (ученый секретарь), доктор философии (PhD) по специальности «Информационные системы», заместитель директора по науке РГП «Институт информационных и вычислительных технологий» Комитета науки МНВО РК (Алматы, Казахстан), **Н=5**

БАЙГУНЧЕКОВ Жумадил Жанабаевич, доктор технических наук, профессор, академик НАН РК, Институт кибернетики и информационных технологий, кафедра прикладной механики и инженерной графики, Университет Саптаева (Алматы, Казахстан), **Н=3**

ВОЙЧИК Вальдемар, доктор технических наук (физ.-мат.), профессор Люблинского технологического университета (Люблин, Польша), **Н=23**

СМОЛАРЖ Анджей, доцент факультета электроники Люблинского политехнического университета (Люблин, Польша), **Н=17**

АМИРГАЛИЕВ Едилхан Несипханович, доктор технических наук, профессор, академик Национальной инженерной академии РК, заведующий лабораторией «Искусственного интеллекта и робототехники» (Алматы, Казахстан), **Н=12**

КЕЙЛАН Алимхан, доктор технических наук, профессор (Doctor of science (Japan)), главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), **Н=6**

ХАЙРОВА Нина, доктор технических наук, профессор, главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), **Н=4**

ОТМАН Мохамед, доктор философии, профессор компьютерных наук, Департамент коммуникационных технологий и сетей, Университет Путра Малайзия (Селангор, Малайзия), **Н=23**

НЫСАНБАЕВА Сауле Еркебулановна, доктор технических наук, доцент, старший научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), **Н=3**

БИЯШЕВ Рустам Гакашевич, доктор технических наук, профессор, заместитель директора Института проблем информатики и управления, заведующий лабораторией информационной безопасности (Казахстан), **Н=3**

КАПАЛОВА Нурсулу Алдажаровна, кандидат технических наук, заведующий лабораторией кибербезопасности РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), **Н=3**

КОВАЛЕВ Александр Михайлович, доктор физико-математических наук, академик НАН Украины, Институт прикладной математики и механики (Донецк, Украина), **Н=5**

МИХАЛЕВИЧ Александр Александрович, доктор технических наук, профессор, академик НАН Беларуси (Минск, Беларусь), **Н=2**

ТИГИНЯНУ Ион Михайлович, доктор физико-математических наук, академик, президент Академии наук Молдовы, Технический университет Молдовы (Кишинев, Молдова), **Н=42**

«Известия НАН РК. Серия информатики».

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Собственник: *Республиканское общественное объединение «Национальная академия наук Республики Казахстан» (г. Алматы).*

Свидетельство о постановке на учет периодического печатного издания в Комитете информации Министерства информации и общественного развития Республики Казахстан **№ 16906-Ж** выданное 14.02.2018 г.

Тематическая направленность: *серия физика-математическая.*

В настоящее время: *вошел в список журналов, рекомендованных ККСОН МОН РК по направлению «информационные коммуникационные технологии».*

Периодичность: *4 раз в год.*

Тираж: *300 экземпляров.*

Адрес редакции: *050010, г. Алматы, ул. Шевченко, 28, оф. 218, тел.: 272-64-39*

<http://www.physico-mathematical.kz/index.php/en/>

© Национальная академия наук Республики Казахстан, 2023
Адрес типографии: ИП «Аруна», г. Алматы, ул. Муратбаева, 75.

Chief Editor:

MUTANOV Galimkair Mutanovich, doctor of technical sciences, professor, academician of NAS RK, acting General Director of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), **H=5**

EDITORIAL BOARD:

KALIMOLDAYEV Maksat Nuradilovich, (Deputy Editor-in-Chief), Doctor of Physical and Mathematical Sciences, Professor, Academician of NAS RK, Advisor to the General Director of the Institute of Information and Computing Technologies of the CS MES RK, Head of the Laboratory (Almaty, Kazakhstan), **H = 7**

Mamyrbayev Orken Zhumazhanovich, (Academic Secretary, PhD in Information Systems, Deputy Director for Science of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), **H = 5**

BAIGUNCHEKOV Zhumadil Zhanabaevich, Doctor of Technical Sciences, Professor, Academician of NAS RK, Institute of Cybernetics and Information Technologies, Department of Applied Mechanics and Engineering Graphics, Satbayev University (Almaty, Kazakhstan), **H=3**

WOICIK Waldemar, Doctor of Technical Sciences (Phys.-Math.), Professor of the Lublin University of Technology (Lublin, Poland), **H=23**

SMOLARJ Andrej, Associate Professor Faculty of Electronics, Lublin polytechnic university (Lublin, Poland), **H= 17**

AMIRGALIEV Edilkhan Nesipkhanovich, Doctor of Technical Sciences, Professor, Academician of NAS RK, Head of the Laboratory of Artificial Intelligence and Robotics (Almaty, Kazakhstan), **H= 12**

KEILAN Alimkhan, Doctor of Technical Sciences, Professor (Doctor of science (Japan)), chief researcher of Institute of Information and Computational Technologies CS MES RK (Almaty, Kazakhstan), **H= 6**

KHAIROVA Nina, Doctor of Technical Sciences, Professor, Chief Researcher of the Institute of Information and Computational Technologies CS MES RK (Almaty, Kazakhstan), **H= 4**

OTMAN Mohamed, PhD, Professor of Computer Science Department of Communication Technology and Networks, Putra University Malaysia (Selangor, Malaysia), **H= 23**

NYSANBAYEVA Saule Yerkebulanovna, Doctor of Technical Sciences, Associate Professor, Senior Researcher of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), **H= 3**

BIYASHEV Rustam Gakashevich, doctor of technical sciences, professor, Deputy Director of the Institute for Informatics and Management Problems, Head of the Information Security Laboratory (Kazakhstan), **H= 3**

KAPALOVA Nursulu Aldazarovna, Candidate of Technical Sciences, Head of the Laboratory cybersecurity, Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), **H=3**

KOVALYOV Alexander Mikhailovich, Doctor of Physical and Mathematical Sciences, Academician of the National Academy of Sciences of Ukraine, Institute of Applied Mathematics and Mechanics (Donetsk, Ukraine), **H=5**

MIKHALEVICH Alexander Alexandrovich, Doctor of Technical Sciences, Professor, Academician of the National Academy of Sciences of Belarus (Minsk, Belarus), **H=2**

TIGHINEANU Ion Mihailovich, Doctor of Physical and Mathematical Sciences, Academician, President of the Academy of Sciences of Moldova, Technical University of Moldova (Chisinau, Moldova), **H=42**

News of the National Academy of Sciences of the Republic of Kazakhstan.

Series of informatics.

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Owner: RPA «National Academy of Sciences of the Republic of Kazakhstan» (Almaty). The certificate of registration of a periodical printed publication in the Committee of information of the Ministry of Information and Social Development of the Republic of Kazakhstan No. 16906-Ж, issued 14.02.2018

Thematic scope: *series physical-mathematical series.*

Currently: *included in the list of journals recommended by the CCSES MES RK in the direction of «information and communication technologies».*

Periodicity: *4 times a year.*

Circulation: *300 copies.*

Editorial address: *28, Shevchenko str., of. 218, Almaty, 050010, tel. 272-64-39*

<http://www.physico-mathematical.kz/index.php/en/>

© National Academy of Sciences of the Republic of Kazakhstan, 2023

Address of printing house: ST «Aruna», 75, Muratbayev str, Almaty.

NEWS OF THE NATIONAL ACADEMY OF SCIENCES
OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES

ISSN 1991-346X

Volume 1, Number 345 (2023), 37-49

<https://doi.org/10.32014/2023.2518-1726.167>

UDK 004.67, 311.218

© A. Bekarystankzy^{1*}, O.Zh. Mamyrbayev²

¹Kazakh National Research Technical University named after K.I. Satpayev,
Kazakhstan, Almaty;

²Institute of Information and Computer Technologies, Kazakhstan, Almaty.
E-mail: akbayan.b@gmail.com

INTEGRATED AUTOMATIC SPEECH RECOGNITION SYSTEM FOR AGGLUTINATIVE LANGUAGES

Bekarystankzy Akbayan — Doctoral student of the Institute of Automation and Information Technologies, specialty «Management information systems», Satbayev University. 050013. Almaty, Satbayev str. 22, Kazakhstan.

E-mail: akbayan.b@gmail.com, <https://orcid.org/0000-0003-3984-2718>;

Mamyrbayev Orken Zhumazhanovich — Deputy General Director in science, PhD, Institute of Information and Computational Technologies. 050010. Almaty, Shevchenko str. 28, Kazakhstan.

E-mail: morkenj@mail.ru. <https://orcid.org/0000-0001-8318-3794>.

Abstract. Nowadays Automatic Speech Recognition (ASR) systems are included in all possible applications which capture different parts of human life: smart assistants like Alica and Siri, smart search engines, Speech To Text (STT) applications for meetings, smart home applications and etc. However, these systems are available only for common languages, like English, Russian and Chinese. For low-resource languages such smart systems are not available, due to the lack of transcribed data for system training and necessity of studying modern effective methods for getting models and building ASR systems. For low-resource languages, like Kazakh language it is necessary to study effective methods for modelling and building effective ASR systems with low error rates. In this paper there were studied the methods of end-to-end Automatic Speech Recognition Systems. In this work it was proved that CTC model works for agglutinative languages. Author conducted experiments with BLSTM neural network using encoder-decoder model based on attention-based model. Training results show Character Error Rate (CER) - 8,01% and Word Error Rate (WER) — 17,91%. This work proves possibility of getting good results without integration of language models to end-to-end ASR models.

Key words: agglutinative languages, integral approach, CTC, neural network, BLSTM, speech recognition.

Acknowledgement. This is funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR18574144).

© А. Бекарыстанқызы^{1*}, Ө.Ж. Мамырбаев², 2023

¹Қ.И. Сәтбаев атындағы ҚазҰТЗУ, Қазақстан, Алматы;

²Ақпараттық және есептеуіш технологиялар институты, Қазақстан, Алматы.

E-mail: akbayan.b@gmail.com

АГГЛЮТИНАТИВТІ ТІЛДЕРГЕ АРНАЛҒАН СӨЙЛЕУДІ АВТОМАТТЫ ТҮРДЕ ТАНУ ЖҮЙЕСІ

Бекарыстанқызы Ақбаян — Докторант, Автоматтандыру және ақпараттық технологиялар институты, мамандығы «Management information systems», Satbayev University, 050013, ул. Сағбаева 22, Алматы, Қазақстан.

E-mail: akbayan.b@gmail.com, <https://orcid.org/0000-0003-3984-2718>;

Мамырбаев Өркен Жұмажанұлы — Бас директордың ғылым жөніндегі орынбасары, PhD, Ақпараттық және есептеуіш технологиялар институты. 050010. Алматы, ул. Шевченко 28, Қазақстан.

E-mail: morkenj@mail.ru, <https://orcid.org/0000-0001-8318-3794>.

Аннотация. Қазіргі уақытта сөйлеуді автоматты түрде тану жүйелері адам өмірінің әртүрлі саласын қамтитын алуан түрлі жүйелерге кіріктірілген. Олар: Алиса және Siri секілді ақылды көмекшілер, ақылды іздеу жүйелері, кездесулер хаттамасын құратын бағдарламалық жабдықтар, ақылды үйлер және т.б. Бірақ аталған жүйелер тек ағылшын, орыс, қытай секілді әлем халықтарының көпшілігі қолданатын тілдерде ғана жұмыс жасай алады. Жүйені үйретуге қажетті мәліметтер санының аздығына байланысты, біртұтас сөйлеуді тану жүйесіне құруға арналған заманауи тиімді әдістерді зерттей түсудің қажет екендігіне байланысты аз ресурсты тілдер үшін мұндай жүйелер әзірше қол жетімсіз болып табылады. Бұл мақалада сөйлеуді автоматты түрде танудың интегралды жүйелері зерттелген. Бұл жұмыста CTC моделінің агглютинативті тілдер үшін тиімді жұмыс жасайтындығы дәлелденген. Автор BLSTM (Bi-Deirectional Long Short-Term Memory) нейрондық жүйесін қолдана отырып назар аудару(attention-based models)механизміне негізделген шифратор-дешифратор моделін қолданып эксперимент жүргізген. Эксперимент нәтижелері 8,01%-ге тең болатын символдарды таңдау қателігін — CER (Character Error Rate)және 17,91%-ге тең сөздерді таңдау қателігін — WER(Word Error Rate)көрсетті. Сөйлеуді танудың біртұтас моделіне тілдік модельді қоспай-ақ жақсы нәтижелер алуға болатындығы дәлелденді.

Түйін сөздер: агглютинативті тілдер, интегралдық әдіс, CTC, BLSTM нейрондық жүйесі, сөйлеуді тану

Алғыс. Бұл жұмыс Қазақстан Республикасының Ғылым және жоғарғы білім министрлігінің Ғылым комитетінің қаржылық көмегі аясында орындалған. (BR18574144).

© А. Бекарыстанқызы^{1*}, О.Ж. Мамырбаев², 2023

¹КазНИТУ имени К.И. Сатпаева, Казахстан, Алматы;

²Институт информационных и вычислительных технологий, Казахстан, Алматы.

E-mail: akbayan.b@gmail.com

ИНТЕГРАЛЬНАЯ СИСТЕМА АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ СЛИТНОЙ РЕЧИ ДЛЯ АГГЛЮТИНАТИВНЫХ ЯЗЫКОВ

Бекарыстанқызы Ақбаян — Докторант, Институт автоматизации и информационных технологий, специальность «Management information systems». Satbayev University, 050013, ул. Сатбаева 22, Алматы, Казахстан.

E-mail: akbayan.b@gmail.com; <https://orcid.org/0000-0003-3984-2718>;

Мамырбаев Өркен Жұмажанұлы — Заместитель генерального директора по науке. PhD. Институт информационных и вычислительных технологий. 050010. Алматы, ул. Шевченко 28, Казахстан.

E-mail: morkenj@mail.ru; <https://orcid.org/0000-0001-8318-3794>.

Аннотация. В настоящее время системы автоматического распознавания речи внедрены во все возможные системы, которые охватывают разные сферы человеческой деятельности: умные помощники, такие как Алиса и Siri, умные поисковые системы, приложения для протоколирования встреч, умные дома и т.д. Но данные системы реализованы для более общих языков, такие как английский, русский и китайский. Для малоресурсных языков очень большая часть таких приложений не доступна в силу недостаточности данных для обучения и необходимости изучения современных методов получения сквозной модели системы автоматического распознавания речи. Малоресурсные языки, такие как казахский, нуждаются в исследовании эффективных методов для получения и построения систем распознавания речи. В статье исследованы современные методы для построения интегральных систем распознавания речи. В работе доказано, что модель CTC эффективно работает для агглютинативных языков. Автором проведен эксперимент с использованием нейронной сети BLSTM (Bi-Directional Long Short-Term Memory) при помощи модели шифратор-дешифратор, которая основана на механизме внимания (attention-based models). Результат эксперимента показал ошибку перебора символов CER (Character Error Rate), равный 8,01% и ошибку перебора слов WER (Word Error Rate), равный 17,91%. Доказано, что сквозная модель и без интегрирования языковых моделей можно достичь хороших результатов.

Ключевые слова: агглютинативные языки, интегральный подход, CTC, нейронная сеть BLSTM, распознавание речи.

Благодарность. Работа выполнена при финансовой поддержке Комитета науки Министерства науки и высшего образования Республики Казахстан (BR18574144).

Введение

Под речью подразумевается целая система звуковых сигналов, символов и письменных знаков, которые применяются человеком для того, чтобы хранить, представлять, перерабатывать и передавать данные. Кроме того, речь является инструментом, который необходим для того, чтобы машина и человек взаимодействовали между собой (Perera и др., 2005). Но для того, чтобы человек мог взаимодействовать с машиной, нужно разработать систему для автоматического распознавания слитной речи. Ранее для того, чтобы голосовой интерфейс был реализован, необходимо было прибегнуть к помощи множества специалистов: программиста по DNN, компьютерного лингвиста, чтобы он на каждый звуковой фрейм ставил метку, и так далее.

Стандартная система распознавания речи делится на ряд модулей: языковые и акустические модели, а также на декодеры (Мамугбаев и др., 2019). Основной для конструкции модульности является большое количество независимых предположений, а стандартная акустическая модель проходит обучение по фреймам, которые находятся в зависимости от модели Маркова. В различных автоматизированных системах распознавания наиболее популярные модели речевых сигналов, а также перечня гауссовских распределений плотностей вероятности для того, чтобы распределить сигналы на протяжении стационарного короткого временного промежутка, как правило, отвечает единице произношения.

Но в последнее время, с появлением сквозных систем распознавания речи, необходимость дополнительных меток для фреймов, также подборка фонем для каждой буквы или сочетания букв отпала. Стало достаточно иметь аудиоданные и их текстовые разметки. При несоответствии длины аудиофайлов с текстовыми данными сквозные системы способны сами провести выравнивание аудиофреймов с буквами. Одним из очень эффективных методов при выравнивании является CTC-функции. Также современные архитектуры используют нейронные сети из LSTM и Bi-LSTM ячеек, которые способны решить проблему зависимостей выражений на больших расстояниях.

Несмотря на эффективность современных методов, для малоресурсных языков вопрос улучшения качества распознавания остается открытым, так как интегральные сквозные системы требуют очень большое количество данных. Множество агглютинативных языков тоже относятся к малоресурсным.

В рамках данной работы была поставлена цель решить задачи

ограниченного речевого ресурса для распознавание агглютативных языков в интегральной архитектуре.

Краткий обзор исследований

В рамках исследований автоматического распознавания речи модель НММ-GMM была наиболее популярной множество лет. В настоящее время нейронная сеть используется в сфере распознавания речи. В большом количестве работ удалось доказать, что применение нейронных сетей в рамках каждого из этапов сценария стандартной системы распознавания делает намного лучше уровень качества данной системы. Самый популярный подход обучения — это гибридная архитектура НММ-DNN, в которой структура НММ остаётся неизменной, а GMM уходит и её место занимает наиболее глубокая нейронная сеть, которая необходима для того, чтобы были смоделированы характеристики речевых сигналов. К примеру, в многочисленных исследованиях Mikolov и др. (2010) конкретные языковые модели обучены благодаря RNN, в труде Rao K. и др. (2015) словарь удалось получить благодаря LSTM сетям. В свою очередь, в работе Jaitly и др. (2011) глубокие нейронные сети продемонстрировали достаточно высокие результаты для формирования акустических моделей, а в труде Smolensky (1986) продемонстрирован способ определения признаков благодаря ограниченным машинам Больцмана. Поэтому возникли идеи для применения искусственных нейронных сетей на различных периодах распознавания речи.

Методы глубокого обучения благодаря высокопроизводительным графическим процессорам используется на практике, и этот подход имеет название интегральный способ. При этом подходе, в процессе обучения нейронной сети исключительно одна из моделей может демонстрировать необходимый результат, не применяя при этом других элементов — интегральная модель. Те или иные интегрированные сети можно сформировать благодаря тому, что будет добавляться ряд рекуррентных и интегральных слоев, которые являются языковыми и акустическими моделями, и благодаря ним сопоставляются друг с другом речевые данные на выходе, с использованием транскрипции. В настоящее время имеется целый ряд способов для того, чтобы были реализованы интегральные модели — шифратор-дешифратор модели, а также коннекционная временная классификация. Они базируются на механизме внимания. Во множестве задач распознавания речи особенный акцент делается на интегральном подходе (Vaněk и др., 2017). В большом количестве работ удалось доказать, что успехи итогов интегрального подхода находятся в зависимости от роста объёма данных для обучения. Имеется множество приложений, которые работают, базируясь непосредственно на интегральном подходе: Baidu Deep Speech, Google Listen, Attend, Spell, Speechto Translator TTS, Voiceto Text Messenger. Главная причина данного подходе состоит в том, что на сегодняшний день интегральные модели проходят обучение, базируясь на данных. На основании анализа выше можно

заметить главную проблематику, которая связана с распознаением мало ресурсных языков, которые входят в группу агглютинативных. Для данных языков не имеется общих корпусов обучающих данных. Иные языки обладают TIMIT, WSJ, Libri Speech, AMI и Switchboard, благодаря которым существует множество языков обучающих данных.

Чтобы улучшить интегральный подход в различных моделях — шифратор-дешифратор, а также CTC, которые базируются на механизме внимания, проводились разнообразные варианты сетей. Чтобы применять локальные корреляции в речевых сигналах были использованы комплексные кодеры, в которые входят сверточные нейронные сети. Эти модели применяют те или иные преимущества для всех подмоделей и влияют на возникновение ограничений для модели. Анализ положительно влияет на степень производительности различных интегральных систем. В прошлых исследованиях удалось выявить, что модели глубокого обучения на разнообразных языках наиболее удачны, а многозадачное обучение наиболее оптимально для интегрального обучения (Kim и др., 2017; Aida-Zade и др., 2009).

Методика и материалы

Модели, которые базируются на коонекционной временной классификации и необходимы для распознавания речи, осуществляют свою деятельность, исключая первоначальное выравнивание выходных и входных последовательностей. CTC сформирован для того, чтобы декодировать язык. Hannun вместе со своей командой применяли для распознавания речи Baidu, который использует параллельный алгоритм обучения с применением CTC.

В труде (Zhang и др., 2016) предлагается применять более глубокие рекуррентные сверточные, а также глубокие остаточные сети вместе с CTC. Наиболее оптимальный результат удалось получить с использованием остаточных сетей с батч-нормализацией. Благодаря этому, удалось получить PER, который на речевом корпусе TIMIT был равен 17,3%.

Альтернативный вариант CTC — это Sequence to Sequence с вниманием (Bahdanau и др., 2016). Данные модели включают в себя как декодировщик, так и кодировщик. Благодаря кодировщику происходит сжатие данных кадров аудио в наиболее компактное представление благодаря сокращению число нейронов в слоях. В свою очередь, декодировщик, базируясь на сжатом представлении, а также рекуррентной нейронной сети, работает над восстановлением этапов слов, символов и фоном.

В работе Zhang и др., (2017) предлагается CTC-модель с применением наиболее глубоких сверточных сетей взамен рекуррентных сетей. Наиболее оптимальная модель, которая базируется на сверточных сетях, обладала десятью сверточными слоями и тремя полно-связными слоями. Наиболее лучшая PER равна 18,2%, учитывая то, что наилучшая PER для двунаправленных LSTM равняется 18,3%. Тестирование осуществлялось на

основании ТИМТ. Удалось определить, что сверточные сети дают возможность скорость обучения сделать быстрее, и наиболее оптимальными они являются для обучения на последовательностях фонем.

В рамках CTC сети, выходные значения нейронной сети — это вероятности перехода. Архитектура нейронной сети — двунаправленные LSTM сети. Между собой проводилось сравнение трёх моделей: RNN-CTC модель, RNN-CTC модель, а также минимизированной переобученной WER и базовой гибридной модели, которая была написана благодаря инструментам Kaldi (Povey и др., 2011).

Soltau и др. (2016) осуществляли контекстно-зависимое распознавание фонем, которые обучили модель, базируясь на CTC в задаче подписи видео на YouTube. В Sequence-to-sequence моделях недостаточное количество распознавания на примерно 13-35%, если сравнивать их с базовыми системами. Имеет место быть «обобщение» CTC моделей, а именно — RNN преобразовательный, на основании которого объединяются друг с другом две RNN в определенную преобразовательную систему (Anandan и др., 2002). Одна сеть является схожей с CTC-сетью, и занимается обработкой одинакового временного периода, который обрабатывается входной последовательностью. В свою очередь, благодаря второй RNN моделируется вероятность будущих метол, принимая во внимание прошлые. В CTC сетях применяются динамическое программирование для того, чтобы вычислять алгоритмы, а также алгоритмы прямого и обратного перехода, при этом принимая во внимание ограничения, присутствующие в двух RNN. Если сравнивать с CTC-сетями, благодаря применению RNN преобразователя можно формировать выходные последовательности, которые являются наиболее длинными, чем входные.

RNN преобразователи продемонстрировали довольно хорошие результаты в процессе распознавании фонем с PER, которая равна 17,7% и базируется на корпусе ТИМТ.

Предлагаемая система автоматического распознавания речи

В работе методология выполнялась следующим способом:

CTC-функция

CTC-функция для обучения нейронной сети применяется в виде функции потерь. Выходная последовательность нейронной сети описывается следующей формулой: $y = f_w(x)$. В выходной слой нейронной сети входит один блок для каждого из символов выходной последовательности, а также дополнительный символ “blank”. Все компоненты выходной последовательности — это вектора распределения вероятности для всех символов G' в конкретный временной промежуток t . Итак, компонентом y^t является вероятность того, что в конкретный временной промежуток t во входной последовательности будет произнесен символ k из множества символов G' .

Итак, пусть α является последовательностью из символов длины T и индексов blanks, согласно по x . Вероятность $P(\alpha|x)$ представляется следующим образом: произведение вероятностей появления символов в каждый временной период:

$$P(\alpha|x) = \prod_{t=1}^T y_{\alpha_t}^t, \forall \alpha \in G^t \quad (1)$$

Пусть B – является оператором, который удаляет повторы символов и blanks.

$$P(\alpha|x) = \sum_{\alpha \in B^{-1}(y)} P(\alpha|x) \quad (2)$$

Вышеописанная формула вычисляется с помощью динамического программирования, а нейронная сеть будет обучена минимизировать CTC-функцию:

$$CTC(x) = -\ln(P(y|x)) \quad (3)$$

Декодирование основывается на следующем предположении:

$$\arg \max P(y|x) \approx B(\alpha^*) \quad (4)$$

где $\alpha^* = \arg \max_{\alpha} P(\alpha|x)$.

Attention-based model

Attention является механизмом Encoder-Decoder, который был сформирован для того, чтобы улучшить уровень производительности RNN в процессе распознавания речи. Шифратором является нейронная сеть, к которым относятся: DNN, BLSTM, CNN. Он изменяет входную последовательность $x = (x_1, \dots, x_{LF})$, чтобы определить признаки в определенное промежуточное представление: $h = (h_1, \dots, h_L)$.

$$h = \text{Encoder}(x_1, \dots, x_{LF}) \quad (5)$$

Дешифратор (Decoder) – это обычный RNN, который использует промежуточное представление для генерации выходных последовательностей:

$$P(y|x) = \text{Attention Decoder}(h, y) \quad (6)$$

В виде дешифратора применялся рекуррентный генератор последовательности, который базируется на механизме внимания.

Данные

Информация для анализа была представлена благодаря лаборатории

«Компьютерной инженерии интеллектуальных систем». В связи с этим применялись: профессиональная звукозаписывающая студия Vocalbooth.com, которая является шумоизоляционной.

Среди дикторов были люди без проблем с произношением. Всего в записи участвовали 380 дикторов разного пола и возраста. Процесс озвучивания и записи одного диктора продолжался около 50 минут. Каждый диктор читал собственный текст, который включал в себя сто предложений, записанные в отдельные файлы. Каждое предложение включало в себя в среднем семь слов с достаточно богатой фонемой. Текстовая информация собиралась из новостных сайтов, которые пишут на казахском языке, но применялись и другие данные, которые хранятся в электронном виде. В общей сложности запись данных достигла 123 ч. В период записи сформированы транскрипции, под которыми подразумевается описание каждого из аудиофайлов, но в текстовом виде. Сформированный корпус позволяет, в первую очередь, осуществлять работу с довольно большими объемами данных, предлагать характеристики системы. Кроме того, можно исследовать то, как воздействуют расширения базы данных на скорость распознавания тексты.

Аудиоматериалы обладают идентичными характеристиками:

1. PCM – это метод образования файла в цифровой вид;
2. 16 битная разрядность файла;
3. Один аудиоканал;
4. Wav – это расширение файлов;
5. Частота 441,1 кГц.

В рамках обучения интегральной системы применялись два корпуса, к которым относятся:

- Корпус турецкого языка: <http://www.tnc.org.tr>;
- Корпус татарского языка: <https://commonvoice.mozilla.org/ru/datasets>.

Реализация

Система интегрального распознавания речи с применением CTC — функции реализовывалась с применением TensorFlow. В этой системе применялся инструмент Eesen в Tensor Flow. Данная система даёт возможность применять языковые модели, которые были сформированы в формате Kaldi, но дополнительная конвертация при этом не применялась. Tensor2 Tensor5 применялась для того, чтобы осуществлялись эксперименты с моделями типа Attention-based models.

Эксперименты осуществлялись с применением сервера Supermicro SYS-7049GP-TRT. Конфигурация данного сервера обладает высокопроизводительной видеокартой NVIDIA TESLA P100.

Результаты

В экспериментах для извлечения признаков мы использовали *мел-частотные кепстральные коэффициенты* (MFCC) с первыми 13 вычисленными коэффициентами. Все данные обучения были разделены на обучающие (90%) и перекрёстную проверку (cross-validation 10%).

На втором этапе эксперимента мы опишем результаты модели на основе функции потерь CTC. Результаты соответствующих CTC-моделей представлены в Таблице 1. В исследованиях мы использовали несколько типов нейронных сетей: ResNet, LSTM, MLP, Bidirectional LSTM. Предварительная настройка нейронных сетей без языковой модели дала нам наилучшие результаты:

MLP: было 6 скрытых слоев с 1024 узлами, при использовании функции активации ReLU с начальной скоростью обучения, равной 0,007, и коэффициентом затухания, равным 1,5.

LSTM: было 6 слоя с 1024 единицами в каждом с выпадением, равным 0,5 с, начальной скоростью обучения, равной 0,001, и коэффициентом затухания, равным 1,5.

ConvLSTM: использован один двумерный сверточный слой с 8 фильтрами, функция активации ReLU. Затем он выпадает с вероятностью удержания, равной 0,5.

BLSTM: использовал 6 слоев с 1024 единицами и выпадал с вероятностью удержания, равной 0,5.

ResNet было 9 остаточных блоков с нормализацией (batch-normalization).

В первом эксперименте для моделей шифратор-дешифратор, основанных на механизме внимания (attention-based models), для извлечения признаков мы использовали алгоритм MFCC.

В первом эксперименте для моделей шифратор-дешифратор, основанных на механизме внимания (attention-based models), для извлечения признаков мы использовали алгоритм MFCC, для обучения нейронной сети применяли функцию CTC. Мы не использовали языковые модели в данной модели. Во втором эксперименте мы использовали MFCC и языковые модели.

Таблица 1 – Результаты CTC-моделей.

Модель	CER%	WER%	Decode	Train
Модели, не использующие языковые модели.				
MLP	48.11	59.26	0.2032	131.2
LSTM	36.43	46.51	0.2152	421.3
Conv+LSTM	34.92	39.31	0.2688	465.2
BLSTM	33.61	37.66	0.2722	491.7
ResNet	32.52	36.57	0.2657	192.6
Модели, использующие языковые модели и MFCC.				
MLP	39.11	63.26	0.0192	146.2
LSTM	24.43	46.51	0.0152	521.3
Conv+LSTM	22.92	39.31	0.0088	465.2
BLSTM	13.61	20.66	0.0022	591.7
ResNet	11.52	19.57	0.0051	242.6

В следующем эксперименте мы применяли нейронные сети LSTM и BLSTM. В нашей модели использовалось 6 слоев по 256 единиц с начальным

уменьшением отсева при вероятности сохранения в кодере 0,7. В качестве декодера мы использовали LSTM и шифратор-дешифратор модели, основанные на механизме внимания (attention-based models). Результаты можно увидеть в таблице 2.

Обсуждение

Проведенные нами эксперименты доказали, что модель CTC работает без языковых моделей непосредственно для агглютинативных языков, но все равно наилучшим является ResNet с результатом CER, равным 11,52% и WER, равным 19,57%, с использованием языковой модели. Таким образом, можно увидеть, что языковая модель является важной частью распознавания речи.

Таблица 2 – Результаты шифратор-дешифратор модели, основанные на механизме внимания (Attention-based models).

Модель	CER%	WER%	Decode	Train
LSTM	8,61	17,58	0,468	476,7
BLSTM	8,01	17,91	0,496	544,3

Модель CTC допускает ошибки в построении слов и предложений из распознанных символов, но полученная фонематическая транскрипция очень похожа на оригинал. Но после эксперимента мы обнаружили, что использование шифратор-дешифратор модели, основанной на механизме внимания (attention-based models) для агглютинативных языков без интегрирования языковых моделей, позволяет достичь хороших результатов. Нейронная сеть BLSTM с помощью шифратор-дешифратор модели, основанной на механизме внимания (attention-based models), показала результат CER, равный 8,01% и WER, равный 17,91%.

Заключение

В этой работе мы рассматриваем задачу распознавания агглютинативных языков с помощью интегрального подхода, таких как модель CTC и шифратор-дешифратор модели, основанных на механизме внимания (attention-based models). При проведении эксперимента мы использовали различного вида архитектуры нейронных сетей: MLP, LSTM и их модификации, а также ResNet. В результате эксперимента мы доказали, что без интегрирования языковых моделей можно достичь хороших результатов. Наилучший результат показали ResNet. В данном эксперименте были достигнуты хорошие результаты, лучшие, чем базовые гибридные модели.

В будущем планируется проведение экспериментов с использованием других типов моделей для извлечения признаков и для распознавания речи. Будет применяться модель условно случайные поля (Conditional Random File).

ЛИТЕРАТУРА

Aida-Zade K., Rustamov S., Mustafayev E., 2009 — *Aida-Zade K., Rustamov S., Mustafayev E. Principles of Construction of Speech Recognition System by the Example of Azerbaijan Language. Int. Symposium on Innovations in Intelligent Systems and Applications. Pp. 378–382. (in Eng.).*

Anandan P., Saravanan K., Parthasarathy R., Geetha T., 2002 — *Anandan P., Saravanan K., Parthasarathy R., Geetha T.* Morphological Analyzer for Tamil. In the proceedings of ICON 2002. Chennai. (in Eng.).

Bahdanau D., 2016 — *Bahdanau D.* et al. End-to-end attention-based large vocabulary speech recognition. *Acoustics, Speech and Signal Processing (ICASSP)*, 2016 IEEE International Conference. Pp. 4945–4949. (in Eng.).

Hannun A., Case C., Casper J., Catanzaro B., Diamos G., Elsen E. et al., 2014 — *Hannun A., Case C., Casper J., Catanzaro B., Diamos G., Elsen E.* Deepspeech: scaling up end-to-end speech recognition. ArXiv: 1412.5567 (2014). (in Eng.).

Jaitly N., Hinton G., 2011 — *Jaitly N., Hinton G.* Learning a better representation of speech soundwaves using restricted boltzmann machines. 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Pp. 5884–5887. (in Eng.).

Kim S., Hori T. & Watanabe S., 2017 — *Kim S., Hori T. & Watanabe S.* Joint CTC-attention based end-to-end speech recognition using multi-task learning. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). (in Eng.).

Mamyrbayev O., Turdalyuly M., Mekebayev N., Alimhan K., Kydyrbekova A., Turdalykyzy T., 2019 — *Mamyrbayev O., Turdalyuly M., Mekebayev N., Alimhan K., Kydyrbekova A., Turdalykyzy T.* Automatic Recognition of Kazakh Speech Using Deep Neural Networks. *ACIIDS 2019, LNAI 11432*. Pp. 465–474, 2019. (in Eng.).

Mikolov T. et al. (2010) — *Mikolov T. et al.* Recurrent neural network based language model. *Interspeech*. vol. 2. Pp.1045–1048. (in Eng.).

Perera F.P., Tang D., Rauh V., Lester K., Tsai W.Y., Tu Y.H., 2005 — *Perera F.P., Tang D., Rauh V., Lester K., Tsai W.Y., Tu Y.H.* Relationship between polycyclic aromatic hydrocarbon–DNA adducts and proximity to the World Trade Center and effects on fetal growth. *Environ Health Perspect*. 113. Pp. 1062–1067. (in Eng.).

Povey D. et al., 2011 — *Povey D. et al.* The Kaldi speech recognition toolkit. IEEE 2011 workshop on automatic speech recognition and understanding. IEEE Signal Processing Society. Pp. 4. (in Eng.).

Rao K., Peng F., Sak H., Beaufays F., 2015 — *Rao K., Peng F., Sak H., Beaufays F.* Grapheme-to-phoneme conversion using long short-term memory recurrent neural networks. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2015. Pp. 4225–4229. (in Eng.).

Smolensky P., 1986 — *Smolensky P.* Information processing in dynamical systems: Foundations of harmony theory. Colorado University at Boulder Dept of Computer Science. Pp. 194–281. (in Eng.).

Soltau H., Liao H., Sak H., 2016 — *Soltau H., Liao H., Sak H.* Neural speech recognizer: acoustic-to-word LSTM model for large vocabulary speech recognition. ArXiv:1610.09975. (in Eng.).

Vaněk J., Zelinka J., Soutner D. & Pstuka J., 2017 — *Vaněk J., Zelinka J., Soutner D. & Pstuka J.* A regularization post layer: An additional way how to make deep neural networks robust. In International Conference on Statistical Language and Speech Processing (ICASSP). (in Eng.).

Zhang Z. et al., 2016 — *Zhang Z. et al.* Deep Recurrent Convolutional Neural Network: Improving Performance For Speech Recognition. arXiv: 1611.07174. URL:<https://arxiv.org/abs/1611.07174>. (in Eng.).

Zhang Y. et al., 2017 — *Zhang Y. et al.* Towards end-to-end speech recognition with deep convolutional neural networks. preprint: arXiv: 1701.02720. URL: <https://arxiv.org/abs/1701.02720>. (in Eng.).

REFERENCES

Aida-Zade K., Rustamov S., Mustafayev E., 2009 — *Aida-Zade K., Rustamov S., Mustafayev E.* Principles of Construction of Speech Recognition System by the Example of Azerbaijan Language. Int. Symposium on Innovations in Intelligent Systems and Applications. Pp. 378–382. (in Eng.).

Anandan P., Saravanan K., Parthasarathy R., Geetha T., 2002 — *Anandan P., Saravanan K., Parthasarathy R., Geetha T.* Morphological Analyzer for Tamil. In the proceedings of ICON 2002. Chennai. (in Eng.).

Bahdanau D., 2016 — *Bahdanau D.* et al. End-to-end attention-based large vocabulary speech recognition. *Acoustics, Speech and Signal Processing (ICASSP)*, 2016 IEEE International Conference. Pp. 4945–4949. (in Eng.).

Hannun A., Case C., Casper J., Catanzaro B., Diamos G., Elsen E. et al., 2014 — *Hannun A., Case C., Casper J., Catanzaro B., Diamos G., Elsen E. et al.* Deepspeech: scaling up end-to-end speech recognition. ArXiv: 1412.5567 (2014). (in Eng.).

Jaitly N., Hinton G., 2011 — *Jaitly N., Hinton G.* Learning a better representation of speech soundwaves using restricted boltzmann machines. 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). –Pp. 5884–5887. (in Eng.).

Kim S., Hori T. & Watanabe S., 2017 — *Kim S., Hori T. & Watanabe S.* Joint CTC-attention based end-to-end speech recognition using multi-task learning. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. (in Eng.).

Mamyrbayev O., Turdalyuly M., Mekebayev N., Alimhan K., Kydyrbekova A., Turdalykyzy T., 2019 — *Mamyrbayev O., Turdalyuly M., Mekebayev N., Alimhan K., Kydyrbekova A., Turdalykyzy T.* Automatic Recognition of Kazakh Speech Using Deep Neural Networks. *ACHIDS 2019, LNAI 11432*. –Pp. 465–474, 2019. (in Eng.).

Mikolov T. et al., 2010 — *Mikolov T. et al.* Recurrent neural network based language model. *Interspeech*. vol. 2. –Pp.1045–1048. (in Eng.).

Perera F.P., Tang D., Rauh V., Lester K., Tsai W.Y., Tu Y.H., 2005 — *Perera F.P., Tang D., Rauh V., Lester K., Tsai W.Y., Tu Y.H.* Relationship between polycyclic aromatic hydrocarbon–DNA adducts and proximity to the World Trade Center and effects on fetal growth. *Environ Health Perspect*.113. –Pp. 1062–1067. (in Eng.).

Povey D. et al., 2011 — *Povey D. et al.* The Kaldi speech recognition toolkit. *IEEE 2011 workshop on automatic speech recognition and understanding*. IEEE Signal Processing Society. –Pp. 4. (in Eng.).

Rao K., Peng F., Sak H., Beaufays F., 2015 — *Rao K., Peng F., Sak H., Beaufays F.* Grapheme-to-phoneme conversion using long short-term memory recurrent neural networks. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2015. –PP. 4225–4229. (in Eng.).

Smolensky P., 1986 — *Smolensky P.* Information processing in dynamical systems: Foundations of harmony theory. *Colorado University at Boulder Dept of Computer Science*. –Pp. 194–281. (in Eng.).

Soltau H., Liao H., Sak H., 2016 — *Soltau H., Liao H., Sak H.* Neural speech recognizer: acoustic-to-word LSTM model for large vocabulary speech recognition. ArXiv:1610.09975. (in Eng.).

Vaněk J., Zelinka J., Soutner D. & Psutka J., 2017 — *Vaněk J., Zelinka J., Soutner D. & Psutka J.* A regularization post layer: An additional way how to make deep neural networks robust. In *International Conference on Statistical Language and Speech Processing (ICASSP)*. (in Eng.).

Zhang Z. et al., 2016 — *Zhang Z. et al.* Deep Recurrent Convolutional Neural Network: Improving Performance For Speech Recognition. arXiv: 1611.07174. URL:<https://arxiv.org/abs/1611.07174>. (in Eng.).

Zhang Y. et al., 2017 — *Zhang Y. et al.* Towards end-to-end speech recognition with deep convolutional neural networks. preprint: arXiv: 1701.02720. URL: <https://arxiv.org/abs/1701.02720>. (in Eng.).

МАЗМҰНЫ

Ж.К. Абдугулова, Г.А. Ускенбаева, М.Н. Тлеген, А.К. Шукирова ҚҰБЫР ЖАБДЫҒЫНДА МАЙДЫ ҚЫЗДЫРУДЫҢ ТЕХНОЛОГИЯЛЫҚ ПРОЦЕСІН АВТОМАТТАНДЫРУ.....	5
Ж.С. Авкурова, С. Гнатюк, Л.М. Кыдыралина, Н.К. Курмангалиева АҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ЖҮЙЕЛЕРДЕ ҚҰҚЫҚ БҰЗУШЫНЫ ЕРТЕ АНЫҚТАУ ЖӘНЕ СӘЙКЕСТЕНДІРУДІҢ ИНТЕЛЛЕКТУАЛДЫ ӘДІСІ.....	22
А. Бекарыстанкызы, Ө. Ж. Мамырбаев АГГЛЮТИНАТИВТІ ТІЛДЕРГЕ АРНАЛҒАН СӨЙЛЕУДІ АВТОМАТТЫ ТҮРДЕ ТАҢУ ЖҮЙЕСІ.....	37
А.С. Еримбетова, Э.Н. Дайырбаева, Л. Черикбаева БИКУБТЫҚ ИНТЕРПОЛЯЦИЯҒА НЕГІЗІНДЕ СУРЕТТЕРГЕ ЖАСЫРЫН АҚПАРАТТЫ ЕНГІЗУ.....	50
М.Б. Есенова, Г.Б. Абдикеримова, А. Толстой, Ж.Б. Ламашева, А.А. Некесова БИДАЙДАҒЫ АРАМШӨПТЕР ОШАҒЫН АНЫҚТАУ ҮШІН ТЕКСТУРАЛЫҚ БЕЛГІЛЕР ӘДІСТЕРІН ҚОЛДАНУ.....	64
Л.З. Жолшиева, Т.К. Жукабаева, Ш. Тураев, М.А. Бердиева, Р.К. Сенгирбаева НАҚТЫ УАҚЫТ РЕЖИМІНДЕ МЕДИАРИРЕ ЖӘНЕ SVM АРҚЫЛЫ ҚАЗАҚ ҰМ ТІЛІН ТАҢУ.....	82
Ж.С. Иксебаева, К. Жетписов, А.Б. Медешова, И.М. Бапиев, Ж.Ж. Багисов ҒАЛЫМДАРДЫҢ ҒЫЛЫМИ ЖОБАЛАР БОЙЫНША ГРАНТТЫҚ ҚАРЖЫЛАНДЫРУҒА ҚАТЫСУҒА ӨТІНІМДЕРІН ДАЙЫНДАУДЫҢ АҚПАРАТТЫҚ ЖҮЙЕСІ.....	94
А.А. Иманберді, Р.Н. Молдашева ӘЛЕУМЕТТІК МЕДИА ТАРАТУ ҮЛГІЛЕРІНЕ ШОЛУ.....	107
Г. Қалман, М.Ғ. Есмағанбет, М.М. Жаманкарин, А.И. Габдулина, Д.В. Плескачев КЛАСТЕРЛЕУ ӘДІСІН ҚОЛДАНЫП КОРЕФЕРЕНЦИЯН ШЕШУ.....	121

Қ.Т. Қырғызбай, Е.Х. Какимжанов ГАЗ ТЕХНОЛОГИЯЛАРЫ НЕГІЗІНДЕ АЛМАТЫ ОБЛЫСЫНЫҢ ГЕОДЕРЕКТЕР БАЗАСЫН ҚҰРУ ВІТСОІН ЖЕЛІСІНДЕГІ КҮДІКТІ ТРАНЗАКЦИЯЛАРДЫ АНЫҚТАУ.....	136
Ш.Ж. Мусиралиева, М.Ж. Шайзат, А.К. Бекетова, Е. Абайұлы, А.Б. Манасова ВІТСОІН ЖЕЛІСІНДЕГІ КҮДІКТІ ТРАНЗАКЦИЯЛАРДЫ АНЫҚТАУ.....	154
А.Ұ. Мұхиядин, Ұ.Т. Махажанова, М.У. Мукашева, А.А. Муханова АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР ТӨТЕНШЕ ЖАҒДАЙДА ҚАШЫҚТАН ОҚЫТУДА ЭКСПЕРИМЕНТТЫҚ ДЕРЕКТЕРДІ ТАЛДАУ ҚҰРАЛЫ РЕТІНДЕ.....	170
А.Б. Тоқтарова, Б.С. Омаров, Г.Н. Казбекова, С.А. Мамиков, Ф.Е. Темірбекова ӘЛЕУМЕТТІК ЖЕЛІДЕГІ ҚАЗАҚ ТІЛДІ БЕЙӘДЕП СӨЗДЕР ҚОРЫН МАШИНАЛЫҚ ОҚЫТУДА ЖИНАҚТАУ.....	191
А.Ә. Шекербек, Г.Б. Абдикеримова, Ж.Б. Ламашева, М.Г. Байбулова, А.К. Токкулиева ТЕРЕҢ ОҚЫТУ АЛГОРИТМІМЕН РЕНТГЕНДІК КЕСКІННІҢ КЛАССИФИКАЦИЯСЫ.....	204
Э.Э. Эльдарова JPEG2000 ҚЫСУЫНАН КЕЙІН ЦИФРЛІК БЕЙНЕЛЕРДІҢ ВИЗУАЛДЫ САПАСЫН ЖАҚСARTУ.....	228

СОДЕРЖАНИЕ

Ж.К. Абдугулова, Г.А. Ускенбаева, М.Н. Глеген, А.К. Шукирова АВТОМАТИЗАЦИЯ ТЕХНОЛОГИЧЕСКОГО ПРОЦЕССА ПОДОГРЕВА НЕФТИ НА ТРУБОПРОВОДНОМ ОБОРУДОВАНИИ.....	5
Ж.С. Авкурова, С.А. Гнатюк, Л.М. Кыдыралина, Н.К. Курмангалиева ИНТЕЛЛЕКТУАЛИЗИРОВАННЫЙ МЕТОД РАННЕГО ВЫЯВЛЕНИЯ ИДЕНТИФИКАЦИИ НАРУШИТЕЛЯ В ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫХ СИСТЕМАХ.....	22
А. Бекарыстанқызы, О. Ж. Мамырбаев ИНТЕГРАЛЬНАЯ СИСТЕМА АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ СЛИТНОЙ РЕЧИ ДЛЯ АГГЛЮТИНАТИВНЫХ ЯЗЫКОВ.....	37
А.С. Еримбетова, Э.Н. Дайырбаева, Л. Черикбаева ВНЕДРЕНИЕ СКРЫТОЙ ИНФОРМАЦИИ В ИЗОБРАЖЕНИИ НА ОСНОВЕ БИКУБИЧЕСКОЙ ИНТЕРПОЛЯЦИИ.....	50
М.Б. Есенова, Г.Б. Абдикеримова, А. Толстой, Ж.Б. Ламашева, А.А. Некесова ПРИМЕНИМОСТЬ МЕТОДОВ АНАЛИЗА ТЕКСТУРНЫХ ИЗОБРАЖЕНИЙ ДЛЯ ВЫЯВЛЕНИЯ ОЧАГОВ СОРНЫХ ТРАВ ПШЕНИЦЫ.....	64
Л.З. Жолшиева, Т.К. Жукабаева, Ш. Тураев, М.А. Бердиева, Р.К. Сенгирбаева РАСПОЗНАВАНИЕ КАЗАХСКОГО ЖЕСТОВОГО ЯЗЫКА В РЕАЛЬНОМ ВРЕМЕНИ С ИСПОЛЬЗОВАНИЕМ MEDIAPIPE и SVM.....	82
Ж.С. Иксебаева, К. Жетписов, А.Б. Медешова, И.М. Бапиев, Ж.Ж. Багисов ИНФОРМАЦИОННАЯ СИСТЕМА ПОДГОТОВКИ ЗАЯВОК ДЛЯ УЧАСТИЯ В ГРАНТОВОМ ФИНАНСИРОВАНИИ УЧЕНЫХ ПО НАУЧНЫМ ПРОЕКТАМ.....	94
А.А. Иманберді, Р.Н. Молдашева ОБЗОР МОДЕЛЕЙ РАСПРОСТРАНЕНИЯ ИНФОРМАЦИИ В СОЦИАЛЬНЫХ СЕТЯХ.....	107

Г. Қалман, М.Ғ. Есмағанбет, М.М. Жаманқарин, А.Г. Габдулина, Д.В. Плескачев РЕШЕНИЕ КОРЕФЕРЕНЦИИ С ПОМОЩЬЮ МЕТОДА КЛАСТЕРИЗАЦИИ.....	121
Қ.Т. Қырғызбай, Е.Х. Какимжанов СОЗДАНИЕ БАЗЫ ГЕОДАНЫХ АЛМАТИНСКОЙ ОБЛАСТИ НА ОСНОВЕ ГИС-ТЕХНОЛОГИЙ О МЕТОДЕ ИДЕНТИФИКАЦИИ ПОДОЗРИТЕЛЬНЫХ ТРАНЗАКЦИЙ В БИТКОИН СЕТИ.....	136
Ш.Ж. Мусиралиева, М.Ж. Шайзат, А.К. Бекетова, Е. Абайұл, А.Б. Манасова О МЕТОДЕ ИДЕНТИФИКАЦИИ ПОДОЗРИТЕЛЬНЫХ ТРАНЗАКЦИЙ В БИТКОИН СЕТИ.....	154
А.Ұ. Мұхиядин, У.Т. Махажанова, М.У. Мукашева, А.А. Муханова ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ КАК СРЕДСТВО АНАЛИЗА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ ПРИ ЭКСТРЕННОМ ДИСТАНЦИОННОМ ОБУЧЕНИИ.....	170
А.Б. Токтарова, Б.С. Омаров, Г.Н. Казбекова, С.А. Мамиков, Ф.Е. Темирбекова СБОР БАЗЫ ДАННЫХ О ЯЗЫКЕ НЕНАВИСТИ В СОЦИАЛЬНОЙ СЕТИ НА КАЗАХСКОМ ЯЗЫКЕ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ.....	191
А.А. Шекербек, Г.Б. Абдикеримова, Ж.Б. Ламашева, М.Г. Байбулова, А.К. Токкулиева КЛАССИФИКАЦИЯ РЕНТГЕНОВСКИХ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ АЛГОРИТМА ГЛУБОКОГО ОБУЧЕНИЯ.....	204
Э.Э. Эльдарова УЛУЧШЕНИЕ ВИЗУАЛЬНОГО КАЧЕСТВА ЦИФРОВЫХ ИЗОБРАЖЕНИЙ ПОСЛЕ СЖАТИЕ JPEG2000.....	228

CONTENTS

J.K. Abdugulova, G.A. Uskenbayeva, M.N. Tlegen, A.K. Shukirova AUTOMATION OF THE TECHNOLOGICAL PROCESS OF HEATING OIL PIPELINE EQUIPMENT.....	5
Z. Avkurova, S. Gnatyuk, L. Kydyralina, N. Kurmangaliev THE INTELLECTUALIZED METHOD OF EARLY DETECTION AND IDENTIFICATION OF THE VIOLATOR IN INFORMATION AND COMMUNICATION SYSTEMS.....	22
A. Bekarystankyzy, O. Zh. Mamyrbayev INTEGRATED AUTOMATIC SPEECH RECOGNITION SYSTEM FOR AGGLUTINATIVE LANGUAGES.....	37
A. Yerimbetova, E. Daiyrbayeva, L. Cherikbayeva EMBEDDING HIDDEN INFORMATION IN IMAGES BASED ON BICUBIC INTERPOLATION.....	50
M. Yessenova, G. Abdikerimova, A. Tolstoy, Zh. Lamasheva, A. Nekessova APPLICABILITY OF TEXTURE IMAGE ANALYSIS METHODS FOR DETECTION OF WHEAT WEED POCKS.....	64
L. Zholshiyeva, T. Zhukabayeva, Sh. Turaev, M. Berdieva, R. Sengirbayeva REAL-TIME KAZAKH SIGN LANGUAGE RECOGNITION USING MEDIAPIPE AND SVM.....	82
Zh.S. Ixebayeva, K. Jetpisov, A.B. Medeshova, I.M. Bapiyev , Zh.Zh. Bagisov AN INFORMATION SYSTEM FOR THE PREPARATION OF APPLICATIONS FOR PARTICIPATION IN GRANT FUNDING OF SCIENTISTS IN SCIENTIFIC PROJECTS.....	94
A. Imanberdi, R. Moldasheva REVIEW OF MODELS OF DISSEMINATION OF INFORMATION IN SOCIAL NETWORKS.....	107
G. Kalman, M.G. Esmaganbet, M.M. Zhamankarin, A.I. Gabdulina, D.V. Pleskachev COREFERENCE SOLUTION USING THE CLUSTERING METHOD.....	121

K. Kyrgyzbay, E. Kakimzhanov CREATION OF A GEODATABASE OF ALMATY REGION BASED ON GIS TECHNOLOGIES.....	136
Sh. Mussiraliyeva, M. Shaizat, A. Beketova, Y. Abayuly, A. Manassova IDENTIFICATION OF SUSPICIOUS TRANSACTIONS IN THE BITCOIN NETWORK.....	154
A. Mukhiyadin, U. Makhazhanova, M. Mukasheva, A. Mukhanova INFORMATION TECHNOLOGIES AS A MEANS OF EXPERIMENTAL DATA ANALYSIS IN EMERGENCY DISTANCE LEARNING.....	170
A.B. Toktarova, B.S. Omarov, G.N. Kazbekova, S.A. Mamikov, F.E. Temirbekova COLLECTING HATE SPEECH DATABASE ON SOCIAL NETWORK IN KAZAKH LANGUAGE BY USING MACHINE LEARNING.....	191
A. Shekerbek, G. Abdikerimova, Zh. Lamasheva, M. Baibulova, A. Tokkuliyeva CLASSIFICATION OF X-RAY IMAGES USING THE DEEP LEARNING ALGORITHM.....	204
E.E. Eldarova IMPROVING THE VISUAL QUALITY OF DIGITAL IMAGES AFTER JPEG2000 COMPRESSION.....	228

**Publication Ethics and Publication Malpractice
the journals of the National Academy of Sciences of the Republic of Kazakhstan**

For information on Ethics in publishing and Ethical guidelines for journal publication see <http://www.elsevier.com/publishingethics> and <http://www.elsevier.com/journal-authors/ethics>.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see <http://www.elsevier.com/postingpolicy>), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service <http://www.elsevier.com/editors/plagdetect>.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:

www.nauka-nanrk.kz

<http://physics-mathematics.kz/index.php/en/archive>

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Заместитель директора отдела издания научных журналов НАН РК *Р. Жалиқызы*

Редакторы: *М.С. Ахметова, Д.С. Аленов*

Верстка на компьютере *Г.Д. Жадыранова*

Подписано в печать 30.03.2023.

Формат 60x881/8. Бумага офсетная. Печать – ризограф.

15,5 п.л. Тираж 300. Заказ 1.