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

«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ
АКАДЕМИЯСЫ» РҚБ

«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ» РҚБ

ХАБАРЛАРЫ

ИЗВЕСТИЯ

РОО «НАЦИОНАЛЬНОЙ
АКАДЕМИИ НАУК РЕСПУБЛИКИ
КАЗАХСТАН»

NEWS

OF THE NATIONAL ACADEMY
OF SCIENCES OF THE REPUBLIC
OF KAZAKHSTAN

SERIES OF PHYSICS AND MATHEMATICS

1 (353)

JANUARY – MARCH 2025

PUBLISHED SINCE JANUARY 1963
PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

БАС РЕДАКТОР:

МҰТАНОВ Ғалымқайыр Мұтанұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының м.а. (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=6506682964, https://www.webofscience.com/wos/author/record/1423665

РЕДАКЦИЯ АЛҚАСЫ:

ҚАЛИМОЛДАЕВ Мақсат Нұрәділұлы, (бас редактордың орынбасары), физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының кеңесшісі, зертхана меңгерушісі (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=56153126500, https://www.webofscience.com/wos/author/record/2428551

МАМЫРБАЕВ Өркен Жұмажанұлы (ғалым хатшы), Ақпараттық жүйелер саласындағы техника ғылымдарының (PhD) докторы, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» директорының ғылым жөніндегі орынбасары (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=55967630400, https://www.webofscience.com/wos/author/record/1774027

БАЙГУНЧЕКОВ Жұмаділ Жаңабайұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Кибернетика және ақпараттық технологиялар институты, Қолданбалы механика және инженерлік графика кафедрасы, Сәтбаев университеті (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=6506823633, https://www.webofscience.com/wos/author/record/1923423

ВОЙЧИК Вальдемар, техника ғылымдарының докторы (физ-мат), Люблин технологиялық университетінің профессоры (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=7005121594, https://www.webofscience.com/wos/author/record/678586

СМОЛАРЖ Анджей, Люблин политехникалық университетінің электроника факультетінің доценті (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=56249263000, https://www.webofscience.com/wos/author/record/1268523

КЕЙЛАН Әлімхан, техника ғылымдарының докторы, профессор (ғылым докторы (Жапония)), ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» бас ғылыми қызметкері (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=8701101900, https://www.webofscience.com/wos/author/record/1436451

ХАЙРОВА Нина, техника ғылымдарының докторы, профессор, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» бас ғылыми қызметкері (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=37461441200, https://www.webofscience.com/wos/author/record/1768515

ОТМАН Мохаммед, PhD, Информатика, Коммуникациялық технологиялар және желілер кафедрасының профессоры, Путра университеті Малайзия (Селангор, Малайзия), https://www.scopus.com/authid/detail.uri?authorId=56036884700, https://www.webofscience.com/wos/author/record/747649

НЫСАНБАЕВА Сауле Еркебұланқызы, техника ғылымдарының докторы, доцент, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» аға ғылыми қызметкері (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=55453992600, https://www.webofscience.com/wos/author/record/3802041

БИЯШЕВ Рустам Гакашевич, техника ғылымдарының докторы, профессор, Информатика және басқару мәселелері институты директорының орынбасары, Ақпараттық қауіпсіздік зертханасының меңгерушісі (Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=6603642864, https://www.webofscience.com/wos/author/record/3802016

КАПАЛОВА Нұрсұлу Алдажарқызы, техника ғылымдарының кандидаты, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты», Киберқауіпсіздік зертханасының меңгерушісі (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=57191242124,

КОВАЛЕВ Александр Михайлович, физика-математика ғылымдарының докторы, Украина Ұлттық Ғылым академиясының академигі, Қолданбалы математика және механика институты (Донецк, Украина), https://www.scopus.com/authid/detail.uri?authorId=7202799321, https://www.webofscience.com/wos/author/record/38481396

МИХАЛЕВИЧ Александр Александрович, техника ғылымдарының докторы, профессор, Беларусь Ұлттық Ғылым академиясының академигі (Минск, Беларусь), https://www.scopus.com/authid/detail.uri?authorId=7004159952, https://www.webofscience.com/wos/author/record/46249977

ТИГИНЯНУ Ион Михайлович, физика-математика ғылымдарының докторы, академик, Молдова Ғылым Академиясының президенті, Молдова техникалық университеті (Кишинев, Молдова), https://www.scopus.com/authid/detail.uri?authorId=7006315935, https://www.webofscience.com/wos/author/record/524462

«ҚР ҰҒА Хабарлары. Физика-математика сериясы».
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Меншіктеуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы).
Ақпарат агенттігінің мерзімді баспасөз басылымын, ақпарат агенттігін және желілік басылымды қайта есепке қою туралы ҚР Мәдениет және Ақпарат министрлігі «Ақпарат комитеті» Республикалық мемлекеттік мекемесі **28.02.2025** ж. берген **№KZ20VPY00113741** Куәлік.
Тақырыптық бағыты: *ақпараттық-коммуникациялық технологиялар*
Қазіргі уақытта: *«ақпараттық-коммуникациялық технологиялар» бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*
Мерзімділігі: *жылына 4 рет.*
Редакцияның мекен-жайы: *050010, Алматы қ., Шевченко көш., 28, 219 бөл., тел.: 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

© «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ, 2025

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), https://www.scopus.com/authid/detail.uri?authorId=6506682964, https://www.webofscience.com/wos/author/record/1423665

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), https://www.scopus.com/authid/detail.uri?authorId=56153126500, https://www.webofscience.com/wos/author/record/2428551

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), https://www.scopus.com/authid/detail.uri?authorId=55967630400, https://www.webofscience.com/wos/author/record/1774027

БАЙГУНЧЕКОВ Жумадил Жанабаевич, доктор технических наук, профессор, академик НАН РК, Институт кибернетики и информационных технологий, кафедра прикладной механики и инженерной графики, Университет Сатпаева (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=6506823633, https://www.webofscience.com/wos/author/record/1923423

ВОЙЧИК Вальдемар, доктор технических наук (физ.-мат.), профессор Люблинского технологического университета (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=7005121594, https://www.webofscience.com/wos/author/record/678586

СМОЛАРЖ Анджей, доцент факультета электроники Люблинского политехнического университета (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=56249263000, https://www.webofscience.com/wos/author/record/1268523

КЕЙЛАН Алимхан, доктор технических наук, профессор (Doctor of science (Japan)), главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=8701101900, https://www.webofscience.com/wos/author/record/1436451

ХАЙРОВА Нина, доктор технических наук, профессор, главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=37461441200, https://www.webofscience.com/wos/author/record/1768515

ОТМАН Мохамед, доктор философии, профессор компьютерных наук, Департамент коммуникационных технологий и сетей, Университет Путра Малайзия (Селангор, Малайзия), https://www.scopus.com/authid/detail.uri?authorId=56036884700, https://www.webofscience.com/wos/author/record/747649

НЫСАНБАЕВА Сауле Еркебулановна, доктор технических наук, доцент, старший научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=55453992600, https://www.webofscience.com/wos/author/record/3802041

БИЯШЕВ Рустам Гакашевич, доктор технических наук, профессор, заместитель директора Института проблем информатики и управления, заведующий лабораторией информационной безопасности (Казахстан), https://www.scopus.com/authid/detail.uri?authorId=6603642864, https://www.webofscience.com/wos/author/record/3802016

КАПАЛОВА Нурсулу Алдажаровна, кандидат технических наук, заведующий лабораторией кибербезопасности РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=57191242124,

КОВАЛЕВ Александр Михайлович, доктор физико-математических наук, академик НАН Украины, Институт прикладной математики и механики (Донецк, Украина), https://www.scopus.com/authid/detail.uri?authorId=7202799321, https://www.webofscience.com/wos/author/record/38481396

МИХАЛЕВИЧ Александр Александрович, доктор технических наук, профессор, академик НАН Беларуси (Минск, Беларусь), https://www.scopus.com/authid/detail.uri?authorId=7004159952, https://www.webofscience.com/wos/author/record/46249977

ТИГИНЯНУ Ион Михайлович, доктор физико-математических наук, академик, президент Академии наук Молдовы, Технический университет Молдовы (Кишинев, Молдова), https://www.scopus.com/authid/detail.uri?authorId=7006315935, https://www.webofscience.com/wos/author/record/524462

«Известия НАН РК. Серия физико-математическая».
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

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

Свидетельство о постановке на переучет периодического печатного издания, информационного агентства и сетевого издания **№ KZ20VPY00113741.** Дата выдачи **28.02.2025**

Тематическая направленность: *информационно-коммуникационные технологии.*

В настоящее время: *вошел в список журналов, рекомендованных КОКСНВО МНВО РК по направлению «информационно-коммуникационные технологии».*

Периодичность: *4 раза в год.*

Адрес редакции: *050010, г. Алматы, ул. Шевченко, 28, оф. 219, тел.: 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

© РОО «Национальная академия наук Республики Казахстан», 2025

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), https://www.scopus.com/authid/detail.uri?authorId=6506682964, https://www.webofscience.com/wos/author/record/1423665

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), https://www.scopus.com/authid/detail.uri?authorId=56153126500, https://www.webofscience.com/wos/author/record/2428551

Mamyrbaev 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), https://www.scopus.com/authid/detail.uri?authorId=55967630400, https://www.webofscience.com/wos/author/record/1774027

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), https://www.scopus.com/authid/detail.uri?authorId=6506823633, https://www.webofscience.com/wos/author/record/1923423

WOICIK Waldemar, Doctor of Technical Sciences (Phys.-Math.), Professor of the Lublin University of Technology (Lublin, Poland), https://www.scopus.com/authid/detail.uri?authorId=7005121594, https://www.webofscience.com/wos/author/record/678586

SMOLARJ Andrej, Associate Professor Faculty of Electronics, Lublin polytechnic university (Lublin, Poland), https://www.scopus.com/authid/detail.uri?authorId=56249263000, https://www.webofscience.com/wos/author/record/1268523

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), https://www.scopus.com/authid/detail.uri?authorId=8701101900, https://www.webofscience.com/wos/author/record/1436451

KHAIROVA Nina, Doctor of Technical Sciences, Professor, Chief Researcher of the Institute of Information and Computational Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=37461441200, https://www.webofscience.com/wos/author/record/1768515

OTMAN Mohamed, PhD, Professor of Computer Science Department of Communication Technology and Networks, Putra University Malaysia (Selangor, Malaysia), https://www.scopus.com/authid/detail.uri?authorId=56036884700, https://www.webofscience.com/wos/author/record/747649

NYSANBAYEVA Saule Yerkebulanovna, Doctor of Technical Sciences, Associate Professor, Senior Researcher of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=55453992600, https://www.webofscience.com/wos/author/record/3802041

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), https://www.scopus.com/authid/detail.uri?authorId=6603642864, https://www.webofscience.com/wos/author/record/3802016

KAPALOVA Nursulu Aldazharovna, Candidate of Technical Sciences, Head of the Laboratory cybersecurity, Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=57191242124,

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), https://www.scopus.com/authid/detail.uri?authorId=7202799321, https://www.webofscience.com/wos/author/record/38481396

MIKHALEVICH Alexander Alexandrovich, Doctor of Technical Sciences, Professor, Academician of the National Academy of Sciences of Belarus (Minsk, Belarus), https://www.scopus.com/authid/detail.uri?authorId=7004159952, https://www.webofscience.com/wos/author/record/46249977

TIGHINEANU Ion Mihailovich, Doctor of Physical and Mathematical Sciences, Academician, President of the Academy of Sciences of Moldova, Technical University of Moldova (Chisinau, Moldova), https://www.scopus.com/authid/detail.uri?authorId=7006315935, https://www.webofscience.com/wos/author/record/524462

News of the National Academy of Sciences of the Republic of Kazakhstan.
Series of Physics and Mathematics
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Owner: RPA «National Academy of Sciences of the Republic of Kazakhstan» (Almaty).
Certificate **No. KZ20VPY00113741** on the re-registration of the periodical printed and online publication of the information agency, issued on **28.02.2025** by the Republican State Institution «Information Committee» of the Ministry of Culture and Information of the Republic of Kazakhstan
Subject area*: information and communication technologies.*
Currently: *included in the list of journals recommended by the CCSES MSHE RK in the direction of «Information and communication technologies».*
Periodicity: *4 times a year.*
Editorial address: *28, Shevchenko str., of. 219, Almaty, 050010, tel. 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

© National Academy of Sciences of the Republic of Kazakhstan, 2025

NEWS OF THE NATIONAL ACADEMY OF SCIENCESOF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 1. Namber 353 (2025). 32–48

https://doi.org/10.32014/2025.2518-1726.323

IRSTI 28.23.15
UDC 004.8

A.B. Aben, N.M. Zhunissov*, G.N. Kazbekova, A.N. Amanov,
A.A. Abibullayeva, 2025.
Khoja Akhmet Yassawi International Kazakh-Turkish University,
Turkistan, Kazakhstan.
*E-mail: nurseit.zhunissov@ayu.edu.kz

DEEPFAKE ARTIFICIAL VOICE DETECTION. COMPARISON OF THE EFFECTIVENESS OF THE LSTM AND CNN MODELS

Aben Arypzhan Baktiarovich – Doctoral student in the educational program «Information systems», Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan, E-mail: arypzhan.aben@ayu.edu.kz, https://orcid.org/0000-0001-8534-3288;
Zhunissov Nurseit Mukhidinovich – PhD, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan, E-mail: nurseit.zhunissov@ayu.edu.kz, https://orcid.org/0000-0001-6531-9408,
Kazbekova Gulnur Nagimetovna – Candidate of Technical Sciences, Associate Professor, International Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan, E-mail: gulnur.kazbekova@ayu.edu.kz, https://orcid.org/0000-0002-2756-7926;
Amanov Anuarbek Nurseyitovich – PhD, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan, E-mail: anuarbek.amanov@ayu.edu.kz, https://orcid.org/0000-0003-0638-6859;
Abibullayeva Aiman Abibullakuzu – PhD, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan, E-mail: aiman.abibullayeva@ayu.edu.kz, https://orcid.org/(0000-0003-2449-2540.

Abstract. This research presents a novel methodology for detecting DeepFake voices, which is based on the effective classification of fake and real audio signals. To enhance the assessment of information in the audience, the voices of 58 politicians and public figures were compiled as fake and real audio files. In the study, fake audio samples were artificially generated, while real samples were obtained from authentic sources. The analysis of the audio signal structure employed Mel-Frequency Cepstral Coefficients (MFCC), Zero-Crossing Rate (ZCR) metrics, and data visualization techniques, including bar charts and histograms.

During the research, the numerical distribution, lengths, MFCC features, and ZCR values of the fake and real audio samples were analyzed. LSTM and CNN models were tested for DeepFake voice detection, resulting in the LSTM model achieving 100% accuracy, while the CNN model was rated at 97.50% accuracy. The findings demonstrated that the LSTM model could accurately and reliably distinguish between fake and real audio, emphasizing the importance of assessing the authenticity of audio signals in light of the dangers posed by DeepFake technology.

This research provides functional methodologies aimed at developing systems for visual individuals while also uncovering new ways to determine the authenticity of audio signals and demonstrating the effectiveness of applying modern deep learning technologies. The study emphasizes that DeepFake plays a significant role in assessing and identifying information in an audience and provides a foundation for future research and practice.

Keywords: DeepFake, Voice Classification, Audio Signals, Mel-Frequency Cepstral Coefficients (MFCC), Zero-Crossing Rate (ZCR), LSTM Model, CNN Model.

А.Б. Абен, Н.М. Жунисов*, Г.Н. Казбекова, А.Н. Аманов,
А.А. Абибуллаева, 2025.
Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті,
Түркістан, Қазақстан.
*E-mail: nurseit.zhunissov@ayu.edu.kz

DEEPFAKE ЖАСАНДЫ ДАУЫСТЫ АНЫҚТАУ. LSTM ЖӘНЕ CNN МОДЕЛЬДЕРІНІҢ ТИІМДІЛІГІ САЛЫСТЫРУ

Абен Арыпжан Бактиярович – «Ақпараттық жүйелер» білім беру бағдарламасы бойынша докторант, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан, E-mail: arypzhan.aben@ayu.edu.kz, https://orcid.org/0000-0001-8534-3288;
Жунисов Нурсейт Мухидинович – PhD доктор, аға оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан· E-mail: nurseit.zhunissov@ayu.edu.kz, https://orcid.org/0000-0001-6531-9408;
Казбекова Гулнур Нагиметовна – техника ғылымдарының кандидаты, доцент м.а., Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан, E-mail: gulnur.kazbekova@ayu.edu.kz, https://orcid.org/0000-0002-2756-7926;
Аманов Ануарбек Нурсейтович – PhD доктор, аға оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан· E-mail: anuarbek.amanov@ayu.edu.kz, https://orcid.org/0000-0003-0638-6859;
Абибуллаева Айман Абибуллақызы – PhD доктор, аға оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан, E-mail: aiman.abibullayeva@ayu.edu.kz , https://orcid.org/(0000-0003-2449-2540.

Аннотация. Зерттеу DeepFake дауысын анықтауға арналған жаңа әдістемені ұсынады, ол фейк және реал аудио сигналдарын тиімді классификациялауға негізделген. Аудиториядағы ақпаратты бағалауды жақсарту мақсатында 58 саясаткер мен танымал тұлғалардың дауыстары фейк және реал аудио файлдары ретінде жинақталды. Зерттеуде фейк аудио үлгілері жасанды түрде жасалған, ал реал үлгілері шынайы дереккөздерден алынды. Аудио сигналдардың құрылымын талдау үшін Mel-Frequency Cepstral Coefficients (MFCC) әдісі, Zero-Crossing Rate (ZCR) көрсеткіші және деректердің визуализациясы қолданылды.

Зерттеу барысында фейк және реал аудио үлгілерінің сандық таралуы, ұзындықтары, MFCC ерекшеліктері және ZCR мәндері талданды. LSTM және CNN модельдері DeepFake дауысын анықтау үшін сыналды, нәтижесінде LSTM моделі 100% дәлдікпен, ал CNN моделі 97.50% дәлдікпен бағаланды. Алынған нәтижелер LSTM моделінің фейк және реал аудионы дәл және сенімді түрде анықтай алатынын көрсетті, бұл DeepFake технологиясының қауіптілігін ескере отырып, аудио сигналдардың шынайылығын бағалаудың маңыздылығын білдіреді.

Бұл зерттеу нәтижелері визуалды тұлғаларға арналған жүйелерді дамытуға негізделген функционалдық әдістемелерді ұсынумен қатар, аудио сигналдардың шынайылығын анықтаудың жаңа жолдарын ашып, қазіргі заманғы терең оқыту технологияларын қолданудың тиімділігін дәлелдейді. Зерттеу DeepFake аудиториядағы ақпаратты бағалау мен идентификациялауда маңызды рөл атқаратынын атап өтеді және болашақ зерттеулер мен практикаға арналған негіздерді қамтамасыз етеді.

Түйін сөздер: DeepFake, Дауыс классификациясы, Аудио сигналдар, Mel-Frequency Cepstral Coefficients (MFCC), Zero-Crossing Rate (ZCR), LSTM моделі, CNN моделі.

А.Б. Абен, Н.М. Жунисов*, Г.Н. Казбекова, А.Н. Аманов,
А.А. Абибуллаева, 2025.
Международный казахско-турецкий университет имени Ходжи Ахмеда
Ясави, Туркестан, Казахстан.
*E-mail: nurseit.zhunissov@ayu.edu.kz

ОБНАРУЖЕНИЕ ИСКУССТВЕННОГО ГОЛОСА DEEPFAKE. СРАВНЕНИЕ ЭФФЕКТИВНОСТИ МОДЕЛЕЙ LSTM И CNN

Абен Арыпжан Бактиярович – докторант по образовательной программе «Информационные системы», Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан, E-mail: arypzhan.aben@ayu.edu.kz, https://orcid.org/0000-0001-8534-3288;
Жунисов Нурсейт Мухидинович – PhD, старший преподаватель, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан, E-mail: nurseit.zhunissov@ayu.edu.kz, https://orcid.org/0000-0001-6531-9408;
Казбекова Гулнур Нагиметовна – кандидат технических наук, и.о. доцента, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан, E-mail: gulnur.kazbekova@ayu.edu.kz, https://orcid.org/0000-0002-2756-7926;
Аманов Ануарбек Нурсейтович – PhD, старший преподаватель, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан, E-mail: anuarbek.amanov@ayu.edu.kz, https://orcid.org/0000-0003-0638-6859;
Абибуллаева Айман Абибуллакызы – PhD, старший преподаватель, Международный

казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан, E-mail: aiman.abibullayeva@ayu.edu.kz, https://orcid.org/(0000-0003-2449-2540.

Аннотация. Данное исследование представляет новую методику для обнаружения голосов DeepFake, основанную на эффективной классификации фейковых и реальных аудиосигналов. С целью улучшения оценки информации в аудитории были собраны аудиозаписи голосов 58 политиков и публичных фигур, содержащие как реальные, так и фейковые аудиофайлы. В исследовании фейковые аудиопримеры были искусственно созданы, в то время как реальные образцы были получены из достоверных источников. Для анализа структуры аудиосигналов использовались коэффициенты мел-частотного кепстра (MFCC), метрика нулевого пересечения (ZCR) и визуализация данных, включая столбчатые диаграммы и гистограммы.

В ходе исследования была проанализирована распределение числовых значений, длины, особенности MFCC и значения ZCR фейковых и реальных аудиопримеров. Модели LSTM и CNN были протестированы для обнаружения голосов DeepFake, в результате чего модель LSTM достигла 100% точности, а модель CNN была оценена на уровне 97,50% точности. Полученные результаты продемонстрировали, что модель LSTM может точно и надежно различать фейковые и реальные аудиозаписи, подчеркивая важность оценки подлинности аудиосигналов с учетом опасностей, связанных с технологией DeepFake.

Данное исследование предлагает функциональные методологии, направленные на разработку систем с визуальными методами анализа данных, а также открывает новые способы определения подлинности аудиосигналов и демонстрирует эффективность применения современных технологий глубокого обучения. Исследование подчеркивает, что DeepFake играет важную роль в оценке и идентификации информации в аудитории и предоставляет основы для будущих исследований и практики.

Ключевые слова: DeepFake, классификация голосов, аудиосигналы, мел-частотные кепстральные коэффициенты (MFCC), нулевая частота пересечения (ZCR), модель LSTM, модель CNN.

Introduction. The digital media and communication technologies of the modern era have developed rapidly, and their impact has deeply penetrated various sectors of society (Karnouskos, 2020). In recent years, advancements in artificial intelligence and deep learning have paved the way for increasingly sophisticated methods of information manipulation (Al-Khazraji, et al., 2023). Among these, DeepFake technologies stand out, as they enable the artificial generation of a person’s voice or image with high accuracy (Mullen, 2023). DeepFake voice synthesis can imitate a real person’s voice and be used for various dangerous purposes, ranging from spreading misinformation to cyberattacks, which significantly threatens information security and public trust (Kumar & Kundu, 2024).

In such circumstances, the issue of detecting DeepFake voices is considered not only a technical challenge but also a socially significant task. The application areas of DeepFake voice detection technologies are very broad, and they play a crucial role in various sectors of society. Below, in Figure 1, the main application areas of these technologies are presented.

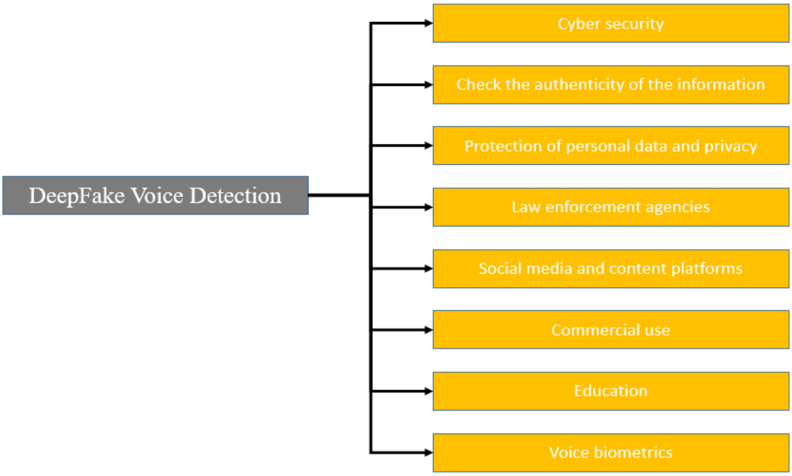

Figure 1. Fields of application of DeepFake voice detection.

Combating voice manipulation goes beyond merely addressing security issues; it also serves the goals of protecting personal privacy, enhancing the effectiveness of law enforcement systems, and ensuring the authenticity of information. The need to explore new methods and models to achieve these objectives is increasing, and in this regard, the potential of deep learning technologies plays a particularly important role.

DeepFake voice detection has become an essential tool with deep learning models such as Convolutional Neural Networks (CNN) (Patel, et al., 2023) and Recurrent Neural Networks (RNN) (Al-Dhabi & Zhang, 2021). CNN models are effective in identifying visual features by processing the spectrograms of audio signals, while RNN models excel in examining dynamic changes over time and understanding the structure of temporal data. However, there is a pressing need for comprehensive comparative studies on the ability of both models to accurately and reliably detect DeepFake voices.

This research addresses significant questions in the field of DeepFake voice detection and compares the performance of CNN and RNN models. The goal of the study is to identify an effective method through an in-depth investigation of the advantages and limitations of these two models, enabling their application in specific fields. Furthermore, the proposed approaches aim to contribute significantly to cybersecurity, combatting misinformation, and verifying the authenticity of audio media.

The results of this study will facilitate the improvement of DeepFake voice detection methods, enhancing the reliability of the information space and providing effective solutions against new threats.

Research on DeepFake Voice Detection Technologies

Research on DeepFake voice detection technologies plays a crucial role in the current field of information security. Various methods and approaches used in this area are essential for identifying the complexities of voice manipulation and effectively combating them. Signal processing techniques are fundamental in the initial phase of DeepFake voice detection, as they form the basis for analyzing and processing audio signals. Tasks such as filtering sound signals, removing noise, and generating spectrograms enhance the effectiveness of detecting manipulated content.

Machine learning methods, including models such as Support Vector Machine (SVM), Random Forest, and k-Nearest Neighbors (k-NN), enable the analysis of large volumes of data (Hamza, et al. 2022). These methods are widely used for pattern recognition and detecting manipulated audio. However, deep learning methods, particularly Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), demonstrate exceptional effectiveness in DeepFake voice detection (Al-Adwan et al., 2024). These models are designed for processing the visual features of audio signals and understanding the structure of temporal data.

A new direction in identifying manipulated content involves leveraging the synchronization of audio and text (Wang, et al, 2022). This approach enables the disruption of DeepFake voice technologies by analyzing the coherence between audio and text data (Agarwal & Farid, 2021). The presence of synchronization, meaning the correspondence between audio and text, serves as a critical indicator in detecting DeepFake manipulations (Bohacek, & Farid, 2024).

Hybrid methods enhance the effectiveness of DeepFake voice detection by combining signal processing, machine learning, and deep learning techniques. These approaches aim to achieve optimal results by integrating the advantages of both traditional and modern technologies(Saikia et al, 2022). Hybrid models offer a comprehensive strategy(Kaddar et al., 2021), enabling improved efficiency in detecting DeepFake voices (Cho, et al., 2023).

In conclusion, the development of DeepFake voice detection technologies requires the integration of signal processing, machine learning, deep learning, audio and text synchronization, and hybrid approaches. Enhancing the effectiveness of these methods contributes to the development of new tools and strategies aimed at ensuring information security and mitigating the risks posed by DeepFake content.

Methods and materials. In this study, we utilize an architecture that leverages two models—LSTM and CNN—to detect DeepFake audio, with a dataset organized into two folders: one for fake audio and one for real audio (**Figure 2**). The performance of these models is compared to identify the best approach for protecting against DeepFake voice threats.

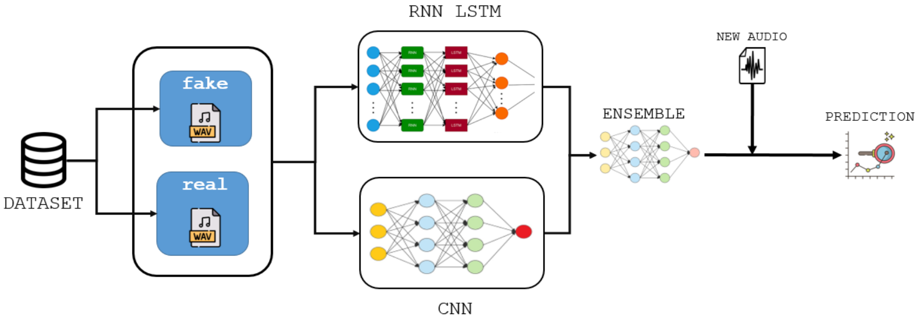

Figure 2. *LSTM Architecture for DeepFake Voice Detection*

Dataset

In our research, we utilized the «In-the-Wild» dataset, which comprises a comprehensive collection of audio deepfakes alongside corresponding bona-fide audio recordings. This dataset was specifically curated to include a diverse array of public figures, including 58 celebrities and politicians, ensuring a wide representation of vocal characteristics and speech patterns.

The dataset was sourced from publicly available platforms such as social networks and video streaming sites, enabling the collection of realistic audio samples reflective of natural speech. In total, the dataset features 20.8 hours of bona-fide audio and 17.2 hours of spoofed audio. On average, each speaker is represented by approximately 23 minutes of bona-fide audio and 18 minutes of spoofed audio, providing a robust foundation for evaluating deepfake detection and voice anti-spoofing machine-learning models (Cavia, et al., 2024).

This dataset serves as a critical resource for assessing the generalization capabilities of various detection models when exposed to realistic, in-the-wild audio samples. Its design facilitates the exploration of model performance across diverse audio scenarios, ultimately contributing to the advancement of deepfake detection technology.

For our experiments, we employed several noteworthy deepfake detection models that are open-source and available on GitHub, including RawNet 2, RawGAT-ST, and PC-Darts. These models were selected based on their relevance and effectiveness in addressing the challenges presented by the dataset (Baxevanakis, et al., 2022).

The dataset, along with its accompanying documentation, is licensed under the Apache License, Version 2.0, ensuring that it remains accessible for further research and development in the field of audio deepfake detection. For additional information regarding the dataset and its application, please refer to our published paper and the provided download link.

Convolutional Neural Networks (CNNs)

In this study, we employed Convolutional Neural Networks (CNNs) as a pivotal method for detecting DeepFake audio(Ahmed, et al., 2022). CNNs have gained prominence in the fields of image and audio processing due to their ability to learn and extract hierarchical features from input data effectively. Their architecture is particularly suited for analyzing audio signals represented as spectrograms, as they can capture both local and global features crucial for distinguishing between bona-fide and spoofed audio samples.

The methodological approach commenced with transforming raw audio recordings from the «In-the-Wild» dataset into spectrogram representations (Mcuba, et al., 2023). Utilizing the Short-Time Fourier Transform (STFT), we generated time-frequency representations that encapsulate the spectral characteristics of the audio signals. This transformation is fundamental, as it enables the CNN to leverage the intricate patterns inherent in the audio, which may indicate manipulation(Li, et al., 2022).

The architecture of the CNN utilized in our experiments is structured to facilitate a comprehensive feature extraction process. The input layer receives spectrograms formatted as $X \in R^{H \times W}$, where H represents the height (frequency bins) and W denotes the width (time frames) of the spectrogram. The subsequent convolutional layers are tasked with applying a series of learnable filters (kernels) K to the input data. Each filter $k \in K$ is convolved with the input X to generate feature maps F, which highlight relevant patterns in the audio:

$$F_{ij} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{i+m,j+n} \cdot k_{m,n}$$

Here, M and N signify the dimensions of the filter, allowing the model to learn distinct features associated with each audio class.

To introduce non-linearity into the model, we employed the Rectified Linear Unit (ReLU) activation function, which has been shown to enhance the learning capacity of deep networks:

$$A(x) = \max(0, x)$$

Following the convolutional layers, max pooling operations were employed to reduce the dimensionality of the feature maps. This reduction not only enhances computational efficiency but also aids in achieving translational invariance, which is crucial for audio processing:

$$P_{ij} = \max_{m,n} F_{(2i+m)(2j+n)}$$

The pooled feature maps are then flattened and fed into fully connected layers. These layers integrate the learned features and generate the final classification probabilities using the softmax function, which outputs a probability distribution over the two classes (bona-fide and spoofed audio):

$$Y = \sigma(WX + b)$$

Where W is the weight matrix, b represents the bias, and σ\sigmaσ denotes the softmax activation function.

To ensure the robustness of our model, we implemented various data augmentation strategies during training, such as random cropping, flipping, and time-stretching of the spectrograms. This approach aimed to enhance the generalization capabilities of the CNN and mitigate overfitting risks. The performance of the CNN was rigorously evaluated using standard metrics, including accuracy, precision, recall, and F1-score, which collectively provide a comprehensive assessment of the model's efficacy in distinguishing between genuine and manipulated audio samples.

The deployment of CNNs in this research highlights their critical role in addressing the challenges posed by audio deepfake detection. Our findings demonstrate that CNNs significantly improve detection accuracy while offering a robust framework for real-world applications in audio forensics and voice authentication systems. Through this investigation, we aim to contribute to the broader field of audio signal processing, paving the way for future advancements in the detection and mitigation of audio manipulations (Figure 3).

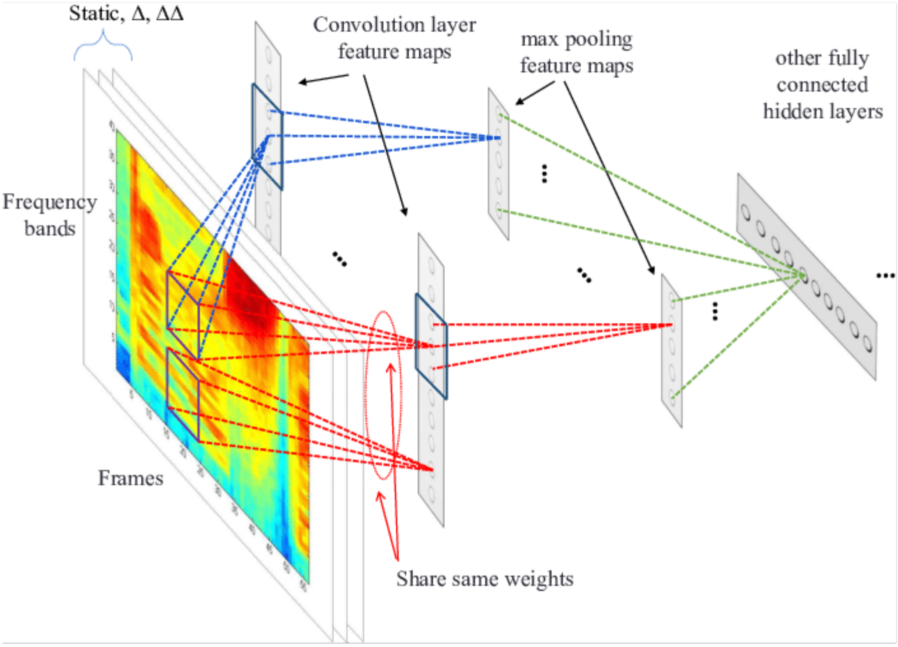

Figure 3. Structure of the CNN Architecture

DeepFake Voice Detection Using RNN LSTM

In our approach to detecting DeepFake audio, we utilized Recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) units. RNNs, and particularly LSTM networks, are well-suited for sequential data processing, making them an effective choice for audio analysis. Their ability to capture temporal dependencies is crucial for distinguishing between bona-fide and spoofed audio, as the subtle patterns indicating manipulation may only become evident over time.

The choice of LSTM over traditional RNNs is driven by the challenges associated with learning long-range dependencies. In standard RNNs, the problem of vanishing or exploding gradients can hinder the model's ability to learn from long sequences (Alshehri, et al., 2024). LSTM networks, on the other hand, address these issues by introducing a memory cell that maintains a persistent state, along with gating mechanisms that control the flow of information. This design enables the LSTM to remember essential features over extended time frames while forgetting irrelevant details, thus improving the model's performance in detecting audio deepfakes.

The raw audio data from the «In-the-Wild» dataset was first converted into Mel-spectrograms to capture the time-frequency characteristics. The Mel-spectrograms serve as a compact representation of the audio signals, highlighting relevant features that the LSTM can leverage for detection. The resulting sequences were then normalized to ensure consistency across different samples.

The LSTM network takes the sequences of Mel-spectrogram features as input, where each time step represents a feature vector corresponding to a specific frame in the audio. The network's architecture includes multiple LSTM layers stacked sequentially to learn both short-term and long-term temporal dependencies within the data. Each LSTM cell comprises three gates: the input gate, the forget gate, and the output gate, which are mathematically expressed as follows:

Controls the extent to which new information is added to the cell state.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

Determines how much of the previous cell state is retained.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

Regulates the output based on the current cell state.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

Combines the new information and the retained information to update the cell state.

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

Generates the hidden state for the current time step.

$$h_t = o_t \cdot \tanh(C_t)$$

I'll also provide a Figure 4 that illustrates the LSTM architecture for DeepFake voice detection.

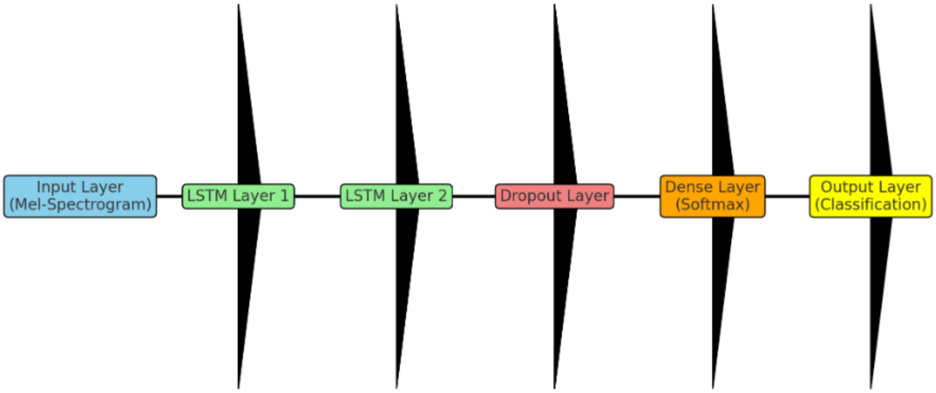

Figure 4. LSTM Architecture for DeepFake Voice Detection.

Here is a diagram depicting the LSTM architecture for DeepFake voice detection:

- **Input Layer**: Takes the Mel-spectrogram features derived from the audio data.
- **LSTM Layers**: Two LSTM layers are used to capture temporal dependencies in the audio features.
- **Dropout Layer**: Added to prevent overfitting by randomly disabling some connections during training.
- **Dense Layer with Softmax Activation**: Converts the output to a probability distribution across the classes.
- **Output Layer**: Provides the final classification, distinguishing between bona fide and spoofed audio.

This architecture is well-suited for detecting patterns in audio sequences, allowing the model to effectively distinguish between real and DeepFake voices.

Results. During the research, several visualization methods were used to analyze the structure and characteristics of the audio data. The results provide the necessary information for training models aimed at detecting DeepFake voices.

Initially, the dataset was composed of fake and real audio files. A total of 58 voices from politicians and public figures were collected. The fake audio samples were artificially created, while the real audio samples were obtained from genuine sources. During data processing, parameters such as the length of the audio files and MFCC features were calculated.

The quantitative distribution of fake and real audio samples in the collected dataset was analyzed. A bar chart was utilized to illustrate the proportion between classes, indicating whether the examined audio files belong to the fake or real category. This chart allows for the determination of the balance level within the dataset (Figure 5).

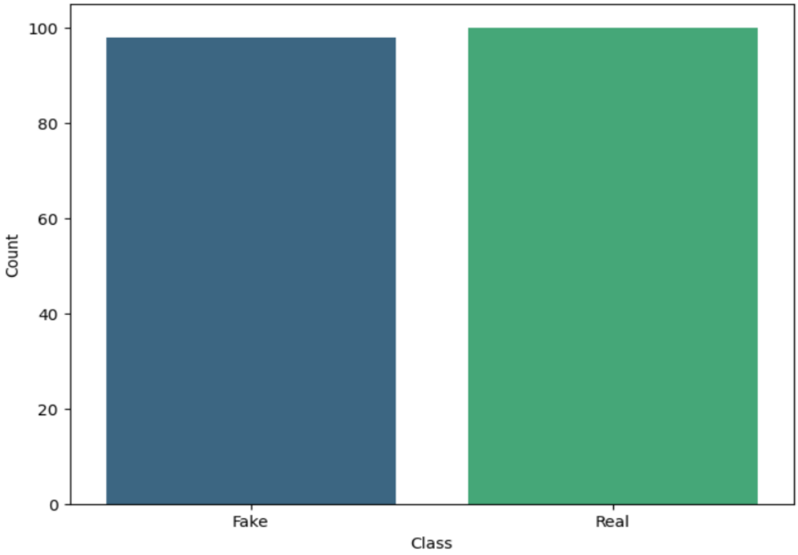

Figure 5. Number of Audio Samples per Class.

The bar chart displays the quantitative distribution of audio samples according to fake and real classes, allowing for the assessment of the impact of data imbalance on the classification model.

The lengths of the audio files vary over time, and analyzing their distribution is crucial for determining the authenticity of the audio. During the study, the distribution of lengths for both fake and real audio files was presented in the form of a histogram. This analysis helps to identify certain length characteristics typical of fake audio.

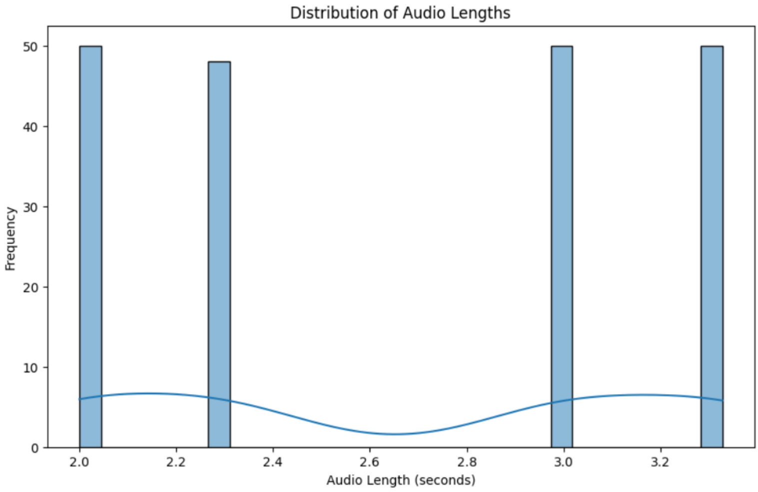

Figure 6. Distribution of Audio Lengths.

The histogram of audio lengths describes the distribution of lengths for fake and real audio samples. This visualization is aimed at identifying potential patterns and differences in audio lengths.

To investigate characteristics specific to fake audio, the Mel-Frequency Cepstral Coefficients (MFCC) method was employed. The heatmap of MFCC features illustrates the energy of the audio signal across various frequencies, depicting how it changes over time. This visualization of features allows for a deeper understanding of the variations within DeepFake audio signals.

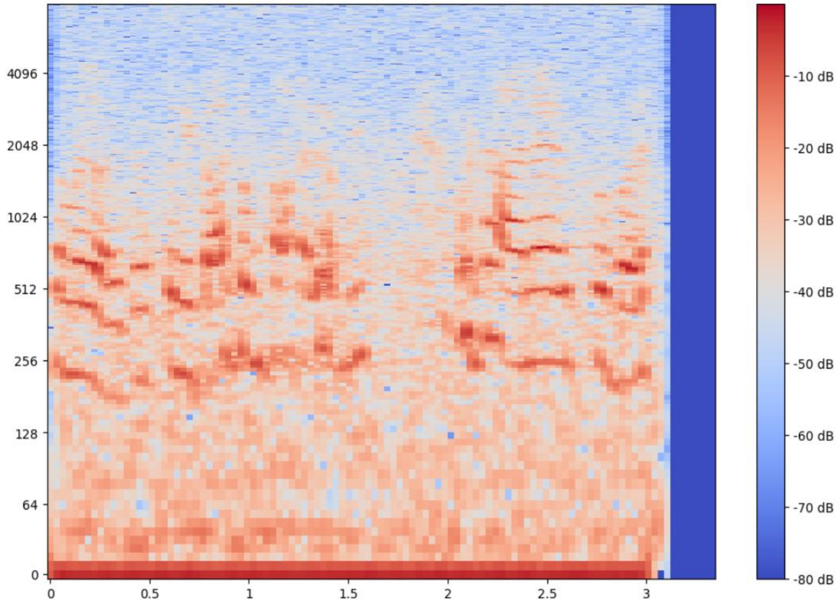

Figure 7. MFCC Features Heatmap.

This heatmap illustrates how the MFCC features of fake audio signals change over time. Analyzing the heatmap revealed differences in the spectral characteristics of fake and real audio signals.

These analysis results provide essential information necessary for training models used to reliably detect DeepFake voices.

To assess the authenticity of audio signals, the Zero-Crossing Rate (ZCR) was employed. ZCR is a metric that describes the frequency of moments when the amplitude of the signal crosses zero over time, aiding in the identification of the audio's spectral features. Comparing the ZCR distributions of fake and real audio signals allows for the identification of differences in their temporal and spectral characteristics.

The histogram of the obtained results shows the distribution of ZCR values among the audio samples, indicating that the features of the samples can be utilized to determine whether they are genuine or fake.

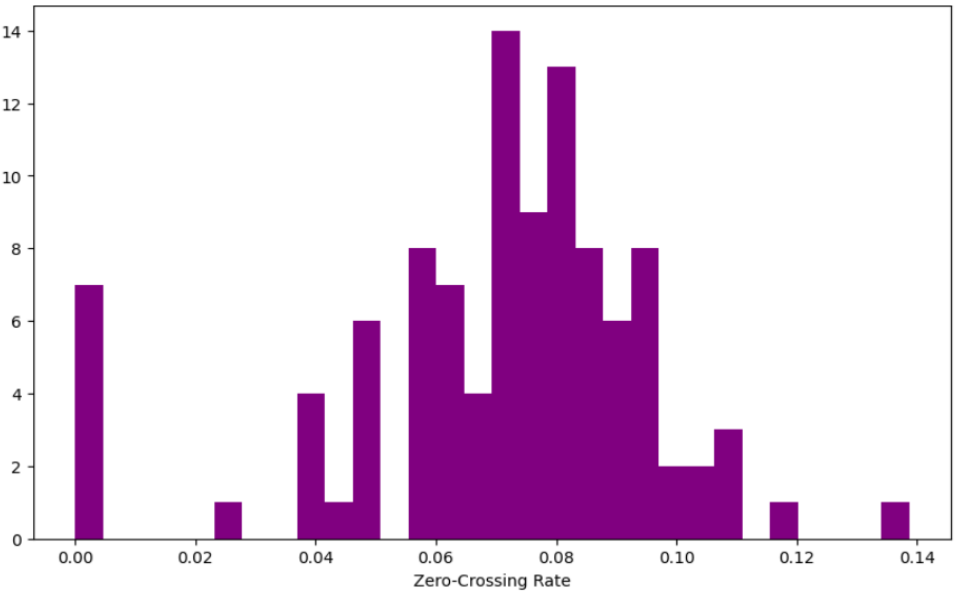

Figure 8. Zero-Crossing Rate in the histogram.

The histogram displays the distribution of ZCR values, enabling the visualization of differences between fake and real audio signals, which aids in determining their authenticity.

The performance metrics for detecting DeepFake voices using LSTM and CNN models are as follows. The results have been compiled in Table 1 below for comparison.

Table 1 – Results of the LSTM and CNN models.

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
LSTM	100.00%	1.00	1.00	1.00
CNN	97.50%	0.98	0.97	0.97

The LSTM model demonstrated excellent results across all metrics, achieving 100% in accuracy, precision, recall, and F1-Score. This confirms the model's ability to accurately and reliably distinguish between fake and real audio.

The CNN model also performed well, but its metrics were slightly lower than those of the LSTM model. The CNN achieved an accuracy of 97.50%, a precision of 0.98, and a recall of 0.97. This indicates that the model may not always correctly identify fake audio samples.

The results show that the LSTM model is significantly more effective for detecting DeepFake voices, while the CNN model, despite its good performance, lags behind LSTM in terms of accuracy.

The primary findings of the study focus on evaluating the effectiveness of the model used for classifying audio files. A key component of the utilized code is the predict_audio function, which loads the audio file and computes its MFCC (Mel

Frequency Cepstral Coefficients) features to make a preliminary assessment with the model.

Once the audio file is loaded, it is processed at a frequency of 16 kHz using the librosa library. The MFCC function is employed to extract the frequency characteristics of the audio signal, utilizing 40 coefficients for this feature extraction. As a result, the obtained MFCC array is averaged, ensuring high accuracy for the model.

The code employs the torch library to convert the MFCC array into a PyTorch tensor, enabling the model to make predictions. The model operates in eval() mode, and predictions are executed within the torch.no_grad() context, eliminating the need to compute gradients during the calculations.

The model's output indicates «Fake» if the prediction equals 1, and «Real» if it equals 0. The result for the audio file tested is presented in the format «The audio is classified as: {result}.» This functional approach allows for the assessment of audio signal quality and their classification, which could be applied in future systems designed for individuals with visual and auditory impairments.

Discussion. The study demonstrates that LSTM outperforms CNN for DeepFake voice detection, achieving perfect scores in Accuracy, Precision, Recall, and F1-Score. This highlights LSTM's strength in handling sequential data, capturing subtle temporal patterns that aid in distinguishing real from fake audio. In contrast, CNN, while accurate (97.50%), showed slightly lower Precision and Recall due to its focus on spatial rather than temporal features, limiting its ability to detect intricate patterns in manipulated audio.

These results underscore LSTM's suitability for DeepFake detection in applications like cybersecurity and media verification. Future work could explore hybrid models combining LSTM and CNN or test other RNN variants to further enhance detection accuracy.

Conclusion. This research utilized various visualization methods to analyze the structure and features of audio data, providing essential information for training models aimed at detecting DeepFake voices. The dataset consisted of fake and real audio files, incorporating the voices of 58 politicians and public figures. While the fake audio samples were artificially generated, the real audio samples were sourced from authentic references. During the data processing phase, parameters such as audio file lengths and MFCC (Mel Frequency Cepstral Coefficients) features were computed.

The analysis of the numerical distribution of fake and real audio samples, particularly through bar charts, enabled the assessment of balance levels within the dataset. Given the varying lengths of audio files over time, analyzing their distribution is crucial for evaluating the authenticity of fake and real audio. A histogram depicted the distribution of lengths for fake and real audio files, aiding in identifying specific length characteristics typical of fake audio.

The MFCC method played a significant role in exploring characteristics specific to fake audio. The heatmap of MFCC features allowed for a deeper understanding of the changes within DeepFake audio signals. Additionally, the Zero-Crossing Rate (ZCR) metric was employed to assess the authenticity of audio signals, helping to identify differences in the temporal and spectral characteristics of fake and real audio signals.

Utilizing LSTM and CNN models, the results of the DeepFake voice detection indicated that the LSTM model achieved an accuracy of 100%, while the CNN model was rated at 97.50% accuracy. The findings demonstrated that the LSTM model is significantly more effective in detecting DeepFake voices, while the CNN model, despite its commendable performance, lagged behind in terms of accuracy.

Overall, the research focused on evaluating the effectiveness of the models used for classifying audio files, ultimately proposing methodologies for reliably detecting DeepFake audio signals. The data obtained can be applied in systems designed for individuals with visual and auditory impairments, enhancing the process of assessing audio signal authenticity and improving information retrieval.

References

Karnouskos S. (2020). Artificial intelligence in digital media: The era of deepfakes. *IEEE Transactions on Technology and Society*, *1*(3), 138-147. DOI:10.1109/TTS.2020.3001312 (in Eng.)

Al-Khazraji S.H., Saleh H.H., KHALID A.I., & MISHKHAL I.A. (2023). Impact of Deepfake Technology on Social Media: Detection, Misinformation and Societal Implications. *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, *23*, 429-441. DOI: 10.55549/epstem.1371792 (in Eng.)

Mullen M. (2022). A new reality: deepfake technology and the world around us. *Mitchell Hamline L. Rev.*, *48*, 210. (in Eng.)

Kumar N., & Kundu A. (2024). Cyber Security Focused Deepfake Detection System Using Big Data. *SN Computer Science*, *5*(6), 752. (in Eng.)

Patel Y., Tanwar S., Bhattacharya P., Gupta R., Alsuwian T., Davidson I. E., & Mazibuko T. F. (2023). An improved dense CNN architecture for deepfake image detection. *IEEE Access*, *11*, 22081-22095. DOI: 10.1109/ACCESS.2023.3251417 (in Eng.)

Al-Dhabi Y., & Zhang S. (2021, August). Deepfake video detection by combining convolutional neural network (cnn) and recurrent neural network (rnn). In *2021 IEEE international conference on computer science, artificial intelligence and electronic engineering (CSAIEE)* (pp. 236-241). IEEE. DOI: 10.1109/CSAIEE54046.2021.9543264 (in Eng.)

Hamza A., Javed A.R.R., Iqbal F., Kryvinska N., Almadhor A.S., Jalil Z., & Borghol R. (2022). Deepfake audio detection via MFCC features using machine learning. *IEEE Access*, *10*, 134018-134028. OI: 10.1109/ACCESS.2022.3231480 (in Eng.)

Al-Adwan A., Alazzam H., Al-Anbaki N., & Alduweib E. (2024). Detection of Deepfake Media Using a Hybrid CNN–RNN Model and Particle Swarm Optimization (PSO) Algorithm. *Computers*, *13*(4), 99. *https://doi.org/10.3390/computers13040099* (in Eng.)

Wang W., Wang Z., Wang G., & Zou Q. (2022, September). Deepfake Video Detection Exploiting Binocular Synchronization. In *International Conference on Artificial Neural Networks* (pp. 101-112). Cham: Springer Nature Switzerland. (in Eng.)

Agarwal S., & Farid H. (2021). Detecting deep-fake videos from aural and oral dynamics. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 981-989). (in Eng.)

Bohacek M., & Farid H. (2024). Lost in Translation: Lip-Sync Deepfake Detection from Audio-Video Mismatch. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4315-4323). (in Eng.)

Saikia P., Dholaria D., Yadav P., Patel V., & Roy M. (2022, July). A hybrid CNN-LSTM model for video deepfake detection by leveraging optical flow features. In *2022 international joint conference on neural networks (IJCNN)* (pp. 1-7). IEEE. DOI: 10.1109/IJCNN55064.2022.9892905 (in Eng.)

Kaddar B., Fezza S. A., Hamidouche W., Akhtar Z., & Hadid A. (2021, December). HCiT: Deepfake video detection using a hybrid model of CNN features and vision transformer. In *2021 International Conference on Visual Communications and Image Processing (VCIP)* (pp. 1-5). IEEE. DOI: 10.1109/VCIP53242.2021.9675402 (in Eng.)

Cho B., Le B.M., Kim J., Woo S., Tariq S., Abuadbba A., & Moore K. (2023, October). Towards understanding of deepfake videos in the wild. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management* (pp. 4530-4537). *https://doi.org/10.1145/3583780.3614729* (in Eng.)

Cavia B., Horwitz E., Reiss T., & Hoshen Y. (2024). Real-Time Deepfake Detection in the Real-World. *arXiv preprint arXiv:2406.09398*. *https://doi.org/10.48550/arXiv.2406.09398* (in Eng.)

Baxevanakis S., Kordopatis-Zilos G., Galopoulos P., Apostolidis L., Levacher K., Baris Schlicht I., ... & Papadopoulos S. (2022, June). The mever deepfake detection service: Lessons learnt from developing and deploying in the wild. In *Proceedings of the 1st International Workshop on Multimedia AI against Disinformation* (pp. 59-68). *https://doi.org/10.1145/3512732.3533587* (in Eng.)

Ahmed S.R., Sonuç E., Ahmed M.R., & Duru A.D. (2022, June). Analysis survey on deepfake detection and recognition with convolutional neural networks. In *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)* (pp. 1-7). IEEE. DOI: 10.1109/HORA55278.2022.9799858 (in Eng.)

Mcuba M., Singh A., Ikuesan R.A., & Venter H. (2023). The effect of deep learning methods on deepfake audio detection for digital investigation. *Procedia Computer Science*, *219*, 211-219. *https://doi.org/10.1016/j.procs.2023.01.283* (in Eng.)

Li M., Ahmadiadli Y., & Zhang X.P. (2022, October). A comparative study on physical and perceptual features for deepfake audio detection. In *Proceedings of the 1st International Workshop on Deepfake Detection for Audio Multimedia* (pp. 35-41). *https://doi.org/10.1145/3552466.3556523* (in Eng.)

Alshehri A., Almalki D., Alharbi E., & Albaradei S. (2024). Audio Deep Fake Detection with Sonic Sleuth Model. *Computers*, *13*(10), 256. *https://doi.org/10.3390/computers13100256* (in Eng.)

CONTENTS

INFORMATION AND COMMUNICATION TECHNOLOGIES

МАЗМҰНЫ

АҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР

СОДЕРЖАНИЕ

ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

Publication Ethics and Publication Malpracticein the journals of the National Academy of Sciences of the Republic of Kazakhstan

For information on Ethics in publishing and Ethicalguidelines 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 otherlanguage, 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
ISSN2518-1726 (Online),
ISSN 1991-346X (Print)

Директор отдела издания научных журналов НАН РК *А. Ботанқызы*
Редакторы: *Д.С. Аленов, Ж.Ш.Әден*
Верстка на компьютере *Г.Д.Жадыранова*

Подписано в печать 20.03.2025.
Формат 60х881/8. Бумага офсетная. Печать –ризограф.
20,0 п.л. Заказ 1.

Национальная академия наук РК
050010, Алматы, ул. Шевченко, 28, т. 272-13-19