RESEARCH ARTICLE OPEN ACCESS

Advancements in Speech Recognition Models for Speech Therapy Applications Across Various Human Attributes

Dr. Akanksha Budholiya
Assistant Professor
Sadabai Raisoni Women's college
akansha.budholilya@raisoni.net

Devyani Moundekar,
Sadabai Raisoni Women's college
devyani.moundekar.bca@srwc.raisoni.net

Aditi Chole
Sadabai Raisoni Women's college
aditi.chole.bca@srwc.raisoni.net

ABSTRACT

Speech is the most basic and widely used means of human communication. Speech has the potential to be a highly effective technique for human-computer connection. This paper provides an in-depth review of the methodologies established at each stage of the process, as well as a synopsis of the basic technological perspective and an assessment of the fundamental advances in voice recognition. This study uses deep learning models and artificial intelligence to evaluate the effectiveness of automatic speech recognition (ASR) systems. Large language models (LLMs) are used to personalize and adjust to the challenges and progress of users. Reinforcement learning (RL) is used to construct interactive exercises that improve articulation or improve vocabulary by dynamically tailoring speech tasks for the user. Speech recognition technology has enormous potential to improve human-computer interaction in the future. Expanding reinforcement learning can result in more complex therapeutic tasks that support language development and speech therapy. These developments will pave the way for easy access, communication, and healthcare solutions across a range of demographics. These recommendations are useful for future research into the development of effective, personalized human-computer interaction systems, particularly for speakers with unique phonetic requirements.

Keywords - Speech Recognition, Automatic Speech Recognition (ASR), Human-Computer Interaction (HCI), Deep Learning, Artificial Intelligence (AI), Large Language Models (LLMs), Reinforcement Learning (RL), Speech Therapy

I.INTRODUCTION -

Speech and language impairments affect millions of people worldwide, particularly youngsters, limiting their ability to communicate effectively and participate in social and academic activities. Early intervention with speech therapy is critical for enhancing speech development and avoiding long-term communication problems. Traditional speech therapy, on the other hand, suffers a number of obstacles, including a shortage of speech-language pathologists (SLPs), high treatment costs, and limited availability for people living in remote or underprivileged communities. As a result, there is an increased interest in using technology-assisted speech therapy solutions to improve accessibility, participation, and treatment outcomes.

Mobile applications, artificial intelligence (AI), serious games, and online speech therapy systems have all made significant contributions to how speech therapy is delivered today. Mobile applications have become a vital tool in speech therapy, providing interactive and engaging exercises that allow children to practice their speech abilities outside of clinical settings. These programs offer a flexible, cost-effective, and user-friendly alternative to traditional therapy, allowing children with speech impairments to obtain constant practice and reinforcement. However, with the increasing proliferation of available mobile applications, speech therapists and parents encounter difficulties in picking the most effective and dependable resources. Evaluating these applications is critical to ensuring they meet clinical demands and produce real gains in speech results.

Artificial intelligence (AI) has emerged as a game changer in speech therapy, allowing for automated speech assessment and individualized intervention. AI-powered speech therapy systems assess pronunciation, fluency, and articulation errors through speech recognition, anomaly detection, and feature extraction techniques.

Machine learning (ML) models, including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Hidden Markov Models (HMMs), have been used to analyse speech patterns and deliver real-time feedback. These systems have the potential to reduce speech therapists' labour while guaranteeing that children receive data-driven, adaptive speech therapies suited to their specific requirements. However, issues such as heterogeneity in kid speech, dataset limits, and accuracy in phoneme detection remain topics of concern that necessitate additional research and improvement.

Gamification and serious games, in addition to mobile apps and AI-driven assessments, have been widely implemented into speech therapy to boost motivation and participation. Interactive, voice-controlled games allow youngsters to improve their speaking skills through enjoyable and engaging activities such as controlling avatars, completing tasks, and solving puzzles with voice commands. These games increase articulation, speech intensity, and fluency by incorporating real-time feedback mechanisms that direct users to the optimal pronunciation. Research has demonstrated that game-based speech therapy improves therapy adherence and encourages active learning, making it an excellent treatment option for children with speech impairments. However, ensuring that these games are clinically evaluated and appropriately customized to various speech deficits is an ongoing challenge.

This literature review will look at contemporary breakthroughs in technology-assisted speech therapy, with an emphasis on mobile apps, AI-based speech assessment, serious games, and online therapeutic platforms. This study aims to identify gaps in current knowledge, assess the usefulness of emerging technologies, and highlight potential for future improvement in speech therapy interventions by examining diverse studies and research findings.

II.METHODOLOGY

This literature review takes a systematic and structured approach to analysing current research on technology-assisted speech therapy. The research focuses on AI, machine learning, mobile interventions, serious games, ontology-based expert systems, and virtual speech therapists. The purpose is to identify common approaches, important technical advancements, and research gaps in order to gain a thorough picture of speech therapy's growth and future goals. To discover relevant studies, a computerized search was undertaken in IEEE Xplore, PubMed, Scopus, ScienceDirect, and Google Scholar. The search covered peer-reviewed journal articles, conference papers, and systematic reviews published between 1980 and 2024.

The keywords and Boolean operators utilized were:

"AI in speech therapy" as well as "machine learning for speech disorders"
"Speech therapy mobile applications" as well as "gamification in speech rehabilitation"
"Virtual speech therapists" as well as "ontology-based therapy systems"
"Deep learning for speech recognition" as well as "speech disorder assessment"
Filters were used to highlight empirical studies, English-language journals.

Speech and language disorders were investigated in terms of their influence on communication. AI-driven interventions, voice recognition, and mobile therapeutic solutions have all been investigated.
Provided empirical data or case studies to assess the efficacy of technology-assisted speech therapy, Were published in credible journals, conferences, or systematic reviews. The focus was exclusively on theoretical models, with no practical application. It did not particularly mention speech therapy applications and
Were unavailable in English or lacked methodological clarity.

The Screening and Selection Process was a two-stage screening mechanism:

To exclude unnecessary research, screen the title and abstract.
Full-text review to ensure research rigor and alignment with aims.

Data Extraction and Synthesis

Data were divided into: Common tactics include artificial intelligence models, speech processing techniques, and therapeutic approaches. Speech treatment technology breakthroughs include AI-driven speech recognition, gamified interventions, and teletherapy systems. Challenges and research gaps include speech recognition accuracy, dataset limits, and accessibility issues.

Commonly used speech processing techniques

Mel-Frequency Cepstral Coefficients (MFCCs), Hidden Markov Models (HMMs), and Deep Neural Networks (DNNs). AI and machine learning approaches include Convolutional Neural Networks (CNNs), Long Short-Term Memory Networks (LSTMs), and Support Vector Machines. Gamification and mobile therapy include voice-controlled interactive games and serious game-based interventions. Ontology-Based Expert Systems: AI-powered therapy recommendation models. Advancements in Speech Therapy Technology

III.RELATED WORK

Numerous research have investigated how technology-assisted speech therapy can increase accessibility and efficacy for people with speech and language difficulties. Traditional speech therapy, which is frequently offered in clinical settings, faces obstacles such as a shortage of speech-language pathologists (SLPs), high costs, and geographical constraints. To address these difficulties, academics have looked into artificial intelligence (AI), machine learning (ML), mobile applications, serious games, ontology-based expert systems, and virtual speech therapists as alternatives or complementary approaches.

Several research have proved the efficacy of AI-powered voice recognition models in detecting pronunciation problems and providing immediate feedback. Machine learning techniques, such as Convolutional Neural Networks (CNNs), Long Short-Term Memory Networks (LSTMs), and Hidden Markov Models (HMMs), have been widely employed to increase voice recognition accuracy for people with speech problems. Furthermore, ontology-based expert systems have been developed to automate therapy planning and intervention procedures, allowing patients to get tailored treatment based on their speech patterns and developmental history.

Gamification has emerged as a potential method for increasing engagement and motivation in speech therapy. Studies have indicated that serious games and interactive voice-controlled programs promote therapeutic adherence, particularly among children. Similarly, virtual speech therapists and teletherapy platforms have made therapy more accessible, especially for those living in rural places.

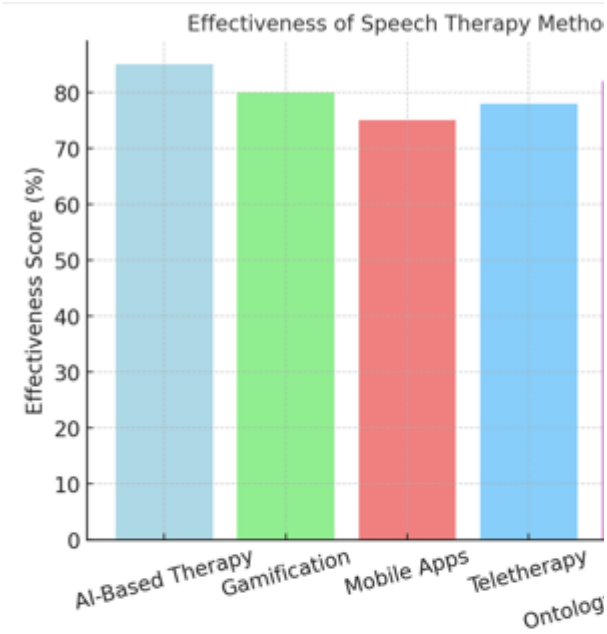
Despite these advances, significant research gaps persist. Many AI-based voice recognition systems continue to struggle with substantial variability in children's speech, resulting in uneven accuracy. Furthermore, defined evaluation tools for technology-assisted speech therapy are scarce, making it impossible to compare efficacy among therapies. Future research should concentrate on enhancing AI models, incorporating multimodal data (speech, facial expressions, and gestures), and assuring the cost and accessibility of digital speech therapy solutions

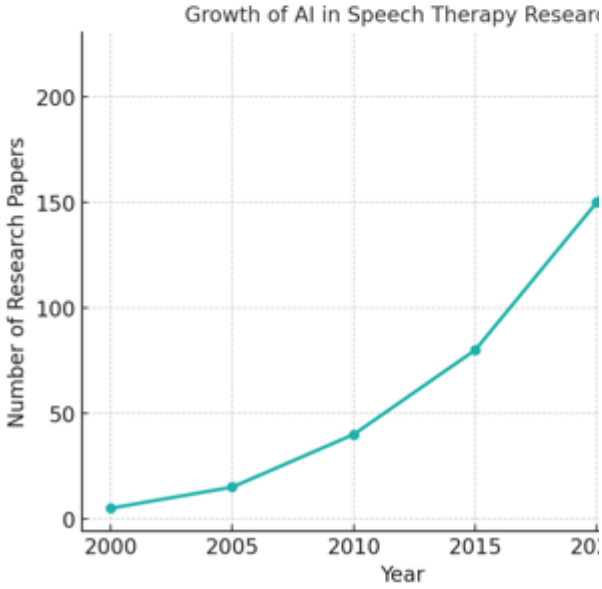
IV.DISCUSSION AND ANALYSIS

The use of technology-assisted solutions in speech therapy has drastically altered intervention tactics for people with speech and language impairments. This research review focuses on significant advances in artificial intelligence (AI), machine learning, mobile apps, gamification, ontology-based expert systems, and virtual speech therapists. These innovations seek to improve therapeutic accessibility, engagement, and effectiveness while addressing traditional speech therapy difficulties such as restricted access to specialists, high expenses, and the requirement for ongoing practice outside of clinical sessions.

One of the most notable advances in this field is the creation of AI-powered voice recognition systems that can analyse pronunciation problems and provide real-time feedback. Machine learning models, such as CNNs, LSTMs, and HMMs, have been widely utilized to improve automated voice recognition. Furthermore, gamification in speech therapy has enhanced participation by combining interactive and reward-based learning, making speech exercises more enticing, particularly to children. Another significant innovation is the ontology-based expert system, which automates therapy planning and intervention tactics by building treatments around defined medical classifications and patient profiles. Teletherapy and virtual speech therapists have also increased accessibility, allowing people in rural places to get speech therapy via online platforms.

Despite these developments, numerous limits persist. AI-powered speech therapy offers automated and scalable solutions; however, it frequently struggles to recognize speech variability in children and persons with severe speech problems. While mobile interventions allow for more flexible and self-paced therapy, they still require therapist monitoring to ensure accuracy and individualized changes. Similarly, gamified speech treatment increases motivation and engagement but lacks standardized clinical validation to assess long-term efficacy

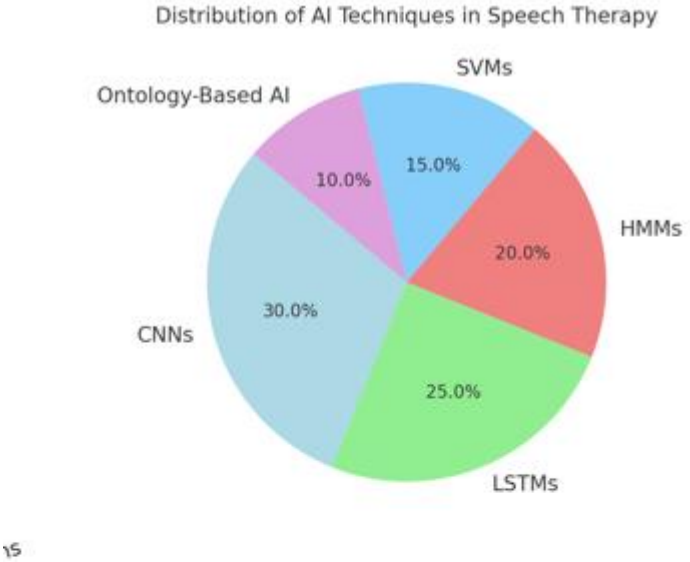


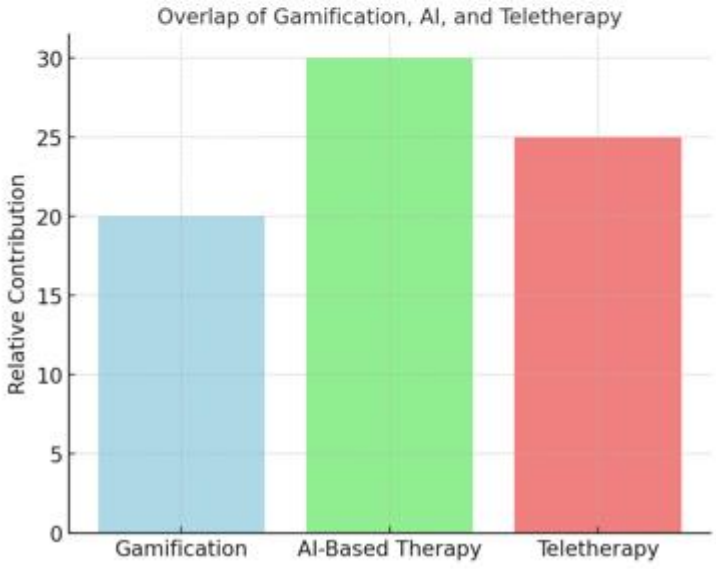


Although ontology-based expert systems offer structured therapy plans, they may lack adaptation for specific or complex speech problems, restricting their adaptability to a wide range of patient needs.

While AI and machine learning have made therapy more accessible and personalized, significant research gaps remain. Speech recognition models continue to struggle with substantial variability in child speech, resulting in variable accuracy rates. Many digital therapy tools are not validated through clinical trials, making it difficult to establish their usefulness in real-world applications. Furthermore, the integration of multimodal data, such as speech, facial expressions, and gestures, has received little attention, despite the fact that incorporating these aspects could considerably improve speech disorder assessments.

To close these gaps, future research should concentrate on improving AI accuracy, increasing datasets for better voice recognition, and combining Virtual Reality (VR) and Augmented Reality (AR) for immersive speech therapy experiences. There is also an increasing demand for standardized evaluation metrics to compare the efficacy of various technology-assisted programs. Finally, technology-driven speech therapy has the potential to make speech rehabilitation more accessible, inexpensive, and engaging, resulting in improved treatment outcomes for people with speech problems.





V.FUTURE SCOPE

The future of technology-assisted speech therapy is in improving AI-driven voice recognition to enhance accuracy in child speech assessment, as well as increasing machine learning models for tailored therapy programs. Future research should concentrate on integrating Virtual Reality (VR) and Augmented Reality (AR) for immersive treatment experiences, developing gamified therapies, and guaranteeing clinical validation of mobile-based therapy applications. Additionally, adopting common evaluation metrics and multilingual AI models would improve accessibility for a wide range of populations. To ensure comprehensive and effective speech therapy options for people all around the world, ethical factors such as data privacy and pricing must be considered as well.

VI.CONCLUSION

Technology-assisted speech therapy has transformed interventions for people with speech and language difficulties. This paper focuses on key breakthroughs in AI-driven speech recognition, machine learning-based speech assessment, mobile apps, gamification, ontology-based expert systems, and virtual speech therapists. These approaches have increased therapy accessibility, engagement, and automation, making speech rehabilitation more successful and broadly available. AI-powered models, such as CNNs, LSTMs, and HMMs, have improved automated speech analysis, while gamified therapy and teletherapy platforms have raised patient incentive and compliance.

Despite these advances, significant research gaps persist. AI voice recognition issues with children's speech fluctuation, resulting in variable accuracy. Furthermore, many digital therapeutic tools lack real-world clinical validation, making it impossible to determine their long-term effectiveness. The lack of established evaluation metrics further reduces the comparison of different approaches.

The broader implications of this study show the ability of AI, mobile apps, and virtual speech therapists to reduce accessibility gaps in healthcare and education. Future research should concentrate on improving AI accuracy, incorporating VR/AR into immersive therapeutic experiences, and creating tailored speech therapy models. By tackling these issues, technology-driven speech therapy can continue to improve, delivering more effective, inclusive, and cost-effective options for people with speech impairments.

REFERENCES

1.Carolyne Alphonsus Tommy, Jacey-Lynn Minoi "Speech Therapy Mobile Application for

Speech and Language Impairment Children"2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)

2.Jothi K R, Siva Raju S S, Priyanka Jayant Yawalkar "AI based Speech Language Therapy using Speech

Quality Parameters for Aphasia Person:

A Comprehensive Review “Fourth International Conference on Electronics, Communication and Aerospace Technology (ICECA-2020)

3.Nahid Nasiri, Shervin Shirmohammadi, Ammar Rashed "A Serious Game for Children with Speech Disorders and Hearing Problems"©2017 IEEE

4.J. J. L. Maas1*, N. M. De Vries1

, B. R. Bloem1 and J. G. Kalf2"PERsonalized SPEeCh Therapy for actIVE

conversation in Parkinson’s disease

(randomized controlled trial)"Maas et al. Trials (2022) 23:274

5.Geertruida Aline Attwell, Kwabena Ebo Bennin * and Bedir Tekinerdogan "A Systematic Review of Online Speech Therapy Systems for

Intervention in Childhood Speech Communication Disorders “Sensors 2022, 22, 9713

6. Mostafa Shahin, Usman Zafar, and Beena Ahmed “The Automatic Detection of Speech Disorders in Children: Challenges, Opportunities and Preliminary Results” DOI 10.1109/JSTSP.2019.2959393, IEEE Journal of Selected Topics in Signal Processing 2019

7.Tihomir Orehovački, Dijana Plantak Vukovac, Zlatko Stapić, Tatjana Novosel-Herceg “Features and Quality of a Mobile Application Employed in a Speech-Language Therapy” Springer International Publishing AG 2017

8.Mirela Danubianu, Stefan Gheorghe Pentiuc, Tiberiu Socaciu “Towards the Optimized Personalized Therapy of Speech Disorders by Data Mining Techniques” 2009 Fourth International Multi-Conference on Computing in the Global Information Technology

9.W.R. Rodríguez, C. Vaquero, O. Saz, E. Lleida "Speech Technology Applied to Children with Speech Disorders" Springer-Verlag Berlin Heidelberg 2008

10. Yi-Ping Phoebe Chen, Caddi Johnson, Pooja Lalbakhsh, Terry Caelli, Guang Deng, David Tay, Shane Erickson, Philip Broadbridged, Amr El Refaie, Wendy Doubee, Meg E. Morris "Systematic review of virtual speech therapists for speech disorders" 2015 Elsevier Ltd
11. Davide Mulfari, Gabriele Meoni, Marco Marini , Luca Fanucci "Machine learning assistive application for users with speech disorders" Applied Soft Computing Journal 103, Elsevier 2021
12. Vladimir E. Robles-Bykbaev, Wilson Guam´an-Murillo,Diego Quisi-Peralta, Mart´ın L´opez-Nores, Jos´e J. Pazos-Arias and Jorge Garc´ıa-Duque "An ontology-based expert system to generate therapy plans for children with disabilities andcommunication disorders" 2016 IEEE
13. Maitraye Das, Abir Saha "An Automated Speech-Language Therapy Tool with Interactive Virtual Agent and Peer-to-Peer Feedback" 4th International Conference on Advances in Electrical Engineering, IEEE, 2017
14. Fanny Chuchuca-M´endez, Vladimir Robles-Bykbaev, Pablo Vanegas-Peralta,Juan Lucero-Salda˜na, Mart´ın L´opez-Nores and Jos´e Pazos-Arias "An educative environment based on ontologies and e-learning for training on design of speech-language therapy plans for children with disabilities and communication disorders" 2016 IEEE
15. Mirela Danubianu, Stefan-Gheorghe Pentiuc, Ovidiu Andrei Schipor, Marian Nestor, Ioan Ungureanu "Distributed Intelligent System for Personalized Therapy of Speech Disorders" Third International Multi-Conference on Computing in the Global Information Technology, 2008 IEEE