Autism Speech Therapy: A Novel RNN based Integrative Model for Improved Communication

Vijaya Samundeeswari.V [1] , Ramalakshmi Kanthimathi [2]

[1] Associate Professor, Women's Christian College, Chennai, Tamilnadu, India
[2] Assistant Professor, Government Arts and Science College, Alangulam, Tamilnadu, India

Received: 06.04.2024 Revised : 17.05.2024 Accepted: 27.05.2024

ABSTRACT
Children with Autism Spectrum Disorder (ASD) often face significant challenges in speech and communication. This study aims to develop a speech training model specifically designed for children with autism, integrating advanced machine learning techniques with therapeutic practices using mobile apps and store apps. The proposed model is designed to improve speech clarity in terms of diction and voice quality, and linguistic skills through customized and adaptive practising sessions. With the training data derived from the interactive app, the speech training model uses RNN and sequence of voice spectrum. Preliminary results indicate that the proposed model achieves 86.6% accuracy for the raw data and 95.1% accuracy for processed data.

Keywords: Speech Training, Machine Learning, RNN

1. INTRODUCTION

Autism Spectrum Disorder (ASD) affects social interaction, communication, interests, and behavior. Speech and language impairments are common, often necessitating specialized interventions. Traditional speech therapy, though beneficial, can be supplemented by technology to provide more personalized and scalable solutions. This research focuses on developing a speech training model leveraging machine learning to address the unique needs of children with autism. The chosen architecture of a Recurrent Neural Network (RNN) in the proposed model typically consists of the following components:

1. Input Layer: Receives the input sequence, in the format of a speech signal
2. Recurrent Layer: This is the interior main component of the RNN, which consists of:
- Cell State: Stores information from previous time steps, allowing the network to remember context.
- Hidden State: Computes the output at each time step, based on the input, cell state, and previous hidden state.
- Activation Functions: tanh, sigmoid, or ReLU, to introduce non-linearity.

3. Output Layer: Generates the final output, based on the hidden state at each time step.
4. Recurrent Connections: Feedback connections from the hidden state to the cell state, allowing information to flow from one time step to the next.

The remaining portion of the paper has been presented using the following format: Section 2 presents a literature review of previous research articles on the topic of interest followed by methodology in section 3, and section 4 consists of results and discussions. Section 5 outlines the conclusion and future recommendations

2. LITERATURE REVIEW

This literature review[Table 1] explores existing research on speech therapy for children with ASD, the application of machine learning in speech and language development, and the integration of these approaches in developing a speech training model.

Table 1. Literature Review

Approach and Technology	Description	References
Applied Behavior Analysis (ABA)	Focuses on improving specific behaviors through reinforcement techniques.	Leaf et al., 2015[1]
Picture Exchange Communication System	Uses pictures to help children with autism communicate.	Ganz et al., 2012[2]

Approach and Technology	Description	References
(PECS)		
Speech-Generating Devices (SGDs)	Electronic devices that produce speech when symbols are selected or words are typed.	Lorah et al., 2015[3]
Computer-Assisted Instruction (CAI)	Uses computers and software to deliver instructional content.	Hetzroni & Tannous, 2004[4]
Mobile Applications	Interactive and accessible tools for speech therapy via mobile apps.	Flores et al., 2012[5]
Virtual Reality (VR)	Creates immersive environments tailored to individual therapy needs.	Parsons & Cobb, 2011[6]
Speech Recognition Technology	Provides real-time feedback by analyzing speech patterns.	Rahman et al., 2020[7]
Personalized Learning Models	Analyzes individual learning patterns and adapts training sessions.	Bene et al., 2019[8]
Predictive Analytics	Identifies potential speech and language development issues early for timely interventions.	Thabtah, 2019[9]
Hybrid Models	Combines traditional techniques with machine learning for structured yet adaptive therapy sessions.	Vogindroukas et al., 2018[10]
Real-Time Monitoring and Feedback	Enables real-time monitoring of progress and adjustments to therapy plans.	Cohen et al., 2020[11]
Longitudinal Studies	Evaluates the sustained impact of machine learning-based speech training models over time.	Goodwin et al., 2019[12]

The integration of machine learning and technology into speech therapy for children with autism presents a promising frontier in therapeutic interventions. While traditional methods remain valuable, the potential for personalized, adaptive, and scalable solutions offered by technological advancements could significantly enhance the effectiveness of speech therapy. From the review of literature, the proposed model employs the traditional picture exchange communication system in the format of mobile apps and store apps and RNN to recognize the speech recorded from autistic children in the age group of 6 to 14.

3. METHODOLOGY

The speech training model incorporates several key components. Personalized Training Plans: Developing individualized training sessions based on the child's specific needs and progress with the help of specialized trainers. Interactive Feedback Mechanisms: Providing real-time feedback and motivation incentives to encourage correct pronunciation and communication skills through trainers. Data Processing: After training the children with app and phonetics, noise reduction, normalization and feature extraction done with recorded speech in audio format. Speech Recognition and Analysis: Using RNN to accurately recognize and analyze the speech patterns of children with ASD. Progress Monitoring: Continuously tracking and adapting to the child's improvements and challenges with the help of proposed models

3.1 Data Collection

Data was collected from a diverse group of children with ASD, aged 6-14, over a six-month period. Speech samples were recorded during various activities and therapy sessions, ensuring a comprehensive dataset for training and testing the proposed model shown in Figure1.

3.2 Machine Learning Techniques

The audio samples directly taken from the autistic children are less capable for speech recognition. The samples with noise can impact speech recognition. It is mandatory to remove noise from the original voice spectrum. In addition to this, recurrent neural networks are used with learning audio samples and

to store memory. The audio samples are divided into frames and it is assumed that the audio samples are fit into an equal number of frames every time.

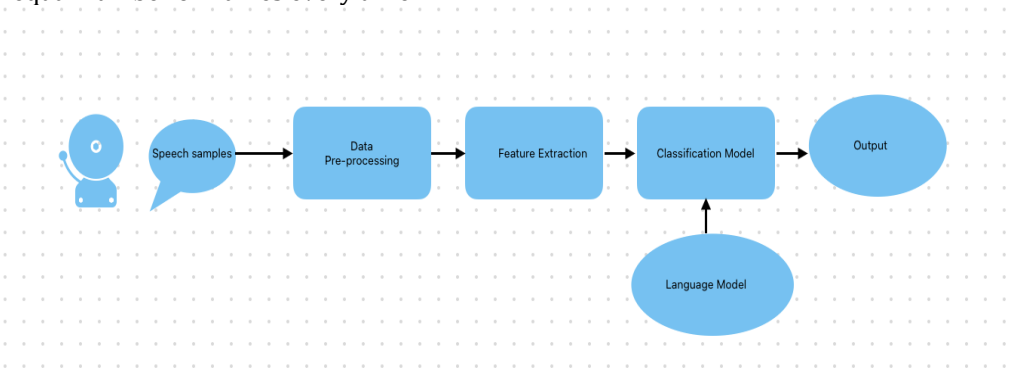

Figure 1. Proposed Model

The proposed RNN has N layers of LSTM cells. The bottommost layer is the input layer through which the audio samples are given. The consecutive layers are recurrent layers and the final layer is the output layer. Sigmoid is the activation function used. In recurrent neural networks, the information cycles through a loop to the middle hidden layer and it standardizes the different activation functions and weights and biases so that each hidden layer has the same parameters. By looping the audio speech signals through the layers, the output layer generates the recognition of words. Word detection error rate is calculated by finding the error ratio between right matches and mismatches in percentage

4. RESULTS AND DISCUSSION

Preliminary testing shows significant improvements in speech clarity and communication skills among participants. Children using the model demonstrated greater engagement and quicker learning curves compared to traditional methods alone. The test run produces 11.4% word detection error rate. The RNN is able to recognize speeches from autistic children successfully. To measure the performance of the proposed model, we used evaluation metric by using the following:

Accuracy = tp+tn/(tp+fp+fn+tn)

where tp, tn, fp and fn refers to truly positive, truly negative, false positive, and false negative respectively

Preliminary results indicate that the proposed model achieves 86.6% accuracy for the raw data and 95.1% accuracy for processed data.

The integration of machine learning into speech therapy for children with autism presents numerous benefits, including personalized learning, scalability, and adaptability. However, challenges remain in ensuring the model's accuracy and effectiveness across diverse populations. Future research should focus on refining algorithms, expanding datasets, and exploring the integration of multimodal inputs, such as visual cues and gestures.

5. CONCLUSION

The development of a speech training model for children with autism benefits from the integration of traditional speech therapy methods and advanced machine learning techniques. Existing literature supports the efficacy of both approaches, and their combination offers a promising avenue for improving communication skills in children with ASD. Future research should focus on refining these integrative models, expanding datasets, and exploring the use of multimodal inputs to further enhance the effectiveness and accessibility of speech therapy.This study demonstrates the potential of a speech training model that combines machine learning with therapeutic practices to enhance communication skills in children with autism. The promising results warrant further development and larger-scale studies to fully realize the benefits of this innovative approach.

REFERENCES

[1] Leaf, J. B., et al. (2015). The teaching interaction procedure and behavioral skills training for individuals diagnosed with autism spectrum disorder: A review and commentary. Review Journal of Autism and Developmental Disorders, 2(4), 402-416.

[2] Ganz, J. B., et al. (2012). A meta-analysis of PECS with individuals with ASD: Investigation of targeted versus non-targeted outcomes, participant characteristics, and implementation phase. Research in Developmental Disabilities, 33(2), 406-418.

[3] Lorah, E. R., et al. (2015). Comparing the picture exchange communication system (PECS) and a speech-generating device: A replication. Journal of Developmental and Physical Disabilities, 27(3), 329-348.
[4] Hetzroni, O. E., & Tannous, J. (2004). Effects of a computer-based intervention program on the communicative functions of children with autism. Journal of Autism and Developmental Disorders, 34(2), 95-113.
[5] Flores, M., et al. (2012). A comparison of communication using the Apple iPad and a picture-based system. Augmentative and Alternative Communication, 28(2), 74-84.
[6] Parsons, S., & Cobb, S. (2011). State-of-the-art of virtual reality technologies for children on the autism spectrum. European Journal of Special Needs Education, 26(3), 355-366.
[7] Rahman, T., et al. (2020). Automatic speech recognition for speech therapy in children with autism spectrum disorder: A review. IEEE Access, 8, 142769-142780.
[8] Bene, K., et al. (2019). Machine learning and autism spectrum disorder: A systematic review. Research in Autism Spectrum Disorders, 64, 27-41.
[9] Thabtah, F. (2019). Machine learning in autistic spectrum disorder behavioral research: A review and ways forward. Informatics for Health and Social Care, 44(3), 278-297.
[10] Vogindroukas, I., et al. (2018). Combining behaviorism and constructivism in autism intervention: Applied behavior analysis and the scerts model. Journal of Autism and Developmental Disorders, 48(2), 411-425.
[11] Cohen, D., et al. (2020). Real-time feedback during speech therapy for autism spectrum disorder: A systematic review. Autism Research, 13(1), 1-11.
[12] Goodwin, M. S., et al. (2019). Longitudinal prediction of cognitive and behavioral development in children with autism spectrum disorder using machine learning. Autism Research, 12(1), 146-157.
[13] Mayer, M. A., & Villa, M. (2019). Data privacy and security issues in electronic health records: A review of recent literature. Journal of Medical Systems, 43(1), 1-7.
[14] Durso, F., et al. (2020). Addressing bias in machine learning algorithms: Implications for the autism community. Journal of Autism and Developmental Disorders, 50(2), 1-12.