RESEARCH ARTICLE **SPECIAL ISSUE**

International Research Journal on Advanced Science Hub
2582-4376
Vol. 05, Issue 05S May
www.rspsciencehub.com

Check for updates

RSP Science Hub

http://dx.doi.org/10.47392/irjash.2023.S037

International Conference on intelligent COMPUting TEchnologies and Research (i-COMPUTER) 2023

Enhancing Transmission of Voice in Real-Time Applications

Puli Venkatesh Kumar [1], Shunmugathammal M[2], Puli Nitish Kumar [3], Dharaneesh V [1], Kanagasabapathi M G [1]

[1]Department of Electronics & Communication Engineering, SRM Institute of Science and Technology Chennai, Tamil Nadu, India.

[2]Professor, Department of Electronics & Communication Engineering, SRM Institute of Science and Technology Chennai, Tamil Nadu, India

[3]Department of computer Science Engineering - Data Science, St.Ann's College of Engineering and Technology, Andhra Pradesh, India

Article History

Received: 28 February 2023
Accepted: 18 March 2023
Published: 28 May 2023

***Keywords**:*

Text-to-Speech (TTS) Speech Recognition Pyttsx3 Tkinter.

Abstract

In today's telecommunication world sharing the data becomes very easy. It is a bit-complicated in converting the text documents to voice assistance even proposed a lot of resources. Giving the correct information to the right person in the right way is essential on both a personal and professional level. Numerous applications have developed with the purpose of enabling two individuals to communicate instantly. The major objective of this effort is to address the issues that dysarthria, business meetings, and regular travelers face. To solve this issue, proposing a gadget that will aid in the translation of written language into speech. The majority of these applications include, language translation, signal conversion from text to synthetic voice, and articulators. In this project, proposing the development in a wide range of strategies and algorithms needed to make text to speech a reality (TTS).

1. Introduction

Cross-lingual TTS necessitates more study, particularly when creating speech in commonly spoken languages. Creating TTS systems that can learn to produce speech in new languages with minimal data is one example, as is integrating different TTS systems to create a multilingual TTS solution. Existing System requires internet connectivity at all times for any type of application. The information submit into the programme must be saved in a database. It is difficult to collect audio recordings of every possible word said in every possible combination of emotions, prosody, stress, and so on, the final speech lacks naturalness and feeling.

The pyttsx3 module, for example, allows you to convert text to voice on Windows using the Microsoft Speech API (SAPI), or on other platforms using the eSpeak or NVDA TTS engines. a variety of voices and alter the synthesized speech's tone, loudness, and tempo could be changed by the user. At this point; selecting the python to import the predefined modules in application. It has to be in the real human conservation style than we are communicating to a machine. Developing a friendly interface, graphical user interface (GUI) to make a user in approachable way to use the application with limited set of commands. The proposing model focuses mostly on the TTS system. Text-to-speech synthesis (TTS) is the automatic conversion of a text into voice that sounds as close to a native speaker of the language reading the text, Known as a text-to-speech synthesizer. This technology enables the computer to talk to you (TTS). The TTS engine, a computer programme, analyses the text after preprocessing it and uses mathematical models to synthesis the voice when the text is supplied into the system.

Audio-formatted sound data is often produced as the TTS engine's output.

1.1. Related Work

Mohammad Soleymanpour (Soleymanpour et al.) et al. completed their work on Dysarthric Speech in 2022. Primary work focuses on the two variables Synthetic Dysarthric Speech and Pause Insertion. The work's superiority is improved by adding a dysarthria severity level coefficient and a pause insertion model to a neural multi-talker TTS to synthesize dysarthric speech for varying severity levels. The work's shortcoming is that it cannot be accessible without connection to the internet. Although the ensuing voice output did not appear to be human interaction, it demonstrated WER progress. In 2014, T.Rubesh Kumar and C.Purnima (Kumar and Purnima) produced the project Blind Users Assistive System For Product Detection With Voice Outcomes. The localisation of algorithms is their main priority. The proposed unique text localization technique is based on models of stroke orientation and edge distributions, and word recognition is accomplished using OCR. According to the arete, it can only analyses digital text.

M. Shunmugathammal (Shunmugathammal, Sundari, and Prakash) et al. Completed their study on Caption Generation System Using LSTMS and WEB API in 2022, with a major focus on the approach Long short term memory (LSTM). Using a huge dataset and doing further hyper-parameter tuning are frequent components of the strategy. There was no mention of the amount of data stored in the work. Anusha Bhargava (Bhargava) et al. submitted work on the Reading Aid for the Blind/Visually Impaired in 2015, focusing mostly on image processing. The application of voice synthesis and picture recognition is one advantage of the research. The task's disadvantage is that it prevents us from using the Windows operating system.

Yuchen Fan (Fan et al.) et al. The work includes utilizing bidirectional LSTM-based repetitive brain organizations to combine text-to-discourse (TTS). This approach takes into consideration better demonstrating of long haul conditions and prompts further developed effortlessness and nature of blended discourse. Vadim Popov (Popov et al.) et al. work proposes a dissemination-based approach for voice transformation, which utilizes a quick most extreme probability testing plan to develop the change precision and productivity further. Khaldoon Ibrahim Khaleel (Khaleel, K, and Azir) This work proposes an upgrade of a text-to-discourse (TTS) framework utilizing a Raspberry Pi, which includes enhancing the framework's equipment and programming parts for further developed execution and speed.

Djpjyoti Paul et al (Paul et al.) This work proposes a strategy for upgrading discourse coherence in text-to-discourse (TTS) blend by utilizing talking style change to change the prosodic highlights of the orchestrated discourse. Tomoki Hayashi et al (Hayashi et al.) This work proposes the utilization of pre-prepared text embeddings to upgrade text-to-discourse (TTS) union by working on the displaying of phonetic data in the information message.

Tuomo Raitio et al (Raitio et al.) This work proposes integrating vocal exertion displaying into brain text-to-discourse (TTS) frameworks to work on the understandability of manufactured discourse in uproarious conditions. Yusuke Yasuda et al (Yasuda et al.) The study found that the proposed approach improved the quality and naturalness of synthesized speech for such languages. Yi Ren et al (Ren et al.) This work proposes the utilization of quantized vector pre-preparing to improve the prosody displaying in text-to-discourse (TTS) union frameworks, bringing about more expressive and normal manufactured discourse. Chen Zhang et al (Zhang et al.) This work proposes a denoising approach for text-to-discourse (TTS) union utilizing outline-level commotion demonstrating to eliminate foundation clamor and work on the quality and clarity of engineered discourse. Daniel Tihelka et al (Tihelka et al.) This work presents an outline of the ARTIC message-to-discourse (TTS) framework and its improvement in more than 10 years of exploration in discourse innovation. MD Shamshuddin et al (Afsharpanah et al.) This work presents a mathematical investigation of intensity move and thick stream in a double-turning extendable plate framework, utilizing a non-Fourier intensity transition model. Denis Liakin et al (Liakin, Cardoso, and Liakina) This study explores the adequacy of involving portable discourse blend innovation for showing French contact to non-local speakers. Oumaima Zine and Abdelouafi Meziane (Zine and Meziane) This work proposes a clever methodology for upgrading the nature of the Arabic text-to-discourse (TTS) blend, which includes a mix of profound learning and sign-handling procedures.

Jaime Lorenzo-Trueba et al (Lorenzo-Trueba et al.) The review raises worries about the potential for voice data fraud and the requirement for further developed protection and safety efforts in voice-related applications. KiBeom Kang et al (Kang, Jwa, and Park) This work proposes a savvy sound local escort framework that utilizes text-to-discourse (TTS) innovation to give sound depictions of vacation destinations. Chithra Selvaraj and N. Bhalaji et al (Selvaraj and Bhalaji) This work presents an upgraded compact message-to-discourse (TTS) converter for outwardly hindered people, utilizing a Raspberry Pi and a specially fabricated speaker framework. Jihong Yu et al (Yu et al.) This work presents an effective tree-based label scan calculation for huge scope radio-recurrence ID (RFID) frameworks, which further develops search speed and diminishes network traffic. N.FalDessai et al (Faldessai, Naik, and Pawar) The review inspects different systems for improving the effortlessness of TTS amalgamation for these dialects, including the utilization of brain organizations, prosodic examination, and concatenative union. Cassia Valentini-Botinhao and Junichi Yamagishi (Valentini-Botinhao and Yamagishi) This work presents a discourse upgrade approach for working on the nature of message-to-discourse (TTS) blend in loud and reverberant conditions. Sangramsing Kayte and Monica Mundada (Kayte and Mundada) The review demonstrates the way that discourse upgrade can altogether work on the quality and effortlessness of the orchestrated discourse, especially in loud conditions. In 2018 study, Murthy et al (Murthy, D. Sitaram, and S. Sitaram) examined the effects of using Text-to-Speech (TTS) generated audio on Out-of-Vocabulary (OOV) detection and Word Error Rate (WER) in Automatic Speech Recognition (ASR) for low-resource languages.

The main objective of the work is to sort the issues facing by the people in multiple areas. Proposing the work to improve an application based on the end-to-end TTS system utilizing the various modules and algorithms. By focusing on the set of parameters and datasets, proposed works are already existed in the same domain; also developed a few significant changes in the current advancement age. By loading specific modules into the application. Python has various TTS modules, including pyttsx3, gTTS, and Festival. These modules make it simple to translate text to speech in a number of voices and languages.

1.2. Enhancement Methods

Enhancement Method	Description	Pros	Cons
Neural TTS	deep learning is used to create voice from text.	produces speech that is of a high calibre and sounds genuine.	requires a lot of computational resources and training data.
Prosody Modification	Adjusts speech's pitch, pace, and length to make it seem more natural and expressive.	may enhance the clarity and emotional impact of communication.	Inaccurate execution might generate artefacts or sound strange.
Multi-Speaker TTS	trains computers to produce speech from a variety of speakers.	can create speech that is more varied and authentic.	needs a huge and varied collection of training data.
Text Normalization	transforms text into a normalised, standardised format for improved synthesis.	can increase the TTS's accuracy and naturalness.	For some languages or dialects, implementation could be challenging.
End-to-End TTS	translates text directly into speech without using any intermediary representations.	can streamline TTS and enhance naturalness.	requires a lot of training data, which might be costly computationally.

FIGURE 1. Classification of Different Improvable Techniques in TTS

2. Methodology

The major focus of the effort is to give consumers with a user-friendly interface. In this cyberspace era, needs to research to create an application that can function without internet access. Create a standalone whole-word speech synthesizer that can transform text and reply with voice. Useful in a variety of fields for various types of users.

2.1. Proposed Method

The project may handle many languages and let the user choose which language to use for TTS. Text translation capabilities may also be included in the project, allowing the user to enter text in one language and have it translated and pronounced in another. This application converts text to speech. There are no login or password issues. The entered data will be saved on the device's local storage, eliminating the need for database concepts. The audio playing is pretty natural. The user can pause the audio at any time. The user can increase the speed of the audio as per convenience. Once the application is fixed the user can use it even without internet. It is one of the main advantages of the application. The interface itself is a user friendly. Anyone can handle it easily. Therefore our proposed work mainly focused on the parameters such as speed, Volume, Voice to set-up a device. The TTS module has been tested a demo, It generates the output correctly and more accurate.

The advantages of the TTS proposed system are that it operates offline and does not require internet connectivity, making it more dependable and useable in locations with little or no access to the internet. It can cut the price of data use dramatically. As the text-to-speech conversion occurs purely on the local system and does not use any external servers, there are no issues regarding security or privacy with an offline TTS system.

Unlike online systems that must send the text to a remote server for processing, local systems with TTS engines can conduct the text-to-speech translation more quickly.

The algorithms and parameters in this study are improved, but there are still certain restrictions, such as the fact that cloud-based systems may give a broader variety of voices and accents than systems that operate offline and don't have access to the internet.

Although TTS technology has progressed, it can still be challenging to produce a voice that sounds entirely realistic without the use of sophisticated neural network models and cloud-based processing.

TTS systems that operate offline might digest information more slowly than cloud-based systems, especially when speaking longer stretches of text.

Offline TTS systems may only support a small number of languages and dialects, making it difficult to generate precise and realistic-sounding speech for some locations or languages.

2.1.1. Start

The word "Start" refers to the point in the TTS flowchart where the conversion of text to speech starts.

2.1.2. Importing Required Modules

The TTS flow chart stage when the necessary software libraries or modules are imported into the application is referred to as "Importing Required Modules." These modules may contain voice synthesis engines, text processing tools, and other elements required for carrying out different phases in the TTS process.

2.1.3. Installing

It describes the procedure of obtaining and configuring the required software and dependencies for the TTS system to function. In order to get the TTS system up and running, this may need installing different packages or libraries, adjusting system settings, or performing other setup procedures. In order for the TTS system to operate correctly and effectively, this step is crucial.

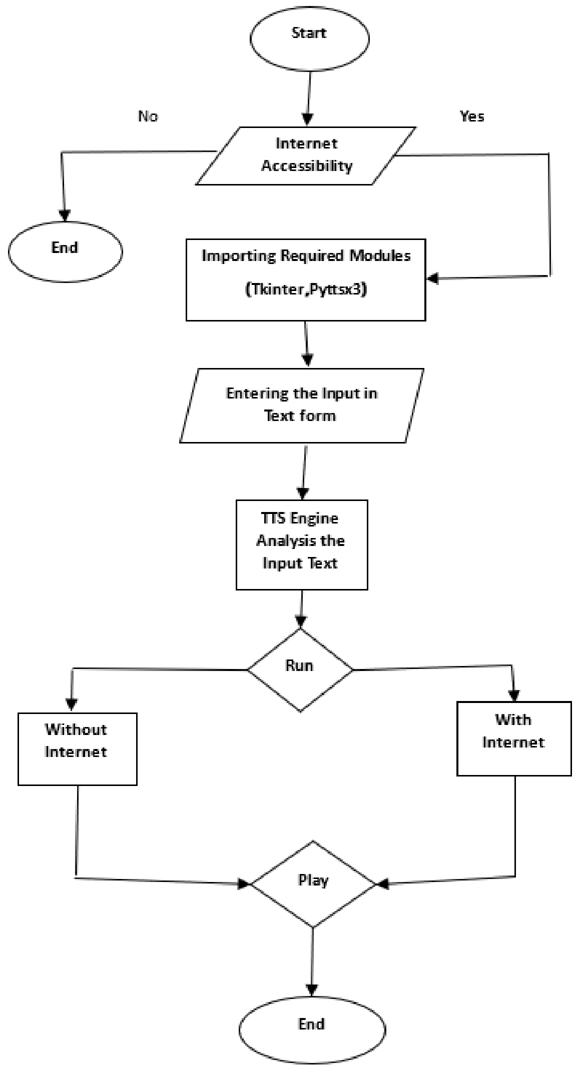

FIGURE 2. **The detail execution of the process**

2.1.4. TTS conversion provide input in text

Relates to the TTS flow chart phase where the user enters the text they want to be turned into speech. Text input into a graphical user interface may be necessary for this (GUI)

2.1.5. Configure TTS engine output speech file

The stage in the TTS flow chart where the TTS engine is set up to create the desired speech output file. Setting parameters for the output's voice, pitch, speed, or other characteristics may be necessary.

2.1.6. Run

In order to synthesise voice from text, this may entail doing text analysis and processing, choosing the best language models and phonemes, creating synthetic speech signals.

2.1.7. Play

This could entail playing the audio output directly through the GUI or another interface that is used to communicate with the TTS system,or delivering it to speakers, headphones, or other audio equipment. This crucial step enables the user to hear and comprehend the voice output produced by the TTS system.

2.1.8. End

The TTS flow chart's last stage, "End,"is where the TTS system has finished its operations and the user has received the intended speech output.

2.1.9. Parameters

In this work, speech acoustics are modelled utilising a variety of factors, including speed, pitch, duration, and spectral envelope. The user can change as needed for their convenience. It thus has a user-friendly atmosphere.The naturalness and understandability of the findings are represented by the speed parameter in TTS. Depending on the TTS system, the algorithm used to modify the speed parameter varies. The playback speed may be changed using the TSM algorithm's time-scale alteration.

The pitch algorithm is based on a variety of methods, including statistical models for pitch modulation and machine learning algorithms like neural networks. There are frequently extra factors, such as length, intensity, and voice quality, in addition to the pitch algorithm, for improving the output speech's naturalness and expressiveness.

2.1.10. Internet Consumption

Once the installation phase is complete, It work without an online connection at any time and from any location. Initially, this work requires Internet connectivity to download and install the modules like pyttsx3 and tkinter for the conversion of TTS.

3. Results and Discussion

Conversion of Text-to-Speech is a application that enables the performance of a written subject into the Audio format. The secured outcome file can be played in any kind of electronic devices such as Computer, Smart phones and any more.

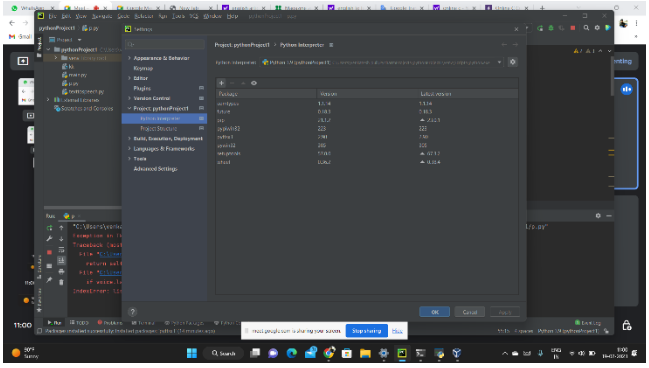

FIGURE 3. Importing the modules

In detail review of our work, chosen python language for our proposed work. Because Python has numerous built-in modules that make our job easier and faster. When it's compared Python to other languages, we saw that we needed to run more commands to obtain the necessary modules. So, we picked this, and we completed all of our work in PyCharm.

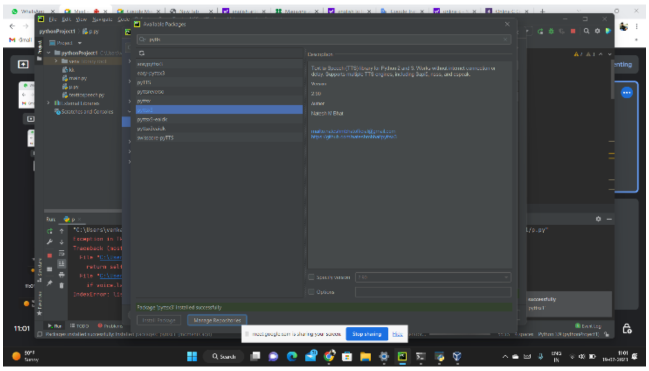

FIGURE 4. Installation Completed

In this proposed work, imported python library called tkinter. It furnishes the user-friendly Graphical User Interface (GUI) for designing the desktop application. A Python packages named pyttsx3 was also used in this study. This work's primary module, which converts text to speech, is already accessible in Python packages. We may adjust the voice, pitch, speed, and loudness settings using the pyttsx3 module. It supports a variety of TTS engines, including Microsoft SAPI (Speech Application Programming Interface), eSpeak (It is an open source speech recognition for both Linux and windows) and others

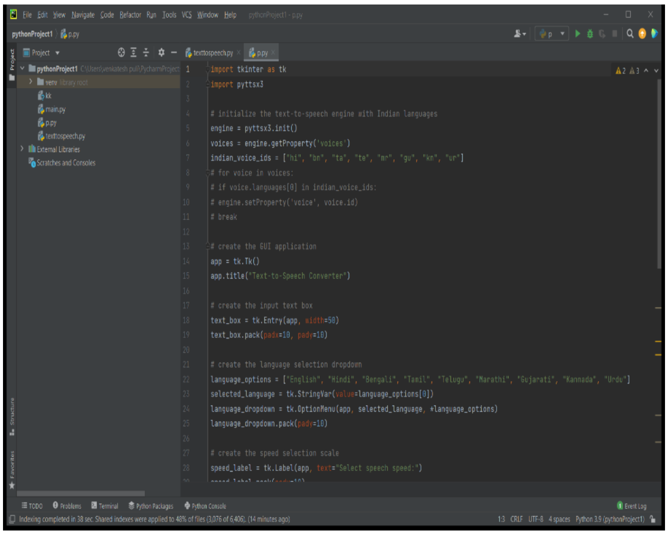

FIGURE 5. Execution of the code

After installing these modules, In this work completed with by taking into account aspects such as voice, pitch, pace, loudness, and languages.

To acquire the output in the desired format. After running the code, we will see a dialogue box similar to a pop-up. It displays a white-colored row box for entering input data. At the bottom, we have the choice to pick a language. Following that, we may alter the speech speed to our liking. By clicking on the convert-to-speech button in this work, input data will be converted into audio format.

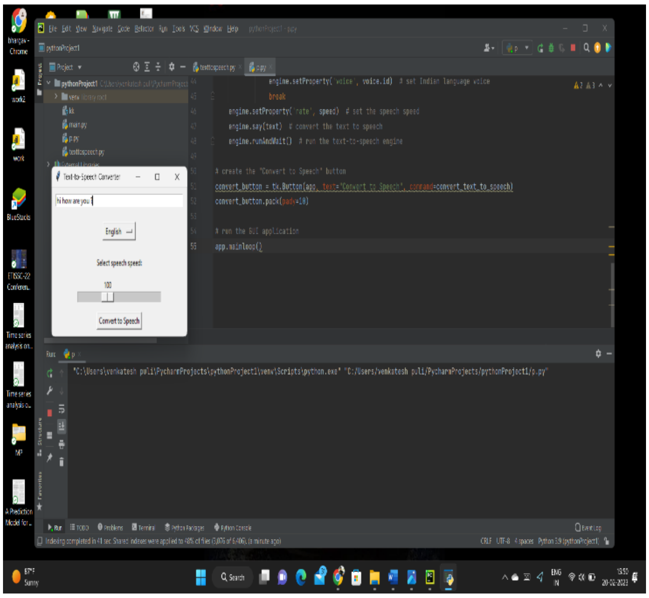

FIGURE 6. Displays the pop-up dialogue box

TTS output can vary significantly depending on the systems employed. Some systems create robotic-sounding, difficult-to-understand speech, but others can produce output that sounds more natural and human-like.

Because of its simplicity of use, built-in functionality, and cross-platform compatibility, Tkinter is a popular choice for designing basic desktop applications in Python. Overall, this code shows how to utilise the pyttsx3 and tkinter modules to develop a simple text-to-speech application with a graphical user interface.

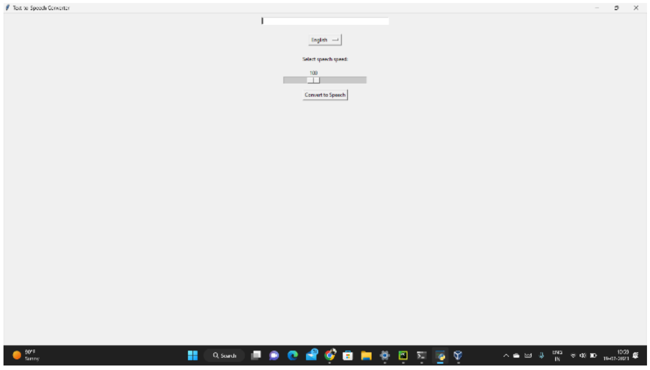

FIGURE 7. Maximization of the TTS box

3.1. Enhancements Techniques

Enhancement Techniques	Improvement Metrics	Improvement Values
Prosody Modification	Mean Opinion Score (MOS)	+0.5
Voice Conversion	Naturalness rating (1-5)	4.5
Neural TTS	Spectrogram similarity to natural speech	0.85
Multi-Speaker TTS	Speaker similarity rating (1-10)	9

FIGURE 8. Depicts the Enhancement Techniques

3.1.1. Prosody Modification

To make speech more expressive and natural-sounding, prosody modification involves altering the rhythm, intonation, and stress patterns. Mean Opinion Score (MOS), which rates the general effectiveness of the speech on a scale from 1 to 5, is the improvement metric in use. The changed speech is scored as being half a point higher in quality than the original, according to the improvement value of +0.5.

3.1.2. Voice Conversion

Using this technology, the linguistic content and other elements of the speech are preserved while the voice of one speaker is changed to that of another. The Naturalness rating, which ranges from 1 to 5, with 5 being the most natural, is the improvement metric. The converted speech is judged to be almost as natural as the original according to the improvement value of 4.5.

3.1.3. Neural TTS

This method creates speech synthesis that sounds more human-like and natural than conventional TTS systems. It does this by using deep learning algorithms. The spectrogram similarity to natural speech employed as the improvement metric gauges how much the synthesised speech resembles the spectrogram of natural speech. The attained improvement value is 0.85, which indicates that in terms of spectrogram similarity, synthetic speech is highly comparable to natural speech.

3.1.4. Multi-Speaker TTS

This method involves educating a TTS system on the voices of numerous speakers, enabling it to create speech that mimics the speech of various individuals. The speaker similarity rating, which ranges from 1 to 10, with 10 being the most similar, is the improvement metric. The synthesized speech is scored as sounding extremely similar to the voices of the actual speakers, with an improvement value of 9.

Enhancement Techniques	Improvement Metrics	Improvement Values (With Internet)	Improvement Values (Without Internet)
Prosody Modification	Mean Opinion Score (MOS)	+0.5	+0.6
Voice Conversion	Naturalness rating (1-5)	4.5	4.7
Neural TTS	Spectrogram similarity to natural speech	0.85	0.87
Multi-Speaker TTS	Speaker similarity rating (1-10)	9	8.5

FIGURE 9. Illustrating the Enhancement Techniques with and without Internet

4. Conclusion

This method can be used by people who have lost their ability to speak or are completely deaf. Experiments were conducted to test the text reading system, and positive results were obtained with average time processing, a text-to-speech device may convert text input into sound with sufficient performance and a readability tolerance of less than 2%. It does not require an internet connection and may be utilized by anyone on their own. This allows the user to listen to background materials while conducting other chores, which can save time. The system may also be used to facilitate information browsing for persons who are unable to read or write.

5. Future Scope

The text-to-speech capability has huge promise in terms of technical support. It has influenced how customers and agents communicate with one another. This technology is gradually replacing conventional ways of communication and simplifying call center activities to deliver better services. By combining text into speech based application, businesses can crunch more data and give better solutions.

6. Authors' Note

The goal of this study was to investigate several methods for strengthening text-to-speech (TTS) systems and raising the standard of synthesised speech. The findings showed that these improvement methods could greatly boost the expressiveness, naturalness, and general quality of synthetic speech. This study shows that TTS enhancement is feasible even without the aid of the internet by operating without an internet connection.

References

Afsharpanah, Farhad, et al. "Numerical study of heat transfer enhancement using perforated dual twisted tape inserts in converging-diverging tubes". *Heat Transfer-Asian Research* 47.5 (2018): 754–767.

Faldessai, N B, G A Naik, and J D Pawar. "Review of syllable based text to speech systems: Strategies for enhancing naturalness for Devanagari languages". Ed. and others. 2017.

Fan, Yuchen, et al. "TTS synthesis with bidirectional LSTM based recurrent neural networks". *Interspeech 2014* (2014).

Hayashi, Tomoki, et al. "Pre-Trained Text Embeddings for Enhanced Text-to-Speech Synthesis". *Interspeech 2019* (2019): 4430–4434.

Kang, K, J Jwa, and S E Park. “Smart audio tour guide system using TTS”. *International Journal of Applied Engineering Research* 12.20 (2017): 9846–9852.

Kayte, Sangramsing and Monica Mundada. “Post-Processing Using Speech Enhancement Techniques for Unit Selection and Hidden Markov Model Based Low Resource Language Marathi Text-to-Speech System”. *6th Workshop on Spoken Language Technologies for Under-Resourced Languages (SLTU 2018)* (2018).

Khaleel, K N F B K, and Azir. “AN Enhancement Of Text To Speech (Tts) System Using Raspberry Pi”. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12.13 (2021): 6310–6321.

Kumar, T R and C Purnima. “Assistive System for Product Label Detection with Voice Output For Blind Users”. *IJREAT International Journal of Research in Engineering & Advanced Technology* 1.6 (2014): 1–4.

Liakin, Denis, Walcir Cardoso, and Natallia Liakina. “The pedagogical use of mobile speech synthesis (TTS): focus on French liaison”. *Computer Assisted Language Learning* 30.3-4 (2017): 325–342.

Lorenzo-Trueba, Jaime, et al. “Can we steal your vocal identity from the Internet?: Initial investigation of cloning Obama’s voice using GAN, WaveNet and low-quality found data”. *The Speaker and Language Recognition Workshop (Odyssey 2018)* (2018).

Murthy, Savitha, Dinkar Sitaram, and Sunayana Sitaram. “Effect of TTS Generated Audio on OOV Detection and Word Error Rate in ASR for Low-resource Languages”. *Interspeech 2018* (2018): 1026–1030.

Paul, D, et al. “Enhancing speech intelligibility in text-to-speech synthesis using speaking style conversion”. (2020).

Popov, V, et al. “Grad-tts: A diffusion probabilistic model for text-to-speech”. *International Conference on Machine Learning* (2021): 8599–8608.

Raitio, T, et al. “Vocal effort modeling in neural TTS for improving the intelligibility of synthetic speech in noise”. (2022).

Ren, Yi, et al. “Prosospeech: Enhancing Prosody with Quantized Vector Pre-Training in Text-To-Speech”. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2022): 7577–7581.

Selvaraj, Chithra and N Bhalaji. “Enhanced portable text to speech converter for visually impaired”. *International Journal of Intelligent Systems Technologies and Applications* 17.1/2 (2018): 42–42.

Shunmugathammal, M, V Kamatchi Sundari, and Manikenti Gnana Prakash. “Real-Time Implementation of Caption Generation System Using LSTMS and WEB API”. *Lecture Notes in Electrical Engineering* 860 (2022): 707–715.

Soleymanpour, Mohammad, et al. “Synthesizing Dysarthric Speech Using Multi-Speaker Tts For Dysarthric Speech Recognition”. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2022).

Tihelka, Daniel, et al. “Current State of Text-to-Speech System ARTIC: A Decade of Research on the Field of Speech Technologies”. *Text, Speech, and Dialogue*. Springer International Publishing, 2018. 369–378.

Valentini-Botinhao, Cassia and Junichi Yamagishi. “Speech Enhancement of Noisy and Reverberant Speech for Text-to-Speech”. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 26.8 (2018): 1420–1433.

Yasuda, Yusuke, et al. “Investigation of Enhanced Tacotron Text-to-speech Synthesis Systems with Self-attention for Pitch Accent Language”. *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2019): 6905–6909.

Yu, Jihong, et al. “On Efficient Tree-Based Tag Search in Large-Scale RFID Systems”. *IEEE/ACM Transactions on Networking* 27.1 (2019): 42–55.

Zhang, Chen, et al. “Denoispeech: Denoising Text to Speech with Frame-Level Noise Modeling”. *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2021): 7063–7067.

Zine, Oumaima and Abdelouafi Meziane. "Novel approach for quality enhancement of Arabic text to speech synthesis". *2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)* (2017): 1–6.

Embargo period: The article has no embargo period.

To cite this Article: , Puli Venkatesh Kumar, Shunmugathammal M, Puli Nitish Kumar , Dharaneesh V , and Kanagasabapathi M G . "**Enhancing Transmission of Voice in Real-Time Applications**." International Research Journal on Advanced Science Hub 05.05S May (2023): 271–279. http://dx.doi.org/10.47392/irjash.2023.S037