iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** VII **Month of publication:** July 2020

DOI: https://doi.org/10.22214/ijraset.2020.29892

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

A Novel Approach for Stutter Speech Recognition and Correction

Amruth V[1], Lavanya K[2], Manoj N[3], Umme Hani[4], Deepika M B[5]
[1]*Assistant Professor, Information Science and Engineering, Maharaja Institute of Technology Mysore, India*
[2, 3, 4, 5]*Information Science and Engineering, Maharaja Institute of Technology Mysore, India*

Abstract: ***Stuttering also called as Stammering is a speech disorder where a person suffering from this disorder repeats or prolongs words, syllables or phrases. He/ She may also stop during speech and make no sound for certain syllables. Current speech recognition systems such as Google Assistant, Apple's Siri etc have very efficient Speech recognition system for normal speech, but they fail to recognise when the person stutters. This paper proposes a system to detect as well as remove stutter from speech by using Speech-to-text Conversion method. This system is implemented in five stages namely, Amplitude thresholding and filtering, Silence ejection, speech to text conversion, repetition removal and text to speech (TTS) conversion.***
Keywords: Stuttering, Silent Pauses, Silence ejection, Speech-to-text, Tokenization, Python.

I. INTRODUCTION

Speech is human vocal communication using language. Each language uses phonetic combinations of vowels and consonant sounds that form the sound of its words, using those words sentences are formed to convey some message. But not everyone is blessed with good speaking capability. Speech disorders or speech impediments are the type of communication disorders where normal speech is disrupted. There are basically three types of Speech impairments such as articulation disorder, Fluency disorders, and voice disorders. Stuttering is a speech disorder which involves frequent problems with normal fluency and flow of speech. About 1% of total world's population suffer from this disorder. Most cases of stuttering begin in childhood. Fewer than 1% of stuttering cases are seen in adults. Stuttering can be caused due to combination of factors such as Family history of stuttering, Intellectual disability, Brain injuries or severe medical causes and it can also occur due to Emotional and mental health problems. For instance, some words are repeated or prolonged while others are preceded by an 'um' or 'uh'. In most cases, stuttering may have major impact on some daily activities. Every Person who stutter may find specific everyday activities to be challenging to perform, which may vary across individuals. For example, for some people, communication difficulties happen only during specific activities, like talking on the phone, talking before large groups, utilizing everyday tools that use speech as inputs. Speech recognition systems such as Google Assistant, Microsoft Cortona, Apple's Siri have good accuracy for normal speech. But these systems fail to recognize speech with repetitions or long involuntary pauses. This problem occurs because these systems are created to stop the identification process when a pause is encountered and are trained with words without any repetitions and hence, when it encounters a stuttered speech, it fails to identify the words. This is because when a person starts stuttering the system assumes that the person has ended speaking and thus it stops recognizing when it encounters stutters such as silent pauses. Hence, this paper aims to detect as well as correct the stutter speech by removing prolongations, repetitions and silent pauses in the system in real time mode. An algorithm can be built using Speech-to-text Conversion system such as Google's API along with a few string operations to convert speech input into Text format and to perform some string operation to remove stutters from the input speech. The language which is used to implement this system is Python.

II. LITERATURE SURVEY

Many papers have been proposed based on stuttered speech recognition and correction. This section discusses about the methodologies proposed in those papers briefly.

Ankit Dash et al. developed an algorithm for detection and correction of speech disfluency using MATLAB. By dividing the sample into short overlapping frames, Then, followed by Amplitude thresholding is done using neural networks to remove prolongations. Repetitions of string are removed using existing Text to Speech (TTS) method. Which was tested for 50 samples out of which an accuracy of 86% is obtained [1].

Arya A Surya et al. proposed three methods for both stuttered speech recognition as well as correction, using a supervised model, stuttering pruning, and automated text-to-speech based recognition [2]. In the supervised model for stuttered speech recognition N audio signals are converted to audio array. The Mel Frequency Cepstral Coefficient (MFCC) features are extracted from the audio

samples. During the classification stage, the extracted features are used to train support vector machine [2]. Since automatic speech recognition is a multiclass problem, SVM is extended to perform multiclass classification though SVM is basically a binary non-linear classifier. During testing, SVM classifies the stuttered input to correct word. Several words were manually segmented from the collected dataset. The SVM model is then trained using the segmented stuttered words. An accuracy of 76% was acquired in classifying the words correctly [2]. Fathima Afroz et al. has implemented a method to recognize and classify the pauses which occur in stuttered speech. Based on energy and zero crossing rate the process of Identification of pauses has been done using automatic blind segmentation approach. The Intra-morphic pauses for stuttered and Inter-morphic pauses for normal speech are compared and an accuracy of 98% is obtained by this method [9].

Lim Sin Chee, proposed the feature extraction method for the recognition of repetition and prolongation using MFCC. Two simple classifiers such as LDA based classifier and K-NN classifier were used for testing the effectiveness of MFCC feature in the recognition of repetitions and prolongations in stuttered speech. The results show the average accuracy of 90% [7].

Adam Kobus, presented a syllable repetition detection method based on Linear Prediction Coefficient obtained by Levinson. The algorithm written is based on Linear Prediction Spectrum where at first the utterance is split into fragments and that corresponds with syllables and at last after dimension reduction by K-means method is compared [5].

P.Mahesha, proposed a new LH-MFCC Feature extraction technique to capture the temporal, instantaneous amplitude and Frequency characteristics of speech. Comparative Analysis of MFCC and LH-MFCC for three dysfluency speech is done here. The LH-MFCC shows the improved performance in all dysfluency classification [8].

P.S. Savin and Pravin B described the focus on repetition and prolongation detection in stuttered speech signal. The acoustic feature like the MFCCs, ZCR in recognizing repetition and prolongation in stammered speech. ANN are used as classifier and a feed forward network using Back propagation algorithm is used for speech pattern classification because of their discrimination and input-output mapping ability and hence the results are evaluated using combination of different features [12].

III. ALGORITHM FOR PROPOSED METHOD

This section explains the procedure to remove stuttering from a stuttered speech sample. The Figure 1 shows the block diagram for the proposed system.

A. Stutter speech is taken from the user as an input through the microphone of the laptop/system.
B. Magnitude filtering is done in which a threshold value of the amplitude is obtained using which the signals that are having frequency less than that threshold value is discarded as noise/ unwanted signals.
C. Speech to text conversion is done for the inputted audio signal to obtain equivalent text for the given input.
D. Silence ejection is done by using word splitting method in which Tokenization of the text is done to extract words from the sentences. By this method silent pauses that exist between any two words are removed.
E. After successful removal of silent pauses and prolongations, Repetitions of word in the text are detected and removed using string comparison method in python.
F. The words free from stutter are combined to form sentence and displayed to the using in text from.
G. The output in form of text is converted back to speech using existing Text to Speech (TTS) system.

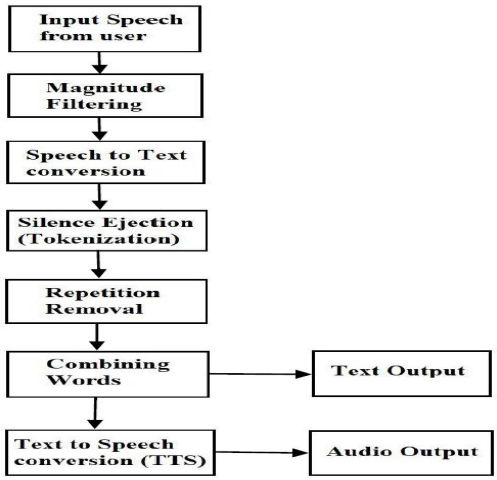


Figure 1. Block diagram for Proposed system

IV. RESULTS DISCUSSION

The results obtained for different types of stutters is discussed here by considering different speech inputs having different types of stutters such as Prolongations, Repetitions and silent pauses.

A. Stutter Speech input containing Prolongations

The system was tested for input "Hello this is my project" which was having "is" and "my" as prolonged words.

1) *Input:* "Hello this iiiiissss mmmmyyyy project"
2) *Output:* "Hello this is my project"

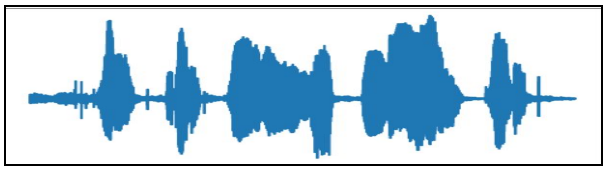


Figure 2(a). Input of a stutter speech signal containing Prolongations

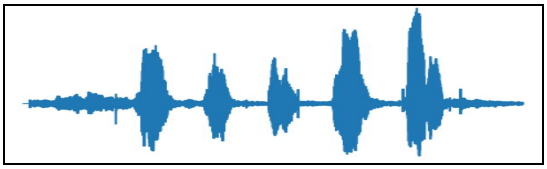


Figure 2(b). Output of a stutter speech signal shown in Fig 2(a)

The above Figure 2(b) shows the output obtained for the stutter speech after removal of prolongations present in input speech. The obtained output was devoid of stutters.

B. Stutter Speech input containing Repetitions

The system was tested for input "This is my project" which was having "this" and "my" as Repeated words.

1) *Input:* "This This is my my project"
2) *Output:* "This is my project"

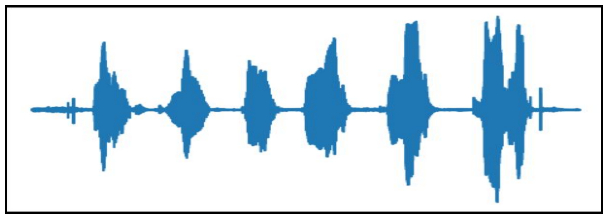


Figure 3(a). Input of a stutter speech signal containing Repeated words

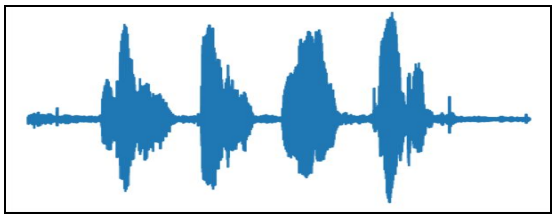


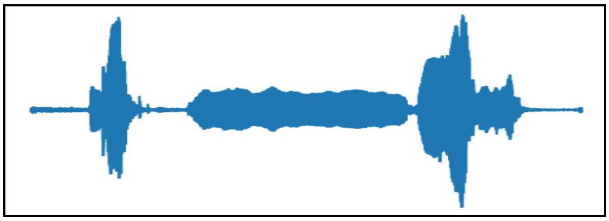
Figure 3(b). Output of a stutter speech signal shown in Fig 3(a)

The above Figure 3(b) shows the output obtained for the stutter speech after removal of all the Repetitions present in input speech. The obtained output was devoid of stutters.

C. Stutter Speech input containing Repetitions

The system was tested for input "Good morning" which was having a silent pause between words "Good" and "morning".

1) *Input:* "Good morning"
2) *Output:* "Good morning"



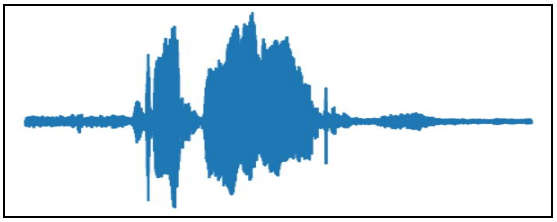


Figure 4(a). Input of a stutter speech signal containing Silent pause Figure 4(b).Output of a stutter speech signal shown in Fig 4(a)

The above Figure 4(b) shows the output obtained for the stutter speech after removal of silent pause present in input speech between two words. The obtained output was devoid of stutters.

By Analysing the outputs, the following results are obtained,

No. of Samples considered = 50

No. of samples corrected efficiently = 43

$$\text{Accuracy} = \frac{\text{No. of samples corrected efficiently}}{\text{No. of samples considered}} \times 100$$

Thus, the accuracy of 86% is obtained from the outputs obtained by the system.

V. CONCLUSIONS

Stuttering is the most common type of speech disorder which can be seen in almost 1% of the total world's population. The proposed system helps to correct stuttered speech sample into a normal speech sample in real time mode with an accuracy of 86%. Repetitions, prolongations and silent pauses are can be removed by using techniques such as Magnitude filtering, Speech to Text conversion, Tokenization and repeated words removal technique. The output can be obtained by the user in two formats i.e Text format and Audio format to make the output both visible and audible by the user.

This system can be integrated with existing systems that use speech as input, to make the system usable by people having stutter disorder. By implementing this, people having stuttering disorder feel free to communicate and interact with others without worrying about their disability.

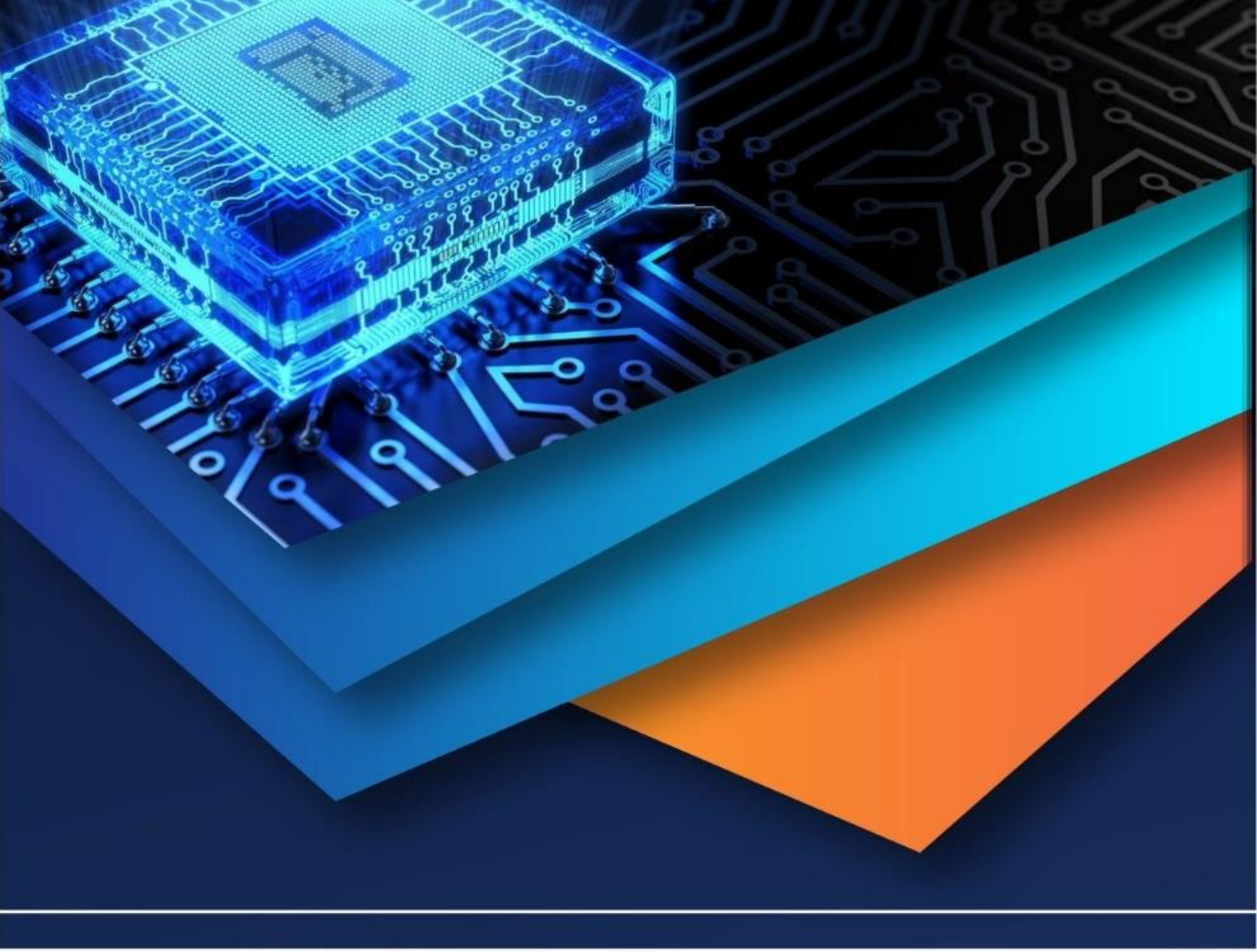
VI. FUTURE WORK

The system to detect as well as correct stuttered speech can be extended to implement in languages other than English. The Existing speech to text and text to speech systems which are based on Neural networks, these systems can be improvised by using more efficient machine learning techniques.

Text to speech system can be implemented by developing speech synthesis in order to get the output in the original speaker's voice. Thus, an application can be developed which internally implements this method and produce stutter free output in his/her original voice within a short time duration.

REFERENCES

[1] Ankit Dash, Nikhil Subramani, Tejas Manjunath, Vishruti Yaragarala and Shikha Tripathi., Speech Recognition and Correction of a Stuttered Speech, IEEE 2018.
[2] Arya A Surya, Surekha Mariam Varghese., Automatic Speech Recognition System for Stuttering Disabled Persons, International Journal of Control Theory and Applications, Volume 9 Number 43, 2016
[3] G. Diwakar and Veena Karjigi, Repetition Detection in Dysarthric Speech, IEEE 2017.
[4] Gunjan Jhawar, Prajacta Nagraj, and P. Mahalakshmi., Speech Disorder Recognition using MFCC, IEEE 2016.
[5] Adam Kobus, Wiesława Kuniszyk-Jó´zkowiak and Ireneusz Codello., Automatic Syllable Repetition Detection in Continuous Speech Based on Linear Prediction Coefficients", Springer International Publishing Switzerland 2016.
[6] Shweta Khara, Shailendra Singh, Dharam Vir., A Comparative Study of the Techniques for Feature Extraction and Classification in Stuttering, IEEE 2018.
[7] Lim Sin Chee, Ooi Chia Ai, M. Hariharan and Sazali Yaacob., MFCC based Recognition of Repetitions and Prolongations in Stuttered Speech using k-NN and LDA, IEEE 2009.
[8] P. Mahesha and D.S. Vinod., LP-Hillbert Transform Based MFCC for Effective Discrimination of Stuttering Dysfluencies, IEEE 2017.
[9] Fathima Afroz and Shashidhar G Koolagudi., Recognition and Classification of Pauses in Stuttered Speech using Acoustic Features, IEEE 2019.
[10] Dawid Połap., Extraction of specific data from a sound sample by removing additional distortion, IEEE 2017.
[11] Pravin B.Ramteke, Shashidhar G. Koolagudi and Fathima Afroz., Repetition Detection in Stuttered Speech, Springer India 2016.
[12] P.S. Savin, Pravin B. Ramteke and Shashidhar G. Koolagudi., Recognition of Repetition and Prolongation in Stuttered Speech Using ANN, Springer India 2016.
[13] Nishant Mishra, Prasun Anand., Performance Evaluation of Neural Networks for Speaker Recognition, IEEE 2019.
[14] Manjula G and Shiva Kumar M., Development of a Recording Protocol for the Assessment of Speech Disorders, IEEE 2016.
[15] Miss. Salma Jabeen, Dr K.M. Ravikumar., Analysis of 0dB and 10dB Babble Noise on Stuttered Speech, IEEE 2015.
[16] Arkadiy Prodeus and Kateryna Kukharicheva., Training of Automatic Speech Recognition System on Noised Speech, IEEE 2016.
[17] Balaji V and Dr. G. Sadashivappa., Speech Disabilities in Adults and the suitable Speech Recognition Software Tools– A Review, IEEE 2015.
[18] Iman Esmailia, Nader Jafarnia Dabanlooa, Mansour Vali., Automatic classification of speech dysfluencies in continuous speech based on similarity measures and morphological image processing tools, Elsevier Ltd 2015.
[19] Jubo Zhang, Bin Dong, Yonghong Yan, A Computer-assist Algorithm to Detect Repetitive Stuttering Automatically, IEEE 2013.
[20] Juraj Palfy, Ji Pospichal., Pattern Search in Dysfluent Speech, IEEE 2012.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY