International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-5, June 2019

Speech Synthesis System using LabVIEW

Uma J, Vanitha P, Renuka R, Sweta S J

Abstract*: The present work is synthesis of voice signal using LabVIEW with the input of text signal in the form of image file, typed text or pdf. The synthesis of signal based on the Optical Code Recognition (OCR) techniques. Even though already so many technologies available in the synthesis of voice signal, the LabVIEW based system is very user friendly and cost effective. The programming of LabVIEW is graphical based enhances the users to be compatible. The proposed system has two parts such as optical character recognition and text to speech conversion (TTS). The Text to Speech interfaces is provided by Microsoft Speech SDK. In addition to that LabVIEW supports special characters and special language characters. The input of either a text document or written text is read and it is converted into speech signal. The output of an audio file is listened by headset or speaker connected with the processor. The above proposed system of a speech synthesis is very useful for educational purpose, for visually impaired people and for shorthand writing.

Index Terms*: LabVIEW, Microsoft Speech SDK, Optical Character Recognition (OCR), Text to Speech Conversion (TTS)

I. INTRODUCTION

Speech synthesis converts the normal language text into the voice signal or speech signal. The synthesizer can have the model of vocal tract. The system of production of speech signal from the text or image file is called as synthesizer. Speech synthesizer is often called as Text To Speech (TTS) system having symbolic linguistic representations like phonetic transcriptions into speech. A TTS comprises of text normalization and synthesizer. In text normalization converts the text into equivalent of written out words and in synthesizer converts the words and numbers into sound signal. TTS can read a text from a document, Web page or e-Book, generating synthesized speech through a computers speaker. TTS is mainly useful for disabled such as poor vision. In [1] TTS is implemented using a hybrid strategy, a minimal set of rules to predict the lexical stress. The present study [2] represents the speech synthesis using concatenative data driven methods and also it compares the musical synthesis and speech synthesis related to tone and naturalness. In the report of [3] uses elicitation approach for improvement of syllabification process and also it compare different spectral smoothing methods. The author proposed [4] unit selection algorithm for Marathi language synthesis and uses units like words, Di-phone and tri-phones as database. The author [5] has used unit selection algorithm Using Bi- grams Model for speech synthesis. The author [6] has introduced the speaker recognition algorithm. Here speech classification is done through Output Probability Clustering algorithm. The paper proposed [7] acoustic units from speech using property rules and text processing stage, by extracting linguistic units using phonetic and lexical rules. In this paper [8] author has uses statistical model based algorithm for minimizing mean squared error from speech enhancement. The author [9] proposed Multilingual Speech Synthesis System for Indian English, Tamil and Telugu by using Concatenative speech synthesis. Authors [10]have used Concatenative Synthesis method for Marathi language and the quality of generated speech depends on the unit size. The author [11] focused on synthesis of speech signal only for Tamil language using Concatenative synthesizer. The author introduces that the speech listeners used combinational information about rhythmic structure, phrasal structure and phone tactics to segments their native language.

The present study helps the people requiring voice signal from the text input especially in banking sector and for blind people. In the present study the possibilities of transforming written text signal into voice or audio signal using LabVIEW programming environment. The TTS is specially provided by Microsoft SDK and in addition LabVIEW supports special characters and special language characters. Input is given either as a text document or written text which is read, converted into speech and the output is produced as an audio file.

II. SUB VI FOR SPEECH SYNTHESIS

A. Initialize Routine

The input of text or image signal is given in the case structure as in the name of initialize. Whenever the new input is fed into the text to be spoken flag box and the power ON terminal in the front panel of main VI is enabled. The process of synthesizing a audio signal will take place .The case structure of first sequence execution is when both the inputs of the selector switch initialize and power ON need to be enabled to take in the input text message of the dialog box.

The Fig. 1 shows the block diagram of Initialize sub VI and Fig. 2 shows the front panel diagram of Initialize sub VI

Manuscript published on 30 June 2019.
***** Correspondence Author (s)
Uma J*, Electrical and Electronics Engineering, M.Kumarasamy College of Engineering, Karur, Tamil Nadu, India.
Vanitha P, Electrical and Electronics Engineering, M.Kumarasamy College of Engineering, Karur, Tamil Nadu, India.
Renuka R, Electrical and Electronics Engineering, M.Kumarasamy College of Engineering, Karur, Tamil Nadu, India.
Sweta S J, Electrical and Electronics Engineering, M.Kumarasamy College of Engineering, Karur, Tamil Nadu, India.

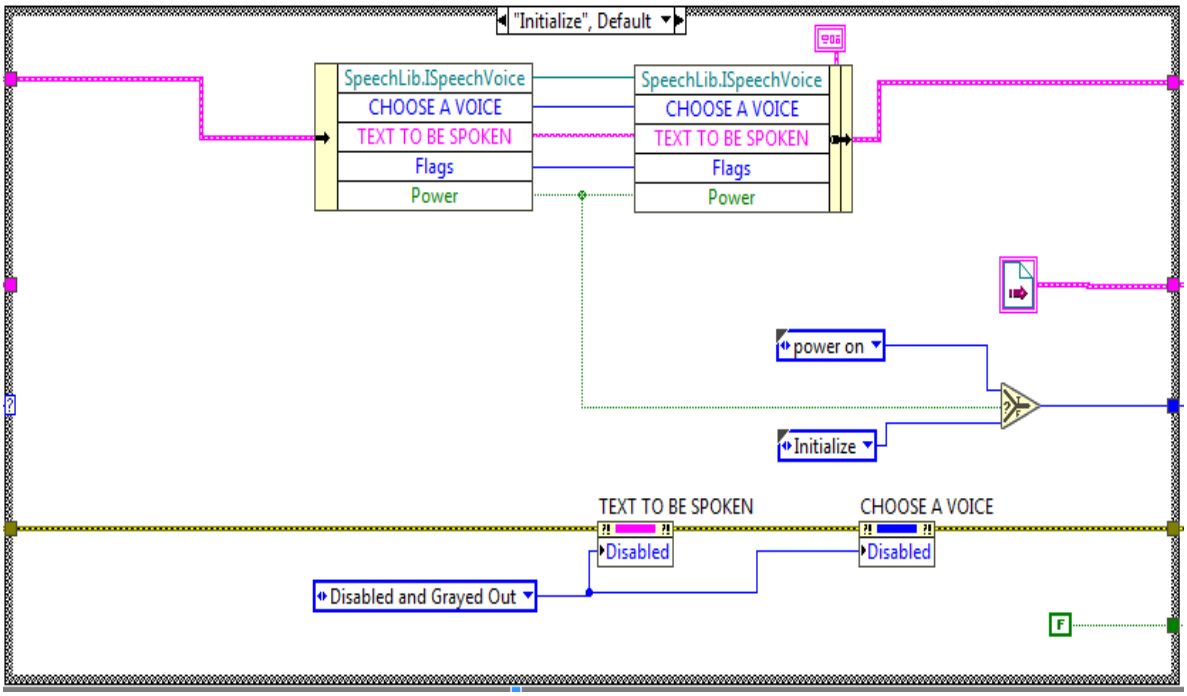

Figure. 1The block diagram of Initialize sub VI

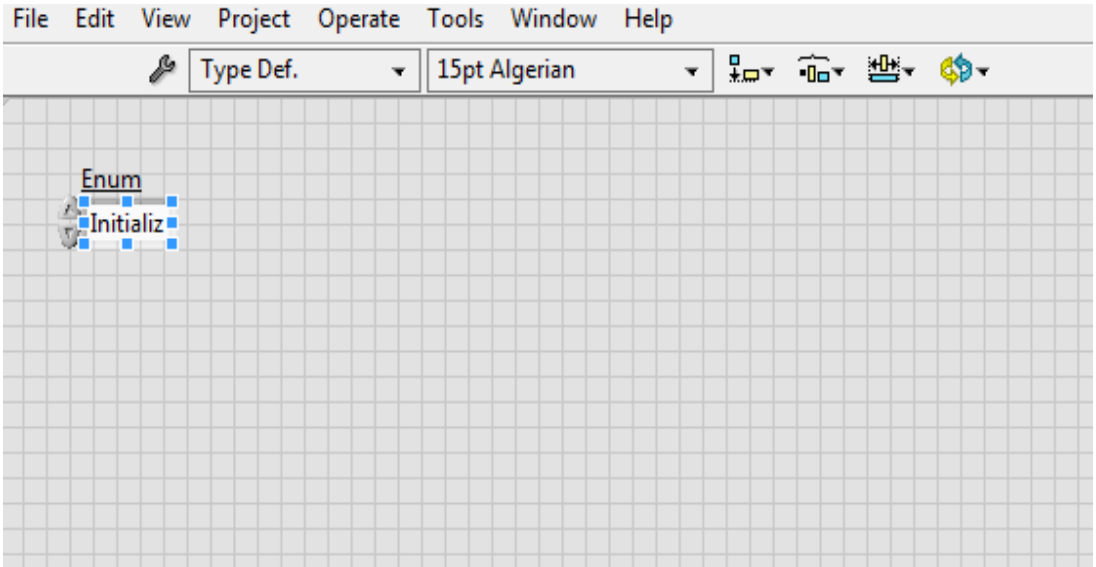

Figure. 2 The front panel diagram of Initialize sub VI

B. Power ON Routine

In the power ON routine the user has to select the text to be spoken and also a type of voice signal .After the input signal is fed it will allow the read dialog bow to read the data. The Fig. 3 shows the power on case module

C. Read Routine

In the Fig. 4 of Read routine, the text to be spoken and the type of voice need to be output is selected. The process of text normalization and synthesize of the audio signal is happening in the particular case structure sequentially. The LabVIEW has icon of Speech Voice, SpeechObjectTokens, SpeechLib, SpeechVoice modules. After a conversion the output is able to listen in the headset or speaker connected to the processor.

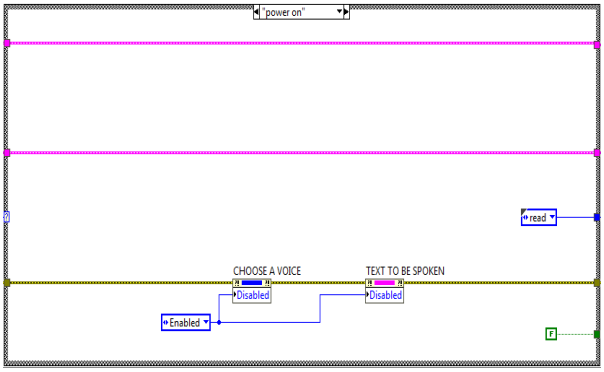

Figure 3 Power on case module

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

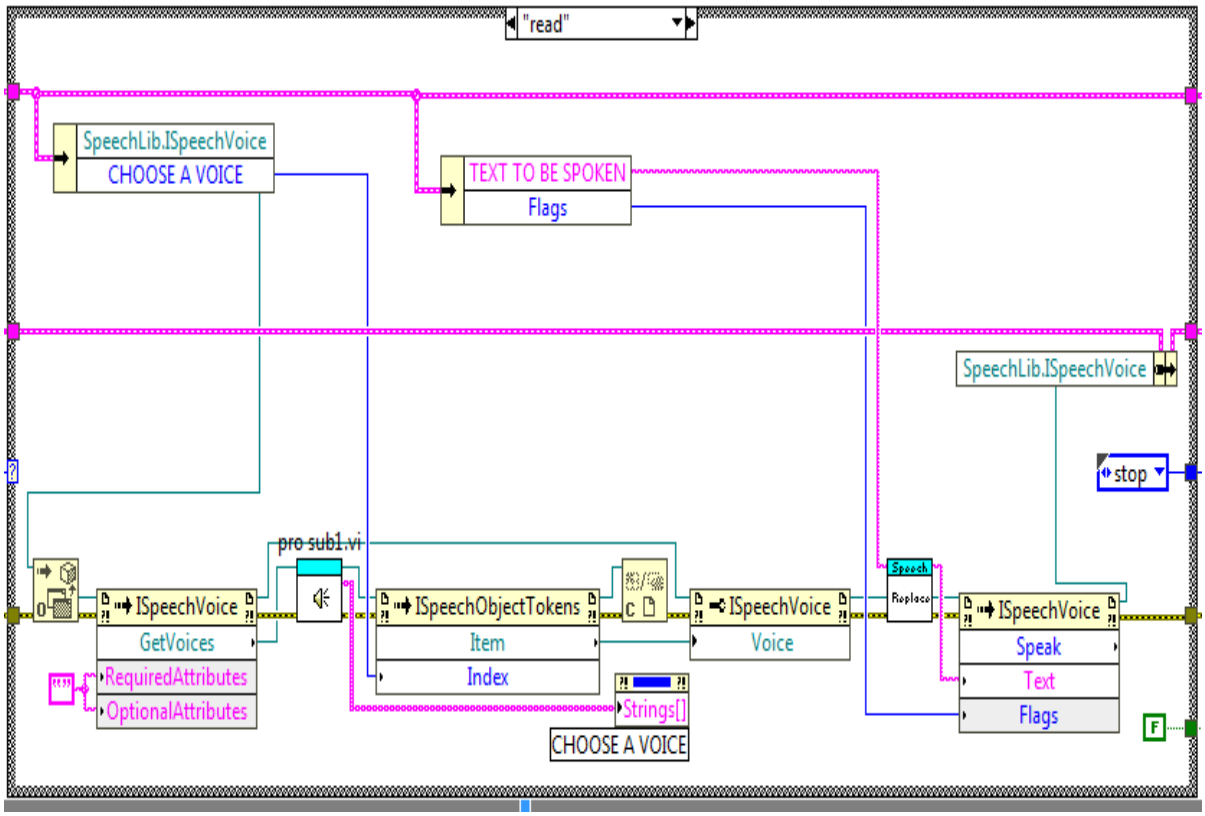

Figure 4. Read routine block diagram

I. MAIN BLOCK EXECUTION OF SPEECH CONVERSION TECHNIQUES

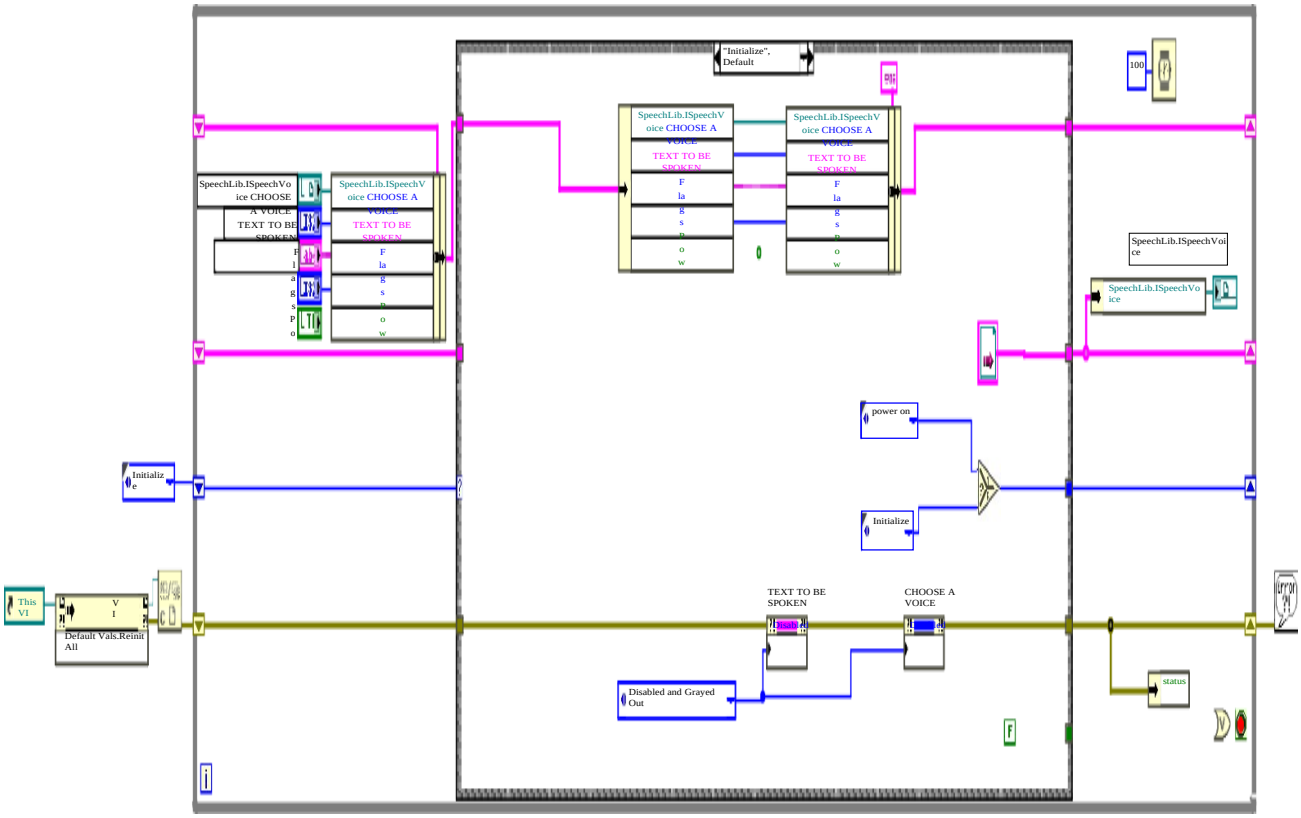

Figure 5 Main block diagram description

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

A. Initialization

In this case, all the controls and indicators are reinitialized to their default values using Invoke node. In order to reduce the numbers of wires Controls Constant Cluster and Indicator constant clusters are created. The voice selection button and the text to be spoken will be in disabled state. When the power button is switched on it will go to the next case else it will wait until the power button is switched on.

B. Power On

In this case, voice selection button and the text to be spoken will be in enabled state and it will wait until you enter the text.

C. Read

In this case, the text to be spoken will be analyzed and it will look for the special characters. Then it will replace all the special characters with its spelling. Now the text to be spoken will not have any special characters, this text will be given to read. Using the invoke node the speech attributes are set. Using the speech library files the voice is set. When the both parameters are available, the application starts to read the text.

D. Stop

When the user press the stop button the application will stop and sets default values to all the objects.Since the application has to run continuously and stop when the user presses the stop button while loop is used. To have a state transition, this application is built using state machine architecture. The Fig. 5 and 6 shows the main block and front panel diagram respectively.

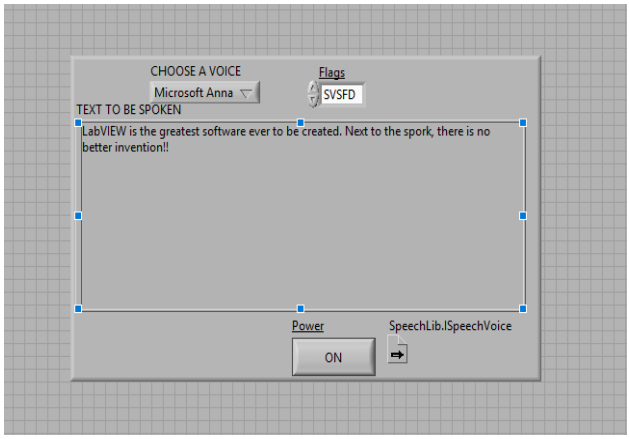

Figure 6 Front panel diagram of Mani VI

IV. CONCLUSION

The LabVIEW instrument permits to execute the Text to Speech change .LabVIEW has its solid inbuilt Speech library to actualize the Text to Speech transformation. Different strategies and procedures are accessible for discourse amalgamation which is talked about in this report. Here, in content to discourse transformation VI ,a content box is made so client can compose content which is to be changed over into discourse in .wav record position and makes a wave document named yield .wav , which can be tune in by utilizing wave document player.

The ACTIVE X sub bed in Communication bed is utilized to trade information between applications. ActiveX innovation gives a standard model to bury application correspondence that distinctive programming dialects can execute on various stages. Microsoft Speech SDK has been utilized to fabricate discourse empowered applications, which recover the voice and sound yield data accessible for PC. This library permits choosing the voice and sound gadget one might want to utilize, enter the content to be perused, and modify the rate and volume of the chose voice.

The application created is easy to understand, savvy and gives the outcome in the constant. Besides, the program has the expected adaptability to be adjusted effectively if the need emerges.

V. FUTURE SCOPE

A discourse blends framework for example a content to discourse change framework is created utilizing the LabVIEW. Still some more work should be possible in this field as referenced underneath:

1. By including some resonation it might be conceivable to build the loveliness of manufactured discourse subsequently
2. Diverse sound wave design records, for example, .mp3 and so on or other configuration required by client, can created utilizing distinctive methods
3. Diverse techniques for right articulation can be utilized to improve better and right
4. Arrangement for controlling the bits/tests, and thus the discourse speed, pitch control and so on can be included as new component.
5. Another module can be included for the voice initiated remote control applications.

REFERENCES

1. Amanpreet Kaur Bains, Nishi Sharma (2008-09) “Automatic Syllable Segmentation for Indian Languages” RIMT-IET,Mandi Gobindgarh, BGIET, Sangrur- 2008-09.
2. Bhuvana Narasimhan, Richard Sproat and George Kiraz (2004) ” Schwa-Deletion in Hindi Text-to-Speech Synthesis” International Journal of Speech Techology 7, 319–333, 2004 Kluwer Academic Publishers
3. Catalin Ungurean, Dragoş Burileanu, Vladimir Popescu and Aurelian Dervis. Hybrid Syllabification and Letter to Phone conversation for TTS Synthesis. U.P.B. Sci. Bull., Series C, Vol. 73, Issue 3, 2011 ISSN 1454-234x.
4. Diemo Schwarz(2000) “A System For Data-Driven Concatenative Sound Synthesis” proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-00), Verona, Italy, December 7-9,2000
5. D.T. Chappell , John H.L. Hansen(2002) “Spectral Smoothing for Speech Segment Concatenation,” Speech Communication, Volume 36, Issues 3-4, March 2002, Pages 343-373.
6. E.Veera Raghavendray, Kishore Prahallad(2008) “A Multilingual Screen Reader In Indian Languages” International Institute of Information Technology - Hyderabad, India. Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA. 7. Hafiz Musa, RabiahA. Kadir, AzreenAzman, M.Taufik Abdullah(2010) “Syllabification Algorithm based on Syllable Rules Matching for Malay Language” Recent Researches in Applied Computer and Applied Computational Science 2010.
7. Kalyan D. Bamane and Kishor N. Honwadkar(2011) “Marathi Speech Synthesized Using Unit selection Algorithm” Computer Engineering and Intelligent Systems ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online). 2011.
8. Massimo Moneglia, TommasoRaso, MaryualêMalvessi-Mittmann, HelianaMello(2006) ”Challenging the perceptual relevance of prosodic breaks in multilingual spontaneous speech corpora: C-ORAL-BRASIL / C-ORAL-ROM” LABLITA, University of Florence, Italy. 2006.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

9. Mohamed Ali Kammoun and Ahmed Ben Hamida (2009) "Unit Selection Algorithm Using Bi-grams Model For Corpus-Based Speech Synthesis" International Journal of Information and Communication Engineering 5:2 2009.
10. M. Sugiyama, J. Murakaml, H.Watanbe "Speech Segmentation and Clustering Based on Speaker Features" ATR Interpreting Telephony Research Laboratories 2-2 Hikaridai, Seika- cho, Soraku-gun, Kyoto 619-02, Japan-1993.
11. Ovidiu Buza, Gavril Toderean, Joszef Domokos, Arpad Zsolt BODO "Building A Text To Speech System for Romanian Through Concatenation" 26-28, Cluj-Napoca. 2009.
12. Philipos C. Loizou, Senior Member, IEEE, and Gibak Kim(2011) "Reasons why Current Speech-Enhancement Algorithms do not Improve Speech Intelligibility and Suggested Solutions" IEEE Transactions On Audio, Speech, And Language Processing, Vol. 19, No. 1, January 2011.
13. S. Sarswathi, R.VishalakshyDesign of Multilingual Speech Synthesis System Intelligent Information Management, 2010, 2, 58-64 ,Scientific Research.
14. Mr. S. D. Shirbahadurkar and Dr. D. S. Bormane. Speech Synthesizer Using Concatenative Synthesis Strategy for Marathi language (Spoken in Maharashtra, India), IJRTE, Vol 2, No. 4, November 2009.

AUTHOR PROFILE

Dr. J. UMA, born in 1983, received the B.E in 2004, M.E in 2009 and Ph.D. degree in 2016 in Electrical Engineering at the University of Bharathiyar, Anna University of Technology and Anna University respectively. Since 2006, I have been working as Associate Professor at the Department of Electrical and Electronics Engineering, in M.Kumarasamy College of Engineering, Karur. She has teaching experience of about 12 years. I am the author and co-author of more than 10 papers. I have received many awards for Teaching and Education Excellence. Few of the outstanding awards are Teaching Awards in Engineering from Staffordshire University, UK and Education Matters and Dr.A.P.J.Abdul Kalam Award for Innovative Research from SEEE in the year 2017. I am well connected with industries like Yokagawa India Private Limited, MAS Solar Systems and Caliber Embedded Technologies. My current research interests are in the areas of Power Electronics and Drives, Artificial Intelligence and Virtual Instrumentation. I am the life member of ISTE and SEEE. I am acting as Research Supervisor under Anna University Chennai.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)