Real Time Machine learning based speech enhancement using multi stage temporal convolutional networks with Time-Frequency attention for hearing aids

File Number : CRG/2023/001203

Submitted By : Dr. Burra Manaswini

Submission Date : 13-Mar-2023

PROPOSAL DETAILS

(CRG/2023/001203)

Dr. Burra Manaswini

manaswini.burra@gmail.com

Associate Professor (Computer Science and Engineering)

Lakireddy Bali Reddy College of Engineering

L.b.reddy nagar, mylavaram, krishna district, Krishna, Andhra pradesh-521230

Technical Details :

Scheme :	Core Research Grant		
Research Area :	Computer Science and Engineering (Engineering Sciences)		
Duration :	36 Months	**Contact No :**	+918897336595
Date of Birth :	08-Aug-1990		
Nationality :	INDIAN	**Total Cost (INR) :**	16,26,832

Is PI from National Laboratory/Research Institution ? No

Project Summary :

Speech enhancement (SE) is an important method for improving speech quality and intelligibility in noisy environments. An effective speech enhancement model depends on precise modelling of the long-range dependencies of noisy speech. Several recent studies have examined ways to enhance speech by capturing the long-term contextual information. For speech enhancement, the time-frequency (T-F) distribution of speech spectral components is also important, but is usually ignored in these studies. The multi-stage learning method is an effective way to integrate various deep-learning modules at the same time. The benefit of multi-stage training is that the optimization target can be iteratively updated stage by stage. In this paper speech enhancement is investigated by multi-stage learning using a multistage structure in which time-frequency attention (TFA) blocks are followed by stacks of squeezed temporal convolutional networks (S-TCN) with exponentially increasing dilation rates. To reinject original information into later stages, a feature fusion (FF) block is inserted at the input of later stages to reduce the possibility of speech information being lost in the early stages. The S-TCN blocks are responsible for temporal sequence modelling task. The time-frequency attention (TFA) is a simple but effective network module that explicitly exploits position information to generate a 2D attention map to characterise the salient T-F distribution of speech by using two branches, time-frame attention and frequency attention in parallel. Extensive experiments have demonstrated that the proposed model consistently improves the performance over existing baselines across two widely used objective metrics such as PESQ and STOI. A significant improvement in system robustness to noise is also shown by our evaluation results using the TFA module.

Objectives :

• To study and analyze the DNN based speech enhancement techniques for hearing impaired people
• To develop robust algorithm in such way that hearing impaired can able to handle stationary noise environment and non-stationary noise environments
• To compare the results of the proposed algorithms with conventional existing speech enhancement algorithms.
• To introduce a multi-stage learning which differs from the single-stage model, where the mapping process is often encapsulated in a single black box. As a result, interpretability assessments are generally poor. A multi-stage structure is used in the proposed model where the present stage receives the estimates from previous stages as input and progressively improves those predictions.
• To introduce TFA which characterises the salient T-F distribution of speech by using two branches, time-frame and frequency channel attentions in parallel. In TFA differential attention weights are assigned for each T-F spectral component and enables the models to focus on where (time frames) and what (frequency-wise channels).
• To introduce squeezed TCNs which are employed to further improve the performance and decrease parameter burden In the TCN used in uses a depth wise dilated convolutions to double the feature dimensions,

Keywords :
speech enhancement, TFA, PESQ

Expected Output and Outcome of the proposal :

High quality and intelligible speech for hearing aids people Development of novel algorithms to improve speech intelligibility and quality for hearing impaired listeners. Implement the proposed algorithms on Raspberry pi.. Publishing SCI journals based on the work. Patents

Suitability of the proposed work in major national initiatives of the Government:

Make in India, Startup India, Digital India

Theme of Proposed Work:

Health, Manufacturing

Collaboration Details for last 5 Years :

S.No.	Name	Type of Collaboration
1	Yannis stylianou Professor University of Crete, Greece Heraklion, Crete, Greece Greece [06-Sep-2017 to 10-Aug-2018]	LEADER- Erasmus Mundus European exchange program

Planned Collaboration for the proposed work with any foreign scientist/ institution ? No

SNo.	CO-PI Details
1	 **SUNNY VANAMBATHINA** sunny.dayal@vitap.ac.in Associate Professor(School of Electronics) **Vellore Institute of Technology (VIT-AP University)** Ainavolu, Amaravati, Guntur District, Andhra Pradesh, ANDHRA PRADESH, GUNTUR D.O.B : 20 Jun, 1986