

# A Deep Learning Approach to Speech Enhancement for Hearing Aids

Dr. Nagaraju C <sup>1</sup>, Deepthi N <sup>2</sup>

<sup>1</sup>Assistant Professor, Dept. of ECE, The National Institute of Engineering, Karnataka, India

<sup>2</sup>PG Student, Dept. of ECE, The National Institute of Engineering, Karnataka, India

\*\*\*

**Abstract** - Speech is the main source of communication in humans. Understanding speech in noisy environments is one of the major challenges in impaired people. Deep neural networks have been useful in solving standard problems in several fields including speech. They have been used to improve the speech quality for hearing impaired people. The main objective is to intensify the speech signals in terms of its nature and comprehensibility that have been manipulated by unwanted noise. This paper makes use of convolution neural networks for the enhancement of speech signals manipulated by real world noise.

**Key Words:** *Speech Enhancement, Deep learning, Convolutional neural networks, Denoising .*

## 1. Introduction

Daily life is full of sound, and people with hearing loss are facing issues in one or other tasks. This is because their ears have become too sensitive that they can hear low-pitched vowel sounds easily and hearing is poorest in high pitched range of speech. Hearing aids are small electronic devices that are given to such people in order to improve the quality of communication in noisy environment. Speech Enhancement means regeneration of original speech that is being manipulated by the real world distractions. Using Machine learning much efforts have been made in the area of speech processing to develop speech enhancement techniques so as to bring back the speech signal by reducing the percentage of unwanted noises. The main problem with any speech processing research work is the environmental noise. The main source of human interaction is speech that is corrupted by the background noise. So the main motto of the speech enhancement is to enhance the speech quality and to recognize the original speech which is been mixed with unwanted sound. Speech Enhancement has numerous applications such as mobile phone, VoIP, teleconferencing systems, hearing aids, speech recognition etc., Spectral subtraction, Minimum mean square error (MMSE), Log MMSE, Optimally modified log spectral amplitude (OM-LSA), Wiener filtering, and other voice improvement techniques can be learned. For the enhancement of speech, we propose the concept of deep. Deep learning is the subset of machine learning, related with set of algorithms concerned with structure & function of artificial neural networks.

## 1.1 Paper Organization

The rest of this paper is organized as follows. In Section 2, problem statement is explained. In section 3 survey of various papers is discussed. Section 4 describes the deep neural networks. Section 5 illustrates the proposed system and in the last part we have results section.

## 2. Problem Statement

For the hearing aid listeners, ears have become too much sensitive that they can hear only lower frequencies speech signals and high frequency speech signals sounds like a mime. The nature and comprehensibility of speech signals are manipulated by the environmental noise. Therefore, the manipulated speech signals needs to be amplified to improve nature and comprehensibility.

## 3. Related works

Surveys conducted in the area of speech processing is as explained in this section.[1]Gyuseok Park et al. explains speech enhancement algorithm to enhance the quality of speech in hearing aid environment by applying noise reduction algorithms with deep neural network learning based on noise classification. The environmental noises were classified by convolution neural network and noise reduction using DNN's were applied on the classified noise. This results in the improved speech quality over the conventional hearing aid algorithm and also objective evaluation was made using various methods like PESQ, STOI, OQCM, and LLR scores. The system proposed gives the best quality of speech in various and irregular noisy environments of hearing aid user. Increase in the noise classification using CNN's improved the performance of noise reduction using DNN's and so this enhances the quality of speech in the hearing aids thereby it gives the increased level of listening satisfaction for the hearing aid listeners.

[2]Gautam Shreedhar Bhat et al. proposes an AutoML based VAD model for hearing aid listeners on smart phone device. This paper describes that the developed VAD can be used to improve the performance of speech processing applications related to hearing aid devices. The AutoML model obtained is computationally fast & has minimal processing delay operates efficiently in real time mode on a smart phone. This paper highlights the usage of AutoML for hearing aid applications and awareness of AutoML on smartphone.

AutoML model related to audio provides highly optimized model based on the network architecture and hyperparameters. The AutoML framework has provided low latency and considered to be efficient architecture in terms of its performance. This proposed model can be an important module to improve the speech processing. Also this paper noted that this AutoML plus smart phone model gives an advantageous & compact platform that can be used by others without any guidance.

[3]Abhishek Sehgal et al. presents a smart phone application that carries out real-time voice activity detection hinge on convolutional neural network. The developed application acts as a switch for noise reduction in the speech processing areas that the noise classification will be conducted only in noise only regions. This paper states that the experimental results shows developed app using convolutional network performs extensive well with previously developed voice detection app and also cited by two popular voice detection algorithms. The smartphone has been created for both android and iOS phones. The convolution neural network architecture was optimized, allowing the processing of audio frames in real time. This is done to ensure that no frames are being skipped and results in higher accuracy in voice detection. The app makes use of the concept of multithreading that runs in parallel to actual audio stream and hence providing computationally efficient framework for working with the speech processing modules in real time. The results generated gives an insight that developed app formed on convolution neural network performs well compared to previously developed which was built on random forest.

[4]Soha A Nossier et al. idea is to develop a smart hearing aid which has the ability to detect noise making the speech hear able. This paper proposes three system designs applying the idea of smart hearing aids. Smart hearing aids means hearing aids having the ability of detecting noise like fire alarm, car horn and making the speech hear able thereby avoiding the unwanted sound heard while actually listening to actual speech during the communication. Speech enhancement approach uses deep neural network to classify the noise and for discarding the unwanted noise. The three system designs vary with its complexity, performance and flexibility. The first system is form on only one FNN and of the least complexity and its is having the negative impact on the speech enhancement process. The third system is said to be of highest complexity because it has 2 FNN. Thus, second and third system proposed performs better than the first system as the speech enhancement and noise classification processes were done in two separate independent networks, this will not affect the quality of the speech and considered more flexible than the first system proposed.

[5]Pallavi D Bhamre and Hemangi H Kulkarni describes speech enhancement is a process of improving the quality and intelligibility of speech in pure environment at low sound to noise ratio(SNR).This paper suggests the speech

enhancement work by using the supervised method deep neural network(DNN),large training set is required for the mapping between noise and actual speech features.

Three strategies worn with the DNN has given the good results. The suggested post-processing,Globalization variance equalization was given in the speech enhancement process in terms of brightening the formants spectra and training methods enhances the performance by lowering the amount of unwanted sound. Thus, DNN based proposed system gives better results when compared with prototypical speech enhancement method called MSME.

#### 4. Deep Neural Networks

Deep neural networks are the subset of Machine learning. They are the artificial neural networks that are based on structure and function of biological neural networks. Deep neural networks are multiple artificial neural networks with multiple layers of networks, they are modern update of ANN's and builds larger and more complex network which accepts very large datasets as input. DNN's consists of 3 layers: input layer, hidden layer & output layer. Input layer accepts input signal, Output layer receives the response.

#### 5. Methodology

The proposed algorithm reduces the noise by classifying its environment. Below figure represent the process flow.

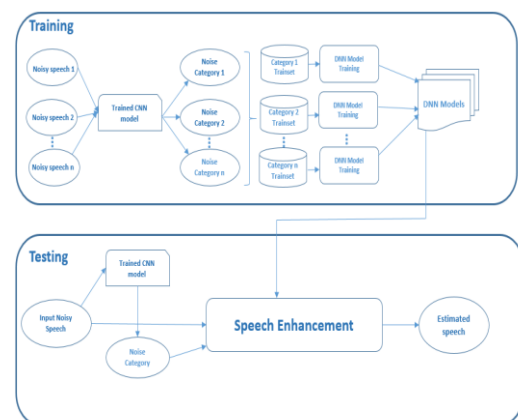


Fig. Proposed System design

All the Machine learning workflow follows two steps training the model to preform required task and testing the model with new data to test to validate model's efficiency. Training:-

Initially, input noisy speech could be labeled primarily based on its surroundings using Convolution neural community (CNN). CNNs are a type of deep mastering techniques and are broadly used for picture category even as maintaining the spatial information of the picture. for you to improve the

classification frequency of the environmental noise, a spectrogram image of the noise turned into used to convert sound indicators inside the time frequency-area into picture alerts. further, a sharpening masks and median filter have been implemented to enhance the noise classification charge.

The characteristic function units of noisy speech had been extracted to generate input statistics for the deep neural network (DNN). enter functions will consist of mel-frequency, cepstral coefficients (MFCCs), amplitude modulation spectrogram (AMS), relative spectral converted perceptual linear prediction coefficients (RASTA-PLP), and sixty four gamma-tone filterbanks (GFs). For hearing aids operation, the DNNs could be educated and saved for every type of environmental noise. The storage factors of the trained model have been the schooling configuration, together with the architecture of the model, the weight of each node in the layers, the value characteristic, and the optimization technique.

Testing:-

In testing phase, new noisy input will be first classified to get noise environment. Based on the classified noise environment, training configuration of respective noise category will be loaded to perform the speech enhancement process.

### Results and Discussion



Fig. Speech Enhancement Software

The above figure shows the user interface of the proposed system. This software is designed in Pycharm IDE using Tkinter toolkit. It has 3 buttons: (i) Load .wav file – this is for loading input noisy audio files. (ii) Play – it is used to play the noisy audio. Spectrogram of the noisy audio with SNR value will be displayed in the software. The type of noise is also shown. (iii) The third button is CNN – The proposed algorithm is implemented here. When user clicks on this button, denoised audio file is played. Spectrogram of the denoise audio signal is displayed, decrease in SNR values

shows that certain amount noise has been eliminated from input. Also, the software detects the type of noise heard.

### CONCLUSION

Unlike any other speech enhancement algorithm, proposed system considers a environmental noise in which hearing aid operates and improves the speech quality. The increased noise classification rate using CNNs improves the noise reduction using DNNs, and through the proposed algorithm, the quality of speech of the hearing aid improves, leads to increase in the listening satisfaction for the people with impaired hearing.

### ACKNOWLEDGEMENT

I am using this opportunity to express gratitude to my guide and all the members for their help and support. I render my sincere gratitude to my family for their support.

### REFERENCES

- [1] Park, Gyuseok, Woohyeong Cho, Kyu-Sung Kim, and Sangmin Lee. "Speech Enhancement for Hearing Aids with Deep Learning on Environmental Noises." *Applied Sciences* 10, no. 17 (2020): 6077.
- [2] Bhat, Gautam Shreedhar, Nikhil Shankar, and Issa MS Panahi. "Automated machine learning based speech classification for hearing aid applications and its real-time implementation on smartphone." In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 956-959. IEEE, 2020.
- [3] A Convolutional Neural Network Smartphone App for Real-Time Voice Activity Detection T2 - IEEE Access SP - 9017 EP - 9026 AU - A. Sehgal AU - N. Kehtarnavaz PY - 2018 DO - 10.1109/ACCESS.2018.2800728 JO - IEEE Access IS - SN - 2169-3536 VO - 6 VL - 6 JA - IEEE Access Y1 - 2018 .
- [4] Soha A. Nossier, M.R.M. Rizk, Nancy Daa Moussa, Saleh el Shehaby, Enhanced smart hearing aid using deep neural networks, Alexandria Engineering Journal, Volume 58, Issue 2, 2019, Pages 539-550, ISSN 1110-0168.
- [5] Bhamre, Pallavi D., and Hemangi H. Kulkarni. "Speech enhancement using deep neural network." *Int. Res. J. Eng. Technol.* 3, no. 7 (2016): 1348-1354.
- [6] Rim Park, Se, and Jinwon Lee. "A Fully Convolutional Neural Network for Speech Enhancement." *arXiv* (2016): arXiv-1609.

[7] Nassif, Ali Bou, Ismail Shahin, Imtinan Attili, Mohammad Azzeh, and Khaled Shaalan. "Speech recognition using deep neural networks: A systematic review." *IEEE Access* 7 (2019): 19143-19165.

[8] Liu, Weibo, Zidong Wang, Xiaohui Liu, Nianyin Zeng, Yurong Liu, and Fuad E. Alsaadi. "A survey of deep neural network architectures and their applications." *Neurocomputing* 234 (2017): 11-26.

[9] Liu, Weibo, Zidong Wang, Xiaohui Liu, Nianyin Zeng, Yurong Liu, and Fuad E. Alsaadi. "A survey of deep neural network architectures and their applications." *Neurocomputing* 234 (2017): 11-26.

[10] Purushotham, U., K. S. Chethan, S. Manasa, and U. Meghana. "Speech Enhancement Using Semi-Supervised Learning." In *2020 International Conference on Intelligent Engineering and Management (ICIEM)*, pp. 381-385. IEEE, 2020.

[11] Khaleelur Rahiman, P. F., V. S. Jayanthi, and A. N. Jayanthi. "Speech enhancement method using deep learning approach for hearing-impaired listeners." *Health Informatics Journal* (2020): 1460458219893850.

[12] Goehring, Tobias, Federico Bolner, Jessica JM Monaghan, Bas Van Dijk, Andrzej Zarowski, and Stefan Bleeck. "Speech enhancement based on neural networks improves speech intelligibility in noise for cochlear implant users." *Hearing research* 344 (2017): 183-194.