Voice Isolation Using Artificial Neural Network

Akshay Khanna[1], Aaftab[2], Agam Jain[3], Akash Verma[4]

[1,2,3,4] *Rajkumar Goel Institute of Technology/AKTU*

Abstract - **We often encounter noises when sounds from different sources in the world mix in the air before arriving to our ear, requiring the brain to estimate individual sources from the received mixture. The problem can be solved by the help of Artificial Neural Networks (ANN) by identifying regularities in natural sound sources. In this project of voice isolation system using ANN we will be using various algorithms like independent component analysis, gradient descent and activation function and the platform used will be Python.**

Index Terms - **Artificial Intelligence, Machine Learning, Independent Component Analysis, Artificial Neural Network and ICA.**

I.INTRODUCTION

We often encounter noises when sounds from different sources in the world mix in the air before arriving to our ear, requiring the brain to estimate individual sources from the received mixture.[1] But how do we recognize what one person is saying when others are speaking at the same time? And how can we program a computer to do that? The problem can be solved by the help of Artificial Neural Networks (ANN) by identifying regularities in natural sound sources.

In this project of voice isolation system using ANN we will be using various algorithms like independent component analysis, gradient descent and activation function and the platform used will be Python.

A. PROJECT OBJECTIVE

People face problem to hear to certain noises when they are at public place or at a party which may be of great importance for them. The challenges that we need to overcome during this process arises when the scenario becomes larger and many more sound sources arises, and due to this it becomes really difficult to determine. Our main goal is to isolate with efficiency the mixed signals by the process of artificial neural network and overcome the 2 most faced challenges.

1. To analyze the mixture of sounds and pick out the one which we have to understand.

2. To direct one's attention to the sound he wants to listen and ignoring the rest.

II.LITERATURE REVIEW

This phenomenon was referred to be the ability to select and isolate one source of audio in a noisy environment from the others so that it can be listened to efficiently. To solve this problem numerous efforts have been made in many fields be it neurobiology, physiology, computer science or engineering. Thus its nature can be considered to be multidisciplinary, as per the survey done in the year 2002 on the title speech separation by simulating the efforts with the neural network it was clear that for the processing of audio signals, they used LPC analysis. This was successfully used for the separation of speech when the interface has broadband noises. There the neural network used was modified in its hidden layer using the extracted data. This data was extracted using LPC algorithm. But the drawback of continuing this approach was the lamination of the samples of audio signals that were supposed to be train. That is why we are working upon this particular paper so that we can exceed its limit of training samples to infinity.

Another publisher Josh McDermott in his article wrote how we can recover different sound sources from repetitions and different mixed signals. They tried to do this with the help of naturalistic properties but this too came with a drawback of how prior knowledge of different sounds were required before isolating them. Thus for this reason we started working upon the unsupervised learning where prior knowledge of the data is not required.[1]

Later on we studied how or in what ways we can isolate the noises using more efficient algorithms under "artificial neural network". ANN is a technology that was inspired by biological neural network. It is the platform where we can process all the data inputs and make several algorithms work

together. Firstly we look into the algorithms we are going to work.

Thus different algorithms are being applied in the project. They are isolating and optimizing our signals based upon their amplitude and frequency. We give a short glace on algorithm which working. ICA that is the analysis we are doing on each component using its frequency and amplitude so as to separate them.

Activation function gives a desired range to the signals being separated. Many other algorithms simultaneously working to give the desired result.

According to IEEE, the neural network is used as a speaker recognition system to control the iterative filter. The neural network is a modified perceptron with a hidden layer using feature data extracted from LPC analysis. The proposed technique has been successfully used for speech separation.[2]

S.NO.	TITLE	YEAR	ALGORITHM	DISADVANTAGES
1.	Tracing and recognising the speech of a specific speaker.	2015	Automatic Speech Recognition.	Our demonstrated design does not reflect the mechanism of the listening of human auditory systems effectively.
2.	Speech separation by simulating the effect with a neural network.	2002	LPC Analysis used in audio signal processing.	It has the limitation of speech samples used for training the system.
3.	Recovering sound sources from embedded repetition	2011	Generative model to synthesize novel sounds with naturalistic properties	Prior knowledge of different sounds is required.

Table 1 Inferences from Literature [3]

III.PROPOSED WORK

Voice isolation System is the separation of mixed signals into two separate components using artificial neural network. After studying various algorithms of ANN we can achieve our goal. So, we make use of the Independent Component Analysis algorithm which helps us to analyze the data and understand the fundamental features of the data, its consistencies and that will help us to isolate the required sound from the mixture.

IV. METHODLOGY

- We are given m sources (conversations for example), and some number of sound receivers, separate out the different signals.
- At every (uniformly spaced) discrete interval of time we record m samples, one at each of our m microphones.
- We simply record the amplitude of the sound at each instant and then we will compare the sounds

pitch with the given frequency and then apply activation function over them.

- We will train the data with every executing epoch to attain the best parameters using ICA algorithm and mathematical calculations.
- Thus at the end we use these trained parameters to isolate voice from noise mixture.
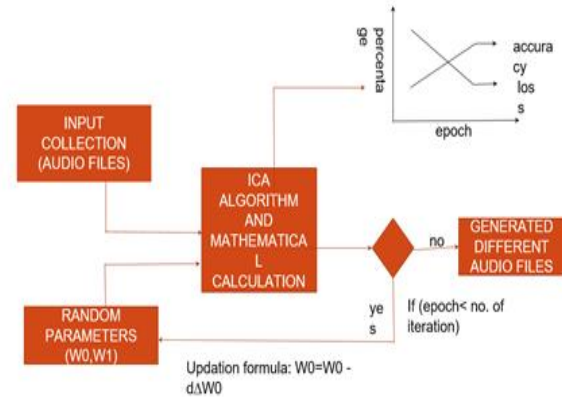

Figure 1 Workflow Diagram

V.CONCLUSION

The main objective of our project is to separate mixture of two voices or signals recorded via two microphones and two speakers into two individual independent speeches. There are two parts of the project – the speech recording part and the computational part. Former part consists of the audio mixtures samples, while latter, the computational part, consists of the algorithms and the computer programs that are used to separate the two voices and provide them as separate outputs. By implementing this project, we will be able to listen and understand one's voice properly without any noise in it and can easily understand what a person was saying. This project can also be used to isolate the sound of musical instruments and the vocals part of the singer and provide them as two different audio files.

REFERENCES

[1] https://www.pnas.org/content/115/14/E3313
[2] Recovering sound sources from embedded repetition-McDermott JH, Wrobleski D, Oxenham AJ
[3] https://ieeexplore.ieee.org/document/595489
[4] https://www.sciencedirect.com/topics/computer-science/speakeridentification