ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Speech Denoising through Deep Learning

Dr Baburao Markapudi[1] , Prof CSE, SR Gudlavalleru Engineering College, baburaompd@gmail.com
Sreeja Pandu[2], SR Gudlavalleru Engineering College, sreejapandu2408@gmail.com
Haritha Pamarthi[3], SR Gudlavalleru Engineering College, harithapamarthi180@gmail.com
Deeksha Sindhu Marrapu[4],SR Gudlavalleru Engineering College, dekshasindhu@gmail.com
Pardha Saradhi Mogili[5],SR Gudlavalleru Engineering College, pardhasaradhi2023@gmail.com
Eswar Datta Narsipally[6], SR Gudlavalleru Engineering College, eswardattanarsipalli2@gmail.com

Abstract:
In speech communication, quality of speech plays a major role to maintain the accuracy of information exchange. To maintain a noise-free environment during communication many speech processing systems are invented However, in a practical situation, the presence of background interference in the form of noise and cumulative background noise abruptly lowers the effectiveness of these devices, resulting in less effective communication and listening strain. To overcome this, many speech stimulation approaches have been introduced like the time domain approach, statistical-based approaches, and transform domain approaches. Here, finally a generalized CNN Single Subspace method for the stimuli of colored noise-corrupted speech is provided. A non-unitary transform based on the simultaneous linearization of the clean speech and noise covariance matrices is used to project the corrupted signal onto a signal-plus-noise subspace and a noise subspace. To evaluate the clear signal, the parts of the signal subspace and the signal parts in the noise subspace are kept. Due to the imposed transform's integrated pre-whitening, it can be utilized for colored sound in common.

Keywords:
CNN signal subspace, Speech Communication, Speech enhancement, deep learning.

Introduction:
Speech Stimulation is the process of removing noises from a speech neither degrading the speech's purity. Due to the current existence of this audio degradation, de-noising has a very important role in increasing human-to-human (like amplifying devices) and human-to-machine (like automatic speech recognition) communications. This is a challenging but common form of the problem in the unknown case of single-channel speech de-noising, due to the complexity of speech processes and the unknown nature of the non-speech material. The difficulty is further compounded by the nature of the data since audio material contains a high density of data samples. Challenges also arise in mediated human-to-human communication, as perception mechanisms can make minimal errors still noticeable to the average user. Here, a comprehensive deep learning method for audio de-noising is utilized. This approach trains the CNN algorithm with clean and noisy audio files and generates a training model. The raw audio file will be subjected to the CNN train model to find and eliminate noise. All noise data from the audio file will be predicted by the CNN algorithm and then replace that noise data with dummy values to remove noise. While removing noise CNN will choose audio sub-signals and then predict noise from that sub-signals. After removing all noise CNN will generate two files called 'noise.wav' contains all noise data and 'clean.wav' contain all clean audio data. Both files will be saved inside the 'Speech Denoise' folder. Here in this project large experiments that compare the presented approach to the recent idea of end-to-end deep learning techniques for de-noising. Our approach transcends them in objective speech quality metrics and big-scale perceptual experiments with human listeners, indicating that our approach is more effective than the baselines. The advantage of this

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

approach is particularly pronounced for the hardest, noisiest inputs, for which de-noising is most challenging.

Related Work:
(2014) A study entitled "An Investigational Study on Speech Stimuli Based on Deep Neural Networks" was published by Yong Xu, Jun Du, Li-Rong Dai, and Chin-Hui Lee. In essence, neural networks are mathematical models that can be used to solve specific issues. Here, a regression-based framework for speech improvement that makes use of deep neural networks (DNNs) with several deep layers. A large training set is necessary for the DNN learning process to build a powerful model that can calculate the intricate nonlinear mapping between the real noisy speech and the desired clean signals.[1]

Using temporal analysis, Shenoy, R.R., Seelamantula, C.S. (2014) predicted frequencies linearly. In order to characterize the contexts of transitory impulses, such as voiced and unvoiced stops, plosives, and other sounds, frequency-domain linear prediction (FDLP) is commonly used in speech coding.[2]

Simon, Timo Gerkmann, Toon van Waterschoot, Ante Juki'c, Simon Doclo1, Timo Gerkmann1, Toon van Waterschoot 2, (2014) Speech Dereverberation Using A Targeted Signal With Sparse Priors and Multi-Channel Linear Prediction. This claims that room reverberation can have a significant impact on the quality of voice signals that are recorded. In this study, concentrate on the highest technique-based blind procedure for voice dereverberation that estimates the model's parameters in the short-time Fourier domain.[3]
For efficient autonomous speech recognition, M Segbroeck, K. Audhkhasi and S. S. Narayanan (2014) incorporated a number of de-noising algorithms. Accordingly a system for combining different de-noising front-ends for effective speech enhancement for identification in noisy contexts.[4]

Hermansky, H.; Peddinti, V. (2013) Optimization of the filter bank for the frequency domain Portions of the semi Frequency are extracted from a voice signal using the discrete cosine transform (DCT). Frequency Domain Linear Prediction (FDLP) approach, It calculates autoregressive models for subband signals' Hilbert envelopes. The window's size and placement in the cosine transform of a signal impact the implicit filtering of the signal. Using Amro's signal subspace technique for speech stimulation, High-rate linear predictive coding for code-excited compression (2013). This article recommends using a linear prediction frame parameter driven by a Hamming Correction Code Compression (HCDC) code. Some of these options include gain, excitation bits, and linear prediction coefficients. The remaining 40 coefficients in the DCT for the signal frame are individually quantized with four bits.[5]

With a transmission rate of 5.22 Kbps, the signals employed in the trials had a total of 261 bits per frame, Meng Guo, J. Jensen, S.H. Jensen, and T.B. Elmedyb (2010) Examination of dynamic echo and feedback cancellation techniques in a typical several-microphone and single-loud speaker system. It explains that in this study, a standard system with many microphones and one loudspeaker was addressed using an optimization technique to suppress audible response and a pulse maker to process the data.[6]

Scordilis, M.S.; Lu Han; Xin Liu (2009) Harmonic Emphasis and Adaptive Comb Filtering for Speech Enhancement. To reduce frequency domain noise, constraint optimization is used to generate the spectral weighting function. Gain and the frequency-dependent noise reduction parameter (FDNFP) are two design parameters that affect rejection gain.[7]

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Research Methodology

CNN for Speech Enhancement:

The suggested convolutional neural network produces a clear voice (CNN). A matrix or a picture is used as the input for a CNN, and its hidden layers process the input by conducting convolution and pooling operations. Finally, they are connected with fully connected layers in a manner akin to conventional DNNs that employ backpropagation techniques before inferring the conclusion. Typically, a convolutional layer in CNNs is formed by several weighted learnable filters, or kernels, that may convolute a tiny portion of an image. Over the whole input space, these kernels are repeated. After each forward pass, each kernel generates a feature map that lists the fundamental characteristics of an image. The max-pooling layer receives these feature maps in order to either increase dimensionality or reduce resolution. Depending on the application, the max pooled output can also be connected to convolutional layers, which enable the network to learn intricate details about a picture. The output of convolutional or max pool layers is flattened before being joined to fully connected layers. A non-linear output layer is connected to the fully connected layer in order to generate the output for classification or regression. When employed for speech enhancement, the network as a whole remains constant. Instead of treating the output as a classification problem, the features are treated as a regression problem.

Here, the network picks up on important time-frequency audio characteristics, such formats. The CNN is invariant to translational variance since the same kernel processes each component of the feature map. This makes it possible for the connect to learn and understand speech from various languages and genders, which typically have distinct fundamental frequencies due to differences in pitch.

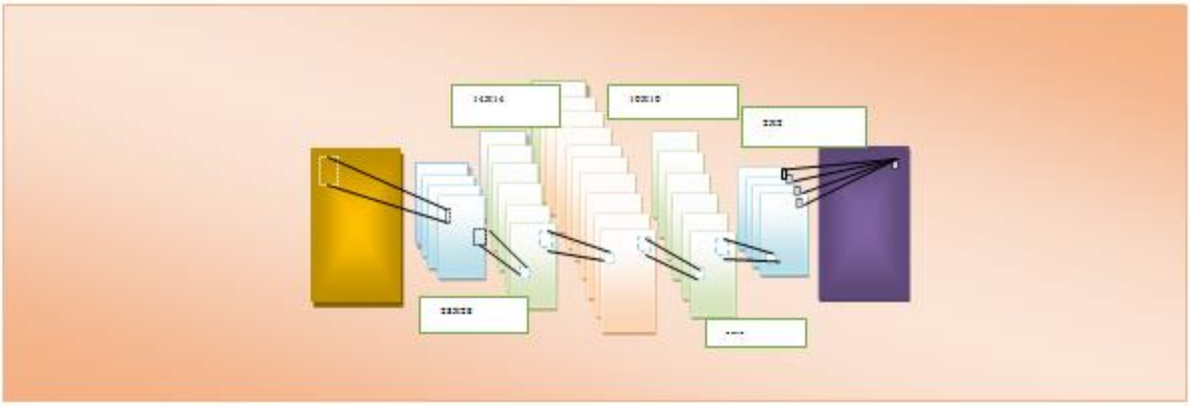

Figure 1 Architecture of A Fully CNN for Speech Enhancement

Combining LPS and MFSC features, the suggested CNN structure shown in Fig. 4 uses these features as the network's input. N/2+1 LPS and L MFSC spotlights are included in the input vector for each time frame, and they are added to form the DT picture, where D=(N/2+1+L) and T is the overall number of frames used to calculate the spectrum. T represents nine consecutive noisy frames while considering multiple frame expansion; 8 frames before this one + this one. The current clean frame is the target vector. T is typically regarded as a big number for recording temporal information. T is set to 9 frames to reduce calculational difficulty, particularly for real-time implementation. A fully linked layer and convolutional layers make up the network. The local patterns in the input image are extracted by the convolutional kernels in terms of both time and frequency.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022**

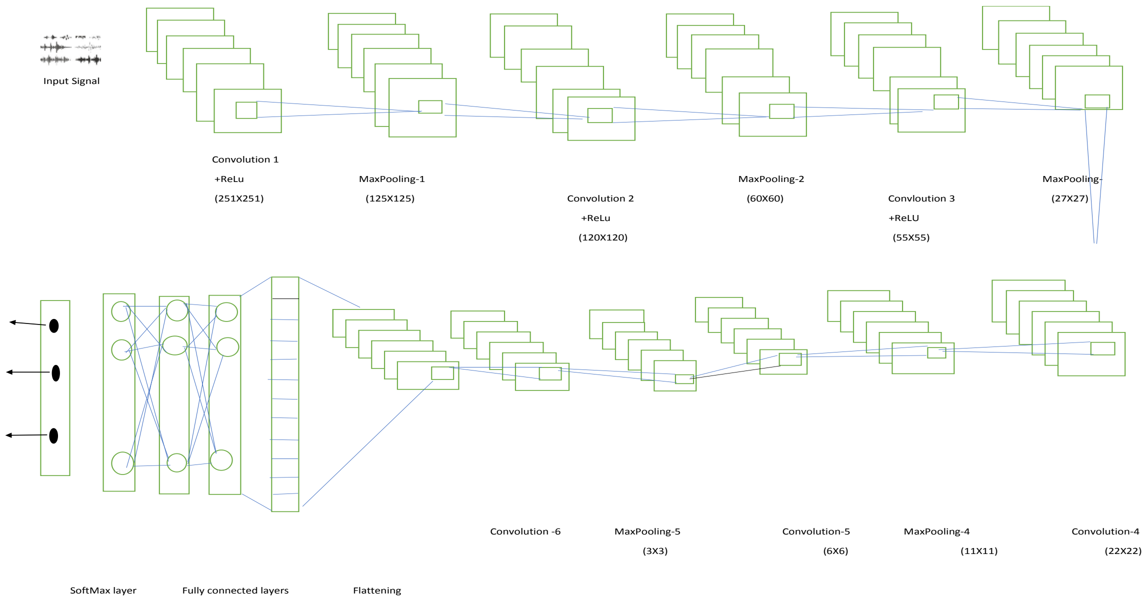

Figure 2 Block diagram of proposed CNN Classification

Subspace Smoothing:

a) Essentials: Basically, assumptions regarding the characteristics of the interfering noise signal are needed for noise reduction techniques. Not all signal amplification methods use subspace-based voice stimulation, although in many situations they do since it also makes some basic assumptions about the target signal's characteristics (pure speech). Speech and noise signals are distinguished based on many features. The characteristics of speech (noise) signals fluctuate over time, and nested frames of analysis are used during the speech stimulation process.

All subspace-based signal improvement algorithms work under the premise that each short speech vector has the form p = [p(1), p(2),...,p(b)].

T is a linear combination of basis functions that can be written as mi, I = 1,..., a, that are linearly independent of one another.

$$p= M\ y, (1)$$

where M is A (b × a) matrix with the fundamental operations listed in columns, and y is a column vector of scales or weights. In general, these fundamental functions' number and shape will change over time (frame dependent).

A logical option for , inspired by the conventional sinusoidal model (SM) for audio signals are (damped) sinusoids. The fact that the following speech vectors will inhabit an a (a<b)-dimensional subspace of the b-dimensional Euclidean space is an important finding in this case (a equals the signal order). Due to the fact that speech signals change over time, the position of this signal subspace (as well as its scale) will vary depending on the chassis.

b) Audio Signal: The extra noise is supposed to be white, zero-mean, and unrelated to the speech signal. Its range should fluctuate gradually over time, allowing measurement from noise-free regions. In contrast to the voice signal, successive noise vectors n may fill the entire b-dimensional space.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

c) Speech and Noise Separation: In accordance with the previous descriptions of the noise and speech signals, the previously stated, b-dimensional monitoring space is divided into two subspaces: an a-dimensional (noise + signal) subspace where the speech signal is interfered with, and a (b-a)-dimensional subspace where the noise is the only thing present (and no speech). The process for improving speech can now be summed up as follows: Eliminate the "noise only" subspace, separate the "signal & noise" subspaces from it, and, if desired, eliminate the noise elements in the "signal & noise" subspace as well. 1 For the white noise scenario the considered way here is, the first operation is straightforward, but further in future if it is in depth, it can be difficult in the case of colored noise. All implementations of use the second operation. To obtain subspace-based signal improvement and better noise suppression, a third procedure is necessary. Due to the advent of voice distortion, the latter operation is occasionally skipped. Due to the noise and speech signals proximity in the signal subspace, the latter issue cannot be avoided. The next sections show how the noise signal's SVD observation matrix, or alternatively, its eigenvalue decomposition (EVD) of the correlation matrix, can be used to conduct orthogonal split into frame-dependent signals and noisy subspaces.

d) Algorithm: Here, in this s(k) represents a pure negative sample and n(k) zero mean added white noise distortion not expected to correlate with pure speech. Then the observed noisy negative x(k) is given by

$$x(k) = p(k) + n(k), \quad (2)$$

Further, let $\bar{R}$x, $\bar{R}$p, and $\bar{R}$n be (a × b) (b>a) true auto-correlation matrices of x(k), p(k), and n(k), respectively. It is clear that the presumption that speech and noise are not connected

$$\bar{R}x = \bar{R}p + \bar{R}n, \quad (3)$$

The EVD of $\bar{R}$p, $\bar{R}$n, and $\bar{R}$x can be written as follows:

$$\bar{R}p = \bar{V}\bar{\Lambda}\bar{V}T, \quad (4)$$

$$\bar{R}n = \overline{V}((\sigma_\omega^2 I) \ \bar{V}T, \quad (5)$$

$$\bar{R}x = \bar{V}\bar{\Lambda} + (\sigma_\omega^2 I) \ \bar{V}T, \quad (6)$$

with I the identity matrix, the noise variation, σ_ω^2 the diagonal matrix holding the eigenvalues of $\bar{\lambda}$i, $\bar{V}$ the orthonormal matrix containing eigenvectors of $\bar{V}i$ and the eigenvalues of I in it. The crucial discovery that the eigenvectors of the noises are identical to the clean-speech eigenvectors is made possible by the white noise hypothesis, enabling the eigenvectors of R¯p to be deduced from the EVD of $\bar{R}$x in (6), on the grounds that the pure speech is believed to be limited to a (a<b)-dimensional subspace(1), the fact that $\bar{R}$p has only a non-zero eigen values $\bar{\lambda}$i.

$$\bar{\lambda}i > (\sigma_\omega^2 \ (i = 1, \ldots, a), \quad (7)$$

If the voice signal and noise can be distinguished, and the EVD of $\bar{R}$x should be expressed as:

$$\bar{R}x = [\bar{V}a \ \ \bar{V}b{-}a] \left(\begin{bmatrix} \bar{\Lambda}a & 0 \\ 0 & 0 \end{bmatrix} \right) + \left(\sigma_\omega^2 \begin{bmatrix} Ia & 0 \\ 0 & Ib-a \end{bmatrix} \right) [[\ \bar{V}a \ \bar{V}b{-}a]\hat{}T] \ , \quad (8)$$

Suspect that the elements $\bar{\lambda}i$ of $\bar{\Lambda}$ are in descending order. The signal and noise subspaces are identified by the subscripts a and b-a, respectively.

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Speech stimuli is now achieved via, regardless of the particular optimization criterion,

(1) limiting the improved speech by nullifying its parts in the noisy subspace so that it only occupies the signal subspace,

(2) modifying the associated signal subspace eigenvalues (i.e. decreasing them).

Mathematically this stimuli process could be represented as a smoothing modifications on the noise speech vector x = [x(1), x(2), … , x(b)]T:

$$\hat{p} = Fx , (9)$$

the screen matrix F should be given as

$$F = \bar{V}a \ Ga \ \overline{(V}\hat{} \ T) *a, (10)$$

The weighting factors gi for the first 'a' eigenvalues of R ̅x are contained in the (a x b) diagonal matrix Ga, where $\bar{V}$ and $\bar{V}$ ^T are referred to as the inverse of the KLT (Karhunen Loeve transform) matrix, respectively. The filter matrix F's alternate is

$$F = \sum_{i=1}^{a} g_i \ \bar{v}i \ \bar{v}T \ i ,(11)$$

It demonstrates how a "filter bank's" outputs might be conceived of as accumulating into the

filtered signal (see below). This filter bank's filters are all depend on just one eigenvector v ̅i and its equivalent gain factor g_i.

e) EVD to SVD Smoothing: The real covariance matrices in (4) to (6) are frequently approximated Rx = HT x Hx, with Hx (= Hs + Hn) an (m × b) (with m>b) a noisy speech vector x, and a noisy Hankel (or Toeplitz)2 signal monitoring matrix.

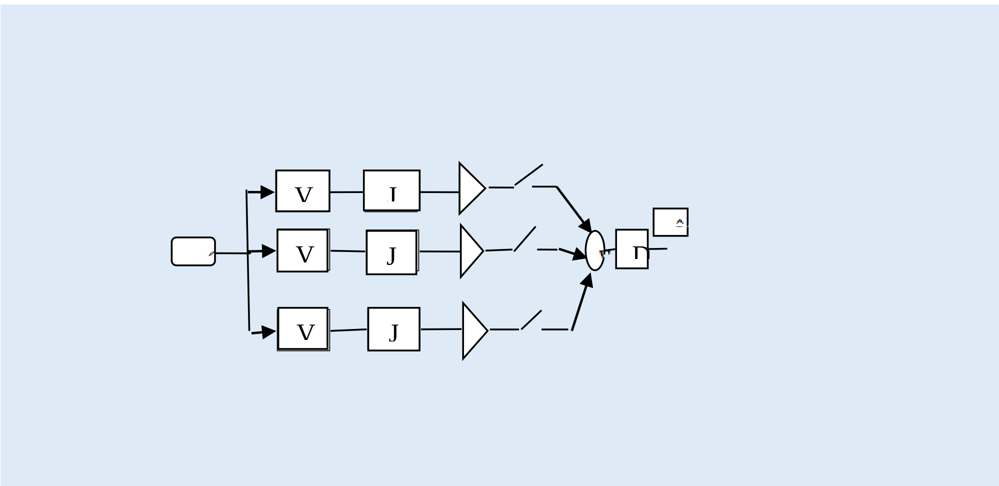

Figure 3 EVD to SVD Smoothing

Consisting of N (N >>b, and m + b = N + 1 specimens of x (k). In that instance, the SVD of can be used to produce an analogous speech improvement Hx [6]. Here is an example of a standard modified SVD-based voice enhancement procedure. Let Hx = UΣVT be the formula for the SVD of Hx. (12) Short-time speech and short-time noise signals that are orthogonal (HT p Hn = 0) have a white short-time noise signal.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

$$H_x = \text{U}\left(\sqrt{\bar{\Sigma}\hat{}2 + \sigma_v^2 I}\right)V^T$$

withΣ¯the matrix that holds the clean Hankel matrix's singular values Hp, and σv the 2-norm of the columns of Hn (notice that in the case of fixed white and for big N noise), in the mean square sense, σ2 ν /m converges to σ2 w). The evidence based covariance matrix HT x Hx / N will inevitably reach the true auto-correlation matrix $\bar{R}$x under weak conditions. To put it another way, if N is large enough, the subspace extended by the p dominating eigenvectors of V will converge to the subspace extended by the vectors of $\bar{V}a$ from (6). improved matrix H

$H_{\hat{p}}$= UaGaΣaVa^T

$$H_{\hat{p}} = g_i \sigma_i u_i v_i^T$$

σi represents ith unique value of Σ.

Averaging along the antidiagonals of Hp allows for the recovery of the augmented signal p(k). Hansen and Jensen [18] and Duologue and Karayiannis [17] later showed that the entire procedure is comparable to a single global FIR-filtering operation on the noisy time signal. The output $g_i\sigma_i u_i v_i^T$is created by filtering the noisy signal x(k) with its matching eigen filter vi and its inverted version J vi. Understand that this leads to a zero-phase filtering operation because of filter theory. The variation in length of the anti-diagonals of the signal inspection matrix is taken into account when the enhanced signal p(k) is produced from the improved observation matrix Hs by multiplying H1, 1/2, 1/3,..., 1/b, 1/b,..., 1/3, 1/2, 1 on the diagonal of D.

This significant discovery provides an intriguing frequency-domain justification for the signal subspace denoising method. The main benefit of adopting the SVD over the EVD is that the covariance matrix does not have to be calculated directly. Then continue to emphasize the SVD description throughout this work. The use of any arbitrary (structured) covariance estimates, including, for example, the empirical Toeplitz covariance matrix, is made possible by the EVD-based technique, which is highlighted as an alternative method for carrying out all estimators.

Working Procedure:

In order to demonstrate how to build a convolutional neural network-based audio noise reduction, a 6-layer neural network has been built that can distinguish between clear and noisy audio. A very small network that a CPU can also manage will be built. Traditional neural networks that are very good at categorizing audio noise contain many more variables and take a lengthy time when trained on a regular CPU. However, to show the functionality on how TENSORFLOW may be used to build convolutional neural networks in practical situations.

In essence, neural networks are computational models that can be used to address optimization issues. Neurons, the fundamental building blocks of processing in neural networks, are used to create them. A neuron creates an output, let's say z= wx + b, by multiplying an input, let's say x, by a variable, let's say w, and adding another variable, let's say b. To induce the final result (activation of a neuron), this value is transferred to a non-linear function called the activation function (f). Different activation mechanisms exist. Sigmoid is a well-liked activation function. A form of brain cell that uses the sigmoid

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

function as an activation function is referred to as a "sigmoid neuron." Names for additional varieties of neurons, including TanH and RELU, vary on the activation roles of those cells.

When a new audio file is uploaded, the CNN train model is applied to the new audio file to locate noise and remove it. The CNN algorithm is trained with clean and noisy audio files to create a train model. The CNN algorithm will anticipate all noisy information from the audio file and then replace it with false values to eliminate noise. CNN will select audio sub signals while eliminating noise and then extrapolate noise from those sub signals. Following the elimination of all noise, CNN will produce two files, "noise.wav," which contains all noise data, and "clean.wav," which contains all clean audio data. They will both be saved in the "SpeechDenoise" folder.

Results Analysis:

A layer, which is the further building block of neural networks, is created by placing neurons in a single sentence “Fig. 4” with layers is shown below.

Input Layer Hidden Layer Output Layer

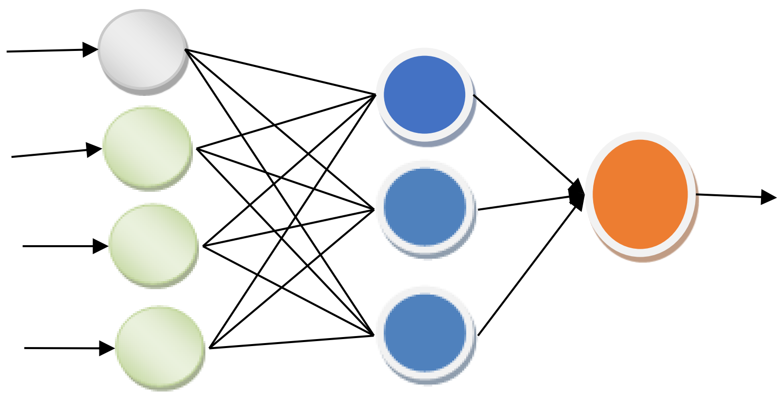

Figure 4 Layers from Input to Output

Multiple layers work together to remove any layers that contain noise in order to anticipate noise class, and this process is repeated until no further progress is possible. To implement this project some audio files have been using which contain noise and all those noise audio files are saved inside a folder. Upload audio file which is corrupted with noise and then extract Features from Audio to retrieve spectrogram from audio file. See the total time it took to retrieve spectrogram from audio file. Now Run deep Learning Speech Enhancement Algorithm to denoise audio file.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

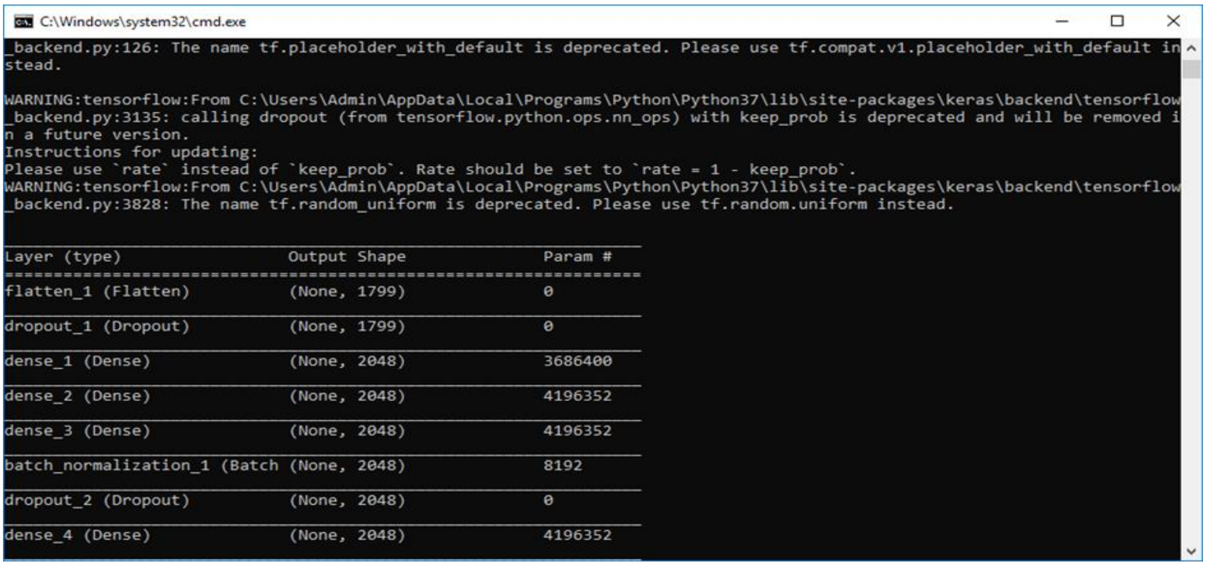

Figure 5 Extracting audio features

In above “Fig. 5” you can see, how the tensorflow deep learning is used to denoise the audio. See below graph with and without noise after enhancement.

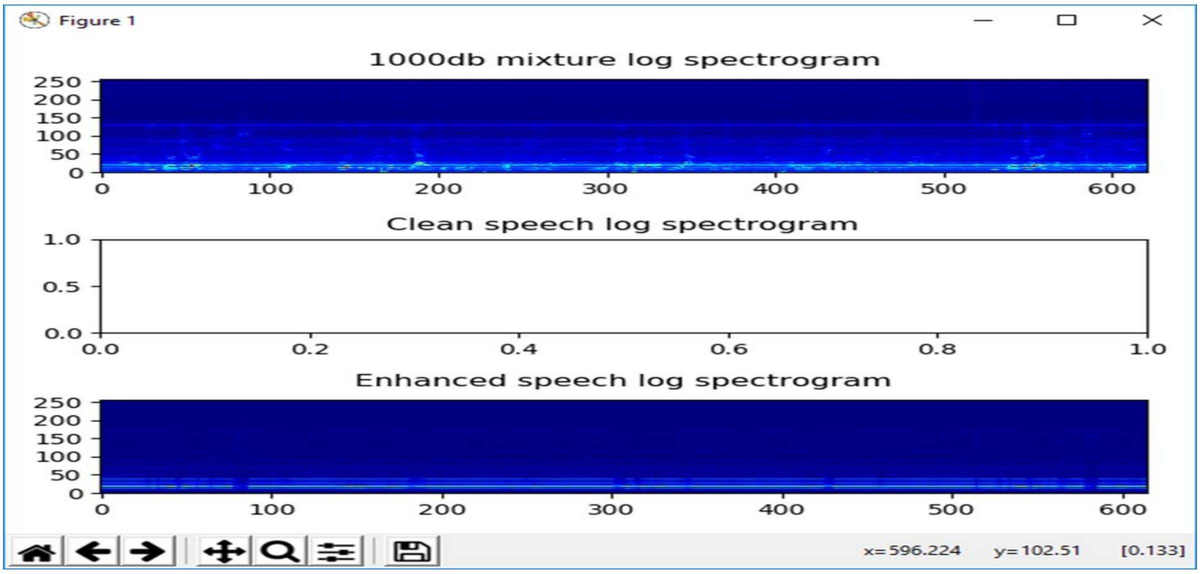

Figure 6 Clean and Enhanced log spectrogram

In above “Fig. 6” in top panel there is lot of noise and in below panel you can see noise is removed after enhance speech. Now close above graph and click on 'View Graph' to view noise in more details.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Figure 7 Graph of online speech enhancement

In above graph “Fig. 7” dark region indicates noise, now close above graph and click on pink color screen to get more graphs.

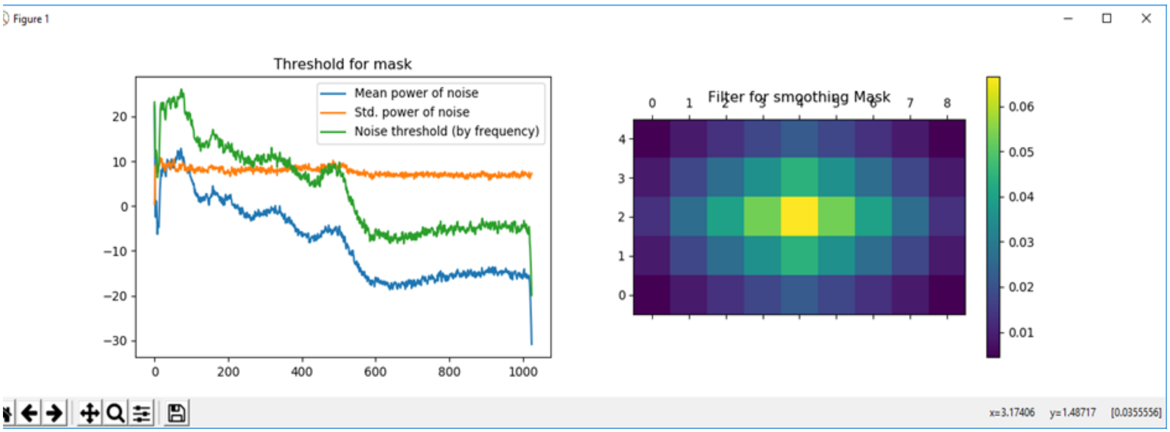

Figure 8 Graph of mean and standard deviation

In above graph “Fig. 8” you can see mean and standard deviation noise graph.

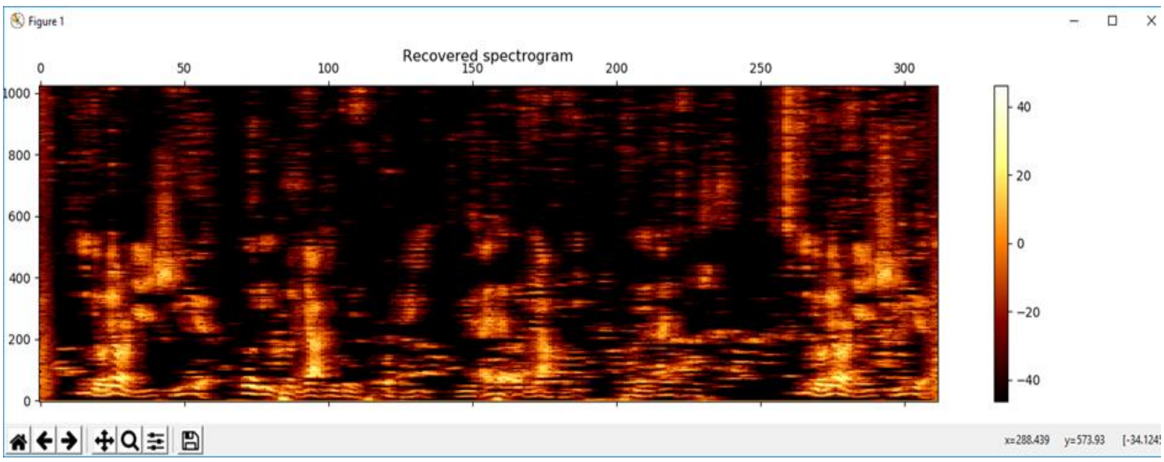

Figure 9 Recovered Spectrogram

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

In the above graph "Fig. 9" you can see the spectrogram recovered with clean audio and less noise is there. The folders noise.wav and clean.wav are saved inside the 'SpeechDenoise' folder. At last, you will get two files one with contains only noise and the other contain clean audio data.

Conclusion:

The suggested CNN-based method is computationally efficient and performs real-time speech enhancement with little audio latency. As a powerful and very customizable method for improving voice interpretation in speech interaction applications as well as the accuracy of recognizing speech automatically in additive noise conditions, signal subspace speech augmentation has come into prominence. In this essay, evaluated had investigated the fundamental idea of subspace smoothing and evaluated how well the most popular optimization criteria performed. For both the white noise and colored noise cases, developed a theoretical estimator to experimentally evaluate an absolute limit on the efficacy that any subspace-based technique may reach. This study evaluates a noise-dominant subspace reduction method for improving voice. By lowering the noise-dominant eigenvalues, this procedure eliminates background noise that is less directed in the subspace domain. The two steps of the strategy that come after that correspond to the different kinds of noise. Less-directional ambient noise is minimized in the first stage by eliminating the subspace that is dominated by noise. It is achieved by increasing the importance of the eigenvalues of the spatial correlation matrix. While the power of directional components is focused on a select few dominant eigenvalues, the energy of less directed noise is dispersed among all eigenvalues. In the second stage, a minimum variance beamformer is used to distinguish the spectrum of the target source from the mixture of spectra from the other multiple directed components. The proposed approach is then tested in both a real-world and a model-simulated setting.

Traditional neural networks (TNNs) and deep learning (DL) are both machine learning techniques that use artificial neural networks to model and classify data. However, there are some key differences between the two:

Depth: The most significant difference between TNNs and DL is the number of layers. TNNs usually have only one or two hidden layers, while deep learning architectures can have many more layers, often in the range of tens to hundreds. This deeper architecture allows deep learning networks to learn more complex representations of the data.

Feature Engineering: TNNs require significant feature engineering to extract meaningful features from raw data before inputting them into the network. Deep learning, on the other hand, learns the relevant features from the raw data through its deep architecture, which can identify patterns and features automatically.

Computational Resources: DL requires more computational resources than TNNs because of its deeper architecture. The added complexity and depth of DL networks require more computing power, memory, and processing time.

Training data requirements: Deep learning requires more training data than traditional neural networks. The additional data allows the deep architecture to learn more accurate representations of the data and reduce overfitting.

Performance: Deep learning often achieves better performance than traditional neural networks, especially in tasks such as image recognition, speech recognition, and natural language processing. The additional layers of the deep architecture allow it to identify complex patterns and features, making it better suited to handle complex tasks.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

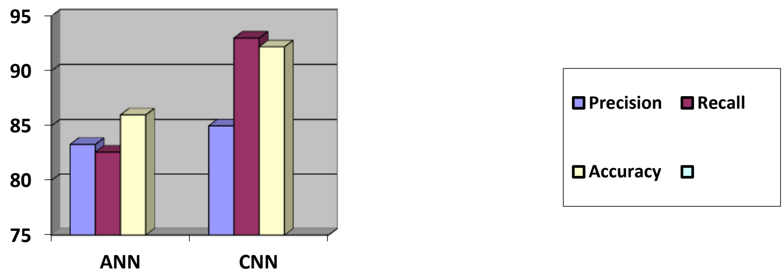

Figure 10 Model Comparison Between ANN and CNN

	ANN	CNN
Precision	83.3%	85%
Recall	82.6%	93%
Accuracy	86%	92.2%

Table I Comparison of values between ANN and CNN

In summary, deep learning is a more complex and powerful extension of traditional neural networks that can automatically learn complex features from raw data, requires more computational resources, and typically performs better on complex tasks.

References:

1. Dr. M Babu Rao, Image Processing, Machine Learning, Deep Learning, “Visual and buying sequence features-based product image recommendation using optimization based deep residual network”, published in the journal of Gene Expression Patterns, Volume 45, 2022, September 2022, ISSN 1567-133X, https://doi.org/10.1016/j.gep.2022.119261.
2. T.W. Lee, Independent Component Analysis. Nobel,Massachusetts: Kluber, 1998.
3. L. J. Griffiths and K.M.Buckley, "Control of stationary patterns in a linearLimited Adaptive Arrays”, IEEE Trans. acoustics speech,signal processing, vol. ASSP-35, pp. 917–926, July 1987.
4. Shaik Salma Begum, Dr. D.Rajya Lakshmi, GLCM of Fuzzy Clustering Means for Textural Future Extraction of Brain Tumor in Probabilistic Neural Networks, DOI: November 2019.
5. R. Roy and T. Kaiath, "ESPRIT Estimating Signal Parameters Using Rotational Invariant Method”, IEEE Trans. acoustic speech, signal Processing, Vol. 37, pp. 984–995, July 1989.
6. H.Wang and M. Kaveh, “Coherent signal processing in subspaces Detect and estimate multiple broadband angles of arrival Source, IEEE Trans. Acoustic, Speech, and Signal Processing, vol. 33, p. 823–831, April 1985.
7. H. Hung and M. Kaveh, “Focus Matrix for CoherenceSignals Spatial Processing”, IEEE Trans. Acoustic,Speech, and Signal Processing, vol. 36, pp. 1272–1281,August 1988.
8. Reshma R, Usha Naidu S, Sathiyavathi V, Sairamesh L, Stock market prediction using machine learning techniques, DOI: 10.3233/APC210156
9. M. Waks and T. Kailat, “Signal Sensing by Information Theory Criteria,” IEEE Trans. Acoustic, Speech, and Signal Processing, vol. ASSP-33, p. 387–392, 4. 1985.
10. J.R. Deller, J. GRAM. Proakis and J. H. L. Hansen,Discrete Time Processing. voice signal. New York: Macmillan, 1993.