International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7059
Article Received: 18 March 2023 Revised: 26 April 2023 Accepted: 20 May 2023

Machine Learning Based Dynamic Band Selection for Splitting Auditory Signals to Reduce Inner Ear Hearing Losses

Sudhir Narsing Divekar[1], Dr. Manoj Kumar Nigam[2]
[1] Department of Electronics Engineering, MATS University, Raipur (C.G.)
sndivekar.pp@gmail.com
[2] Professor, Department of Electrical & Electronics Engineering, MATS University, Raipur (C.G.)
drmanojk@matsuniversity.ac.in

Abstract—Quality of hearing has been severely impacted due to signal losses occurs in the human inner ear explicitly in the region of cochlea. Loudness recruitment, degraded frequency selectivity and auditory masking are the major outward effects of inner ear hearing losses. Splitting auditory signals into frequency bands and presenting dichotically to both ears became a comprehensive solution to reduce inner ear hearing losses. However, these methods divide input signal into the fix number of frequency bands, this limits their applicability where signals have large variations in their spectral characteristics. To address this challenge, we have proposed machine learning based intelligent band selection algorithm to split auditory signals dynamically. Proposed algorithm analyze input speech signal based on spectral characteristics to determine the optimum number of bands required to effectively present major acoustic cues of the signal. Further, dynamic splitting algorithm efficiently divides signal for dichotic presentation. Proposed method has been examined on large number of subjects from different age groups and gender having cochlear hearing impairment. Qualitative and quantitative assessment shown significant improvement in the recognition score with substantial reduction in the response time.

Index Terms— Inner ear hearing loss, Intelligent band selection, Spectral splitting, Cochlear hearing impairment, Machine learning algorithm.

I. INTRODUCTION

HUMAN ear mainly comprises of outer ear, middle ear and inner ear. Hearing losses associated with outer ear and middle ear are due to loud sound exposure, contagious infections, or due to the head injury. Most often these problems are resolved using proper medication or in case of severity they can be treated surgically [1]. However, hearing losses associated with inner ear are difficult to address using prevailing medical treatments. Inner ear losses are resulted from damage to inner or outer hair cells, hardening of organ of corti or aging can also be a substantive cause for signal losses as auditory functions degenerates more hastily in elderly. Loudness recruitment, degraded frequency selectivity and auditory masking are the major outward effects of cochlear hearing losses. Unfortunately, there is no comprehensive technique available to comprehensively address hearing losses occurs in the human inner ears [2].

To reduce hearing losses in the inner ears explicitly occurring in the region of cochlea, various band splitting methods has been proposed in the literature. These methods split auditory signals to form different bands based on their frequency contents. To split auditory signals, either bank of filters or transforms such as wavelets has popularly been used in the literature. Simultaneously presenting alternate split bands in even-odd manner to both ears significantly minimized the frequency overlapping and helped in improving the audibility in hearing impaired [3]. Though, such methods are efficient when spectral cues of signal are flat and consistent, but they underperform in real time scenario where interference of the background noise introduces unwanted variations and fluctuations in the spectral information of signal.

To address this issue, we proposed machine learning based intelligent band selection algorithm to split auditory signals. Proposed algorithm analyze input speech signal based on spectral characteristics such as place and manner of articulation, relations between formant frequencies, voice and unvoiced distinction, and vowel differentiation. Based on this spectral analysis, algorithm determines optimum number of bands required to effectively present major acoustic cues of the signal. Algorithm also fixes the frequency range of each band dynamically. Proposed method has been examined on large number of subjects from different age groups and gender having cochlear hearing impairment. Qualitative and quantitative assessment shown significant improvement in the recognition score with substantial reduction in the response time when compared with state-of-the-art methods.

The rest of the paper is organized as follows: Section II discusses existing band splitting methods and their band

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7059
***Article Received**: 18 March 2023 **Revised**: 26 April 2023 **Accepted**: 20 May 2023*

formation strategies. Proposed methodology has been elaborated in Section III with various novel algorithms involved at each step of implementation. Section IV provides qualitative and quantitative assessment of proposed methodology. Further, testing materials used for validation has been detailed in this section. This section deliberates performance analysis of proposed method in comparison with state-of-the-art methods. Finally, outcomes of this research have been concluded in section V.

II. EXISTING BAND SPLITTING METHODS

In literature, various band splitting methods based on different frequency bands has been reported. Eighteen bands of constant bandwidths, nineteen bands of one-third octave bandwidths and eighteen bands of critical bandwidths using pair of complementary filter sets were few among well performing methods [4]. These methods split signals in spectral domain using parallel dual filters with 512 number of coefficients. Technique of frequency sampling has been used to design filters having reciprocal amplitude response with phase linearity. Each of these filters have ripples < 1 dB in the pass band with negative gain crossover of ~6dB. Sequence of bands with range of frequencies in each band has been depicted in the Table I.

In constant bandwidth method, there are eighteen number of bands with bandwidth of 270 Hz respectively. Lowest attenuation in the stop band was 64dB having transition band of 75-80Hz. Level of loudness with constant bandwidth filter observed to be same for both ears when number of bands were above 16. Amplitude response do not reflect explicit differentiation between stop bands and pass bands. Further, noncognitive balance could not be achieved as there is unwanted variations in the crossover gain between constant bands [5]. Also, compared to critical bands and one-third octave filters, constant band filter induces higher distortion [6].

Critical bandwidth method also has eighteen bands with lowest of 29dB attenuation in stop bands and ~55Hz width of transition bands. At bottom frequencies, width of bands is almost constant and therefore at these frequencies amplitude response of critical band looks similar to constant band filter [7]. Whereas at higher frequencies it matches with amplitude response of one-third octave filter since critical bandwidths are approximately proportional to centre frequencies [8].

Table I: Range of frequencies in each of pass band (KHz)

Bands	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Constant Bandwidth	0.001-0.028	0.027-0.056	0.056-0.084	0.084-1.111	1.111-1.396	1.396-1.167	1.167-1.950	1.950-2.227	2.227-2.505	2.505-2.782	2.782-3.059	3.059-3.336	3.336-3.613	3.613-3.891	3.891-4.168	4.168-4.445	4.445-4.722	4.722-5.000	--
1/3 Octave Bandwidth	0.0708-0.089	0.089-0.112	0.112-0.141	0.141-0.178	0.178-0.224	0.224-0.282	0.282-0.355	0.355-0.447	0.447-0.562	0.562-0.708	0.708-0.891	0.891-1.120	1.120-1.410	1.410-1.780	1.780-2.240	2.240-2.820	2.820-3.550	3.550-4.470	4.470-5.000
Critical Bandwidth	0.01-0.20	0.20-0.30	0.30-0.40	0.40-0.51	0.51-0.63	0.63-0.77	0.77-0.92	0.92-1.08	1.08-1.27	1.27-1.48	1.48-1.72	1.72-2.00	2.00-2.32	2.32-2.70	2.70-3.15	3.15-3.70	3.70-4.40	4.40-5.00	--

In one-third octave method, number of bands are nineteen with minimum of 22dB attenuation in stop bands which is lowest compared to constant and critical band filters. Pass bands of octave filters are narrow at bottom frequencies and becomes wider with increasing frequencies. Here, width of transition band is same as that of constant band and that is around 80Hz [9]. Though, critical and octave filters are producing lower distortion with small inter-band overlapping but in presence of noise their speech intelligibility got highly deteriorated. Major drawback of these methods was lateralization of signal towards one ear which causes high level of perceptual loudness [10].

III. PROPOSED METHODOLOGY

To overcome the limitations of existing static band formation methods such as reduced speech intelligibility, high lateralization of signal and elevated level of perceptual loudness. Here, we proposed intelligent band selection algorithm that splits input auditory signal in dynamic number of bands based on spectral characteristics of the signal. Also, while forming the bands it takes level of noise contents into the consideration, which allows to effectively present important acoustic cues to ears. Following sections elaborates each step involved in the proposed methodology that includes; acquisition of auditory materials, extraction of spectral cues, intelligent band selection algorithm, dynamic splitting algorithm and stimulus presentation to ears. Block

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7059
***Article Received**: 18 March 2023 **Revised**: 26 April 2023 **Accepted**: 20 May 2023*

schematic of proposed methodology has been shown in figure 3.1.

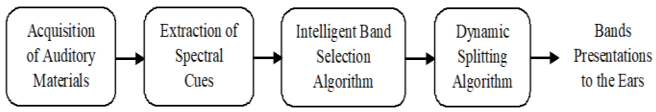

Figure 3.1: Block Schematic of Proposed Methodology

A. Acquisition of Auditory Materials:

In different researches different auditory materials has been used by the investigators in order to examine their methodology under varying conditions. Such speech materials include phonetic, linguistic and acoustic variables, nonsense syllables, and single letter to even sentences. Nonsense syllables in different combinations such as vowel-consonant, consonant-vowel, vowel-consonant-vowel and consonant-vowel-consonant have been widely used for analysis of many investigations [11].

For this study, we have screened phonetically balanced English monosyllables as testing materials since these monosyllables have intervocalic consonants and they are considered to be most appropriate to validate speech intelligibility of any algorithm. Being phonetically balanced, these words have equivalent phonetic composition that are best representative for everyday English speech since they contain most occurred frequencies of speech sounds. Total 60 phonetically balanced English monosyllables have been synthetically generated in three female and two male voices. Sampling frequency has been set to 10 KHz with pre-set voice volume of 0dB, 2 channels, 16 bits standard. For varying the performance of algorithm in the presence of noise, different signal to noise ratio has been adjusted by mixing of noises into the auditory materials.

B. Extraction of Spectral Cues:

The purpose of extraction of spectral characteristics is to analyze acoustic energy distribution of auditory signals across range of frequencies. From this analysis, we locate zero crossing or low energy moments from signal spectrum and these moments are the key junctures where bands were split-off. Cues that we identified most effective for spectral analysis were linear prediction cepstral coefficients, formant frequencies, zero crossing rate, spectral centroid, spectral bandwidth, spectral flatness, spectral roll off and spectral contrast. To extract aforementioned cues, two algorithms has been applied to input auditory signals; first was linear prediction algorithm used to extract two features; linear prediction cepstral coefficients (LPCC) and formant frequencies. Second was fast fourier transform (FFT) that has been used to extract six features; zero crossing rate and other five spectral features derived from spectrogram of auditory signal.

Linear prediction is one of the most efficient algorithms for estimation of vocal tract filter that shapes the human voice. At any instant of time, linear prediction approximates speech signal as linear combination of their past values [12][14]. Linear predictor coefficients (LPC) have ability to represent the behaviour of speech signals at their steady state. The logarithm of the magnitude of the LP coefficients, denoted as log(|A(k)|), can be calculated as:

$$\log(|A(k)|) = \log\left[\sqrt{\sum_{m=1}^{p} A(m)^2}\,\right], \text{ for } k = 1, 2, ..., p \quad \ldots (3.1)$$

LPC precisely estimates sounds parameters and comparatively fast in terms of computational time, therefore cepstral coefficients are derived from it and hence termed as linear prediction cepstral coefficients (LPCC). LPCC are less sensitive to quantization noise invoking low error rate compared to LPC. Figure 3.2 shows the steps involved in computation of LPCC.

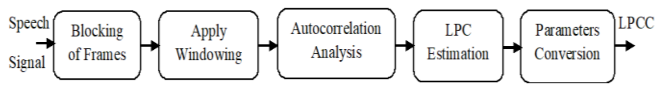

Figure 3.2: Computation of linear prediction cepstral coefficients (LPCC)

Formant frequencies are dependent on the spectral envelope of the speech signal, and therefore, are not constant. They can change dynamically as the speech is produced, so linear prediction (LP) analysis has been used to estimate the formants. Firstly, signal has been windowed into overlapping frames and processed frame by frame. Then the autocorrelation coefficients of each frame calculated to estimate the autoregression AR coefficients using the Levinson-Durbin recursion algorithm. Further, estimated AR coefficients has been used to generate the linear prediction (LP) polynomial, which is a representation of the prediction filter. The roots of the LP polynomial extracted, which correspond to the poles of the prediction filter. Finally, the formant frequencies estimated as the negative reciprocals of the real parts of the poles of the prediction filter.

Zero crossing rate (ZCR) is a measure of the number of times a signal crosses the zero-amplitude axis in any given time interval. To compute ZCR using FFT, signal frames obtained during LP analysis has been used to count the number of zero crossings in each frame. FFT of each frame was calculated and the magnitude of the FFT of each frame squared to get the spectral energy of each frame. The ZCR is

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7059
***Article Received**: 18 March 2023 **Revised**: 26 April 2023 **Accepted**: 20 May 2023*

weighted by the spectral energy of each frame to obtain an energy-based ZCR. The energy-based ZCR of each frame is averaged to get the overall ZCR of the signal. Mathematical formula used to obtain ZCR of a signal x(t) expressed in equation 3.2.

$$ZCR = \frac{1}{T}\int\left|sign\big(x(t)\big) - sign\big(x(t-\Delta t)\big)\right|dt \qquad \ldots (3.2)$$

Here, T is the total time period and Δt represents frame interval.

Fast Fourier transform provides robust representation for audio signal and can be characterized effectively using five spectral features. First spectral cue obtained from signal spectrum was spectral centroid which is a measure of the center of mass of the power spectrum of a signal. Next is the spectral bandwidth that is a measure of the spread of the spectral power distribution of a signal. Spectral flatness measures uniformity of the spectral power distribution of a signal. Spectral roll off is the frequency below which a specified percentage of the total spectral energy lies. At the end, Spectral contrast was measured from the variation of the spectral power distribution of a signal [13][15]. Note that these spectral features we used not individually rather in combination with other features to obtain a more complete representation effective for next step of band selection.

C. Intelligent Band Selection Algorithm:

Band selection algorithm is the crucial segment of our proposed methodology. Machine learning based algorithm outputs even number of bands appropriate to divide input speech signal based on extracted spectral cues. Considering non-stationary nature of auditory signals, random forests algorithm has been found to be suitable due to its ability to handle non-linear decision boundaries for complex datasets and robust behaviour towards outliers. Also, as random forest uses bagging (bootstrapped samples) and feature randomness when building trees, it helps reduce overfitting in the model. Further, it is highly effective in identifying redundant or irrelevant features and thereby helped in reducing the dimensions of feature vector. Following are the steps involved in the design of intelligent band selection algorithm:

i. Split data: To build random forests machine learning (ML) model, a set of training and testing data has been formed by splitting 300 observations obtained from recordings of 60 phonetically balanced English monosyllables synthetically generated in the 5 different voices (3 females and 2 males). Total size of feature vector is 300 X 8 with number of predictors as 8. To split the feature vector into training and testing sets, we have used 5-fold cross-validation method where each fold has 60 observations. During each iteration, four folds were used for training and remaining one has been used for testing. Result of all five iterations has been averaged to determine the fitting accuracy of a model.

ii. Bootstrapping of decision trees: Next, the algorithm generates multiple decision trees from the training data. Each tree is generated by selecting a random subset of the training data and a random subset of the features. Tree is then grown using a recursive binary splitting process, where the feature that provides the best split is chosen at each node. Mathematical formulation of decision trees involves calculating Gini impurity at each node in the tree has been given in the equation 3.3.

$$Gi = 1 - \sum_{i=0}^{C}\left(\frac{n_i}{N}\right)^2 \qquad \ldots (3.3)$$

where C is the number of classes in the set

iii. Optimize the algorithm: Once the individual decision trees have been generated, the algorithm combines them to make a prediction. Based on the results of the evaluation, optimized the algorithm by adjusting its parameters such as number of trees, minimum number of samples per leaf, maximum tree depth to improve its performance.

iv. Band selection: Finally, use the optimized algorithm to perform band selection on new, unseen speech signals. The goal of the training process is to learn a mapping from the features to decide number of bands required to represent a signal. Most effective way to address inner ear hearing issues is to present the auditory signals in alternate bands to left and right ears simultaneously. By considering the same, our band selection algorithm has been tweaked to outputs even number of bands covering major acoustic cues required for speech recognition.

D. Dynamic Splitting Algorithm:

In this last stage, auditory signal has been divided into number of sub-bands recommended by intelligent band selection algorithm. Dynamic splitting algorithm works by taking an input audio signal and passing it through a bank of filters that separate the signal into different frequency bands. To split speech into dynamic bands, the Mel-frequency filter bank has been used due to its ability to mimic the human auditory system's response to different frequencies.

The Mel-frequency filter bank works by dividing the frequency spectrum into a series of overlapping triangular

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7059
***Article Received**: 18 March 2023 **Revised**: 26 April 2023 **Accepted**: 20 May 2023*

filters. The center frequencies of the filters are spaced evenly on a Mel-frequency scale, which is a logarithmic scale that approximates the human perception of pitch. The width of each filter is determined by its center frequency and the frequencies of its neighboring filters. Equation 3.4 applied to design the Mel-frequency filter bank.

$$h_{k(n)} = w(n)\cos[\frac{2\pi k(m)}{M} + \phi\ (n)] \quad \ldots (3.4)$$

where $h_{k(n)}$ is the filter coefficient at time n and frequency k, w(n) is the window function, $\phi\ (n)$ is the phase offset, M is the total number of filters, and m is the index of the center frequency of the kth filter.

Each filter was designed to let only a specific range of frequencies pass through, while blocking or attenuating all other frequencies. The output of each filter represents a sub-band of the original signal and the sum of the outputs of all filters gives the overall spectral content of the speech signal. Lastly, these even bands were presented to the human ears in dichotic manner by considering its effectiveness to reduce the frequency masking.

IV. PERFORMANCE ANALYSIS

Evaluation and assessment of the proposed methodology has been carried out on subjects from different age groups and gender. Performance has been systematically examined using various qualitative and quantitative measuring parameters. Results obtained from presentation of proposed dynamic bands methodology has been compared with that of unprocessed speech signals, processed signals using constant bands, critical bands and octave bands.

A. Subjects and Testing Set-up:

Around eighteen subjects having mild to severe cochlear hearing impairment has been participated in these listening tests. Out of 18, there were 11 males and 7 females having age between 13 to 78 years. For qualitative assessment, we divided them into five standard age groups; 0-14 years (children), 15-24 years (early age), 25-49 years (prime age), 50-64 years (mature age), 65 years and over (elderly).

Testing set-up built for conducting the listening tests contains interactive user interface to record subject's responses over speech materials presented. Total 60 synthetically generated phonetically balanced English monosyllables have been used as testing materials. Listening tests has been conducted in an acoustically isolated room. Processed bands were presented simultaneously to both the ears over headphones. Intensity of sound has been adjusted to normal range of hearing. UI has been designed in a such way that different stimulus was linked to a different push button on the subject's terminal. Subject has to respond by pressing appropriate push button based on the signal he/she interpreted. Presentation of stimulus was in random order and results were aggregated based on number of stimuli presented.

B. Qualitative Assessment:

For qualitative assessment of proposed methodology two metrics; perception score and latency has been used. Perception score has been computed based on rank given by the subject over listening the unprocessed and processed signals. There are four ranks ranging from 1 to 4 representing the quality of signal as poor, average, good and outstanding respectively. Results has been aggregated for all 60 responses recorded against 60 monosyllables presented to each subject. At last, rank values grouped together under five age groups.

Table 2 depicts the perception score under each age group obtained from 4 processing schemes including the proposed scheme of dynamic bands. In the table, grp1 to grp5 represents five age groups mentioned in the earlier section. Results shown that, almost under each age group perception score has been improved with processing scheme over the unprocessed, particularly for children and elderly it was exceedingly well. Comparing the perception score of processing schemes, dynamic bands has distinctly outperformed under each of the age group. Average perception score reported by proposed methodology was 3.93 on the scale of 4. Next to dynamic bands, performance of critical bands was better than constant and octave band methods except under prime age group where octave is outperforming over preceding two. Bar graph in figure 4.1 represents the percentage improvement in the perception score over unprocessed signal. Average improvement of 36.93% has been reported by dynamic bands over that of unprocessed signal and 12.28% over critical bands method. Constant and octave bands shown negative performance under prime and mature age groups.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7059
***Article Received**: 18 March 2023 **Revised**: 26 April 2023 **Accepted**: 20 May 2023*

Table 2: Perception Scores

Age Groups	Unprocessed Signal	Constant Bands	Critical Bands	Octave Bands	Dynamic Bands
grp1	3.26	3.32	3.84	3.52	4
grp2	3.18	3.21	3.56	3.24	3.87
grp3	2.69	1.94	2.97	3.63	3.89
grp4	3.31	3.33	3.38	3.19	3.93
grp5	1.91	3.04	3.75	3.41	3.98
Average	2.87	2.97	3.5	3.4	3.93

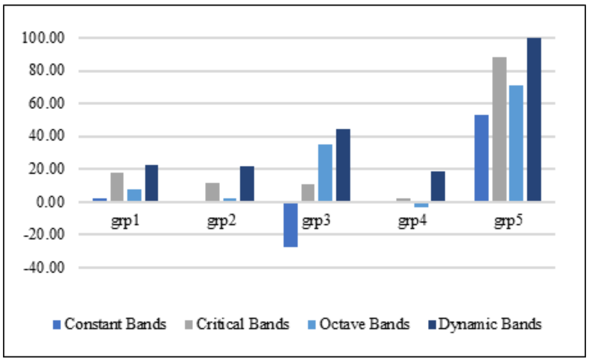

Fig. 4.2 Percentage improvement in Perception Scores

Table 3: Latency chart of four processing schemes and unprocessed signal

Age Groups	Method	Low	High	Mean	S.D.	P.D (%)
grp1	Unprocessed Signal	1.31	2.59	2.08	0.36	-
	Constant Bands	1.23	2.26	1.74	0.35	15.91
	Critical Bands	1.07	2.12	1.54	0.36	25.96
	Octave Bands	1.07	2.48	1.68	0.49	18.91
	Dynamic Bands	0.96	1.62	1.29	0.24	37.64
grp2	Unprocessed Signal	1.87	2.65	2.25	0.21	-
	Constant Bands	1.46	2.96	2.13	0.41	5.14
	Critical Bands	1.37	2.56	2.07	0.35	7.90
	Octave Bands	1.53	2.62	2.08	0.32	7.68
	Dynamic Bands	1.37	2.67	1.86	0.40	17.45
grp3	Unprocessed Signal	1.26	2.96	1.88	0.52	-
	Constant Bands	1.42	2.92	2.33	0.41	-1.58
	Critical Bands	1.04	2.53	1.61	0.43	14.32
	Octave Bands	1.09	2.73	1.53	0.51	18.49
	Dynamic Bands	1.09	2.73	1.51	0.49	19.58
grp4	Unprocessed Signal	1.87	2.59	2.30	0.21	-
	Constant Bands	1.70	2.56	2.17	0.23	5.51
	Critical Bands	1.40	2.61	2.02	0.37	12.26
	Octave Bands	1.37	2.76	1.89	0.40	-0.43
	Dynamic Bands	1.48	2.71	1.95	0.38	15.28
grp5	Unprocessed Signal	1.89	3.56	2.73	0.53	-
	Constant Bands	1.50	2.65	2.07	0.31	24.26
	Critical Bands	1.25	2.87	1.96	0.41	28.33
	Octave Bands	1.21	2.73	1.96	0.51	28.10
	Dynamic Bands	1.25	2.51	1.86	0.36	31.75

Figure 3.2 shows the latency chart for processing schemes and unprocessed signal under each of the age group. Latency is the time delay between signal presented and response submitted by the subject, measured in seconds. It has been computed to assess utility of the processing schemes in minimizing load on perception. Proportional decrease (P.D.) is the reduced response time in percentage compared to unprocessed signal. Lowest latency has been accomplished using dynamic bands over other processing schemes.

C. Quantitative Assessment:

For quantitative assessment of proposed methodology recognition accuracy, proportional improvement (P.I.) and non-recognition rate has been computed. Recognition accuracy in percentage has been calculated by taking ratio of number of correctly recognized stimuli to the total number of stimuli presented to the subject. Table 4 shows the recognition accuracy obtained for each of the methods considered in this assessment. Dynamic bands shown the average improvement in the recognition accuracy of 12.98% over the unprocessed whereas its 5.84% over the critical bands. Under each age group, proportional improvement (P.I.) representing the percentage improvement of processing

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7059
***Article Received**: 18 March 2023 **Revised**: 26 April 2023 **Accepted**: 20 May 2023*

schemes over unprocessed has been shown in the bar graph of figure 4.2. Almost under all age group, processing schemes are outperforming except under grp3 and grp4 where constant bands and octave bands underperformed.

Non recognition rate represents the events where subject does not recognize the presented stimuli. It counts the number of occasions where subject unable to categorize presented stimuli under any of 60 monosyllables. Figure 4.2 shows that non recognition rate of constant bands is highest and that is followed by octave and critical bands respectively. Dynamic bands managed to lower the non-recognition rate significantly, especially for elderly it has reported zero non-recognition rate.

Table 4: Evaluation of Recognition Accuracy (%) for Processing Schemes

Age Groups	Unprocessed Signal	Constant Bands	Critical Bands	Octave Bands	Dynamic Bands
grp1	85.55	86.66	93.33	91.11	97.77
grp2	78.88	82.22	88.88	86.66	91.74
grp3	81.1	76.67	83.33	85.55	92.06
grp4	80	83.33	87.77	78.88	94.44
grp5	82.22	84.45	90.15	89.18	96.66
Average	81.55	82.67	88.69	86.28	94.53

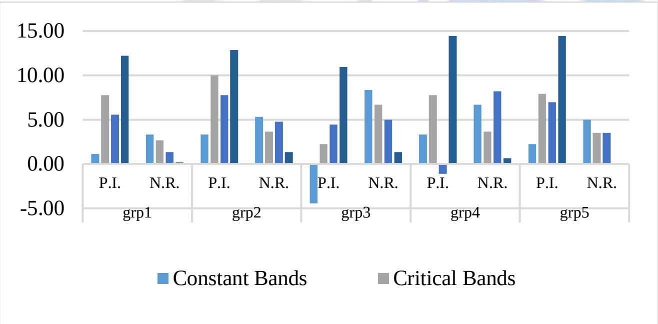

Figure 4.2: Proportional Improvement (%) and Non-Recognition Rate (%)

V. CONCLUSION AND DISCUSSIONS

This study proposed an efficient methodology to address the inner ear hearing losses. Existing methods based on static band allocation such as constant bands, critical bands and octave bands performs well in the absence of background noise however fails to interpret when distortion is high. Moreover, these methods also introduce high inter-band overlapping results in losing the signal information associated with corner frequencies of adjacent bands. Proposed methodology has efficiently addressed these issues by generating the frequency bands dynamically based on the spectral contents of input auditory signals. It also minimized lateralization of signal towards one ear that helped in improving the speech intelligibility.

Qualitative and quantitative assessment using different measuring parameters has been carried out to analyze the overall performance of proposed methodology under different scenarios. Synthetically generated phonetically balanced English monosyllables in female and male voices has been used as data for training the intelligent band selection algorithm, further this material has been used for audibility tests of the subjects. Interactive user interface has been built to carry out unbiased extensive testing on large number of subjects belonging to different age groups. Overall results, notably recognition accuracy and latency has shown the outperformance of proposed methodology over existing state-of-art methods.

Scope for further enhancement still exists in the few areas such as collection of wide range of databases can help in carrying out exhaustive fitting of machine learning model. Random forests observe slowness with large number of trees and makes it inadequate for real-time predictions. Hence, a novel machine learning algorithm can be designed that may suits for real time applications. Recent growth in the computing world with micro scaling of electronic devices provisioned the development of wearable's that can provide comfort in long-term usage. However, with existing methodologies it is challenging to obtain a high degree of precision, particularly under changing environmental conditions.

REFERENCES

[1] Brian C. J. Moore, "Introduction to Psychology of Hearing", Sixth Edition, Academic Press, 2012.

[2] F. Henry, M. Glavin and E. Jones, "Noise Reduction in Cochlear Implant Signal Processing: A Review and Recent Developments," in IEEE Reviews in Biomedical Engineering, vol. 16, pp. 319-331, 2023.

[3] Divekar S, Nigam MK., "Minimize Frequency Overlapping of Auditory Signals using Complementary Comb Filters", SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, 14(3), 333-336, 2022.

[4] Khatri, K. ., & Sharma, D. A. . (2020). ECG Signal Analysis for Heart Disease Detection Based on Sensor Data Analysis with Signal Processing by Deep Learning Architectures. Research Journal of Computer Systems and Engineering, 1(1), 06–10. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/11

[5] P. N. Kulkarni, P. C. Pandey, "Optimizing the Comb Filters for Spectral Splitting of Speech to Reduce the Effect of Spectral Masking", IEEE-International Conference on Signal processing, Communications and Networking Madras

Institute of Technology, Anna University Chennai India, Jan 4-6. pp.69-73, 2008.

[6] AGYEI , I. T. . (2021). Simulating HRM Technology Operations in Contemporary Retailing . International Journal of New Practices in Management and Engineering, 10(02), 10–14. https://doi.org/10.17762/ijnpme.v10i02.132

[7] Justin T. Lui, Katie de Champlain, Justin K. Chau, “Management of Adult Sensorineural Hearing Loss”, Evidence-Based Clinical Practice in Otolaryngology, Pages 15-24, Elsevier, 2018.

[8] Jongwoo Lim; Yeongjin Kim; Namkeun Kim, “Mechanical Effects of Cochlear Implants on Residual Hearing Loss: A Finite Element Analysis”, IEEE Transactions on Biomedical Engineering, Volume: 67, Issue: 11, pp. 3253 – 3261, Nov. 2020.

[9] Luca Remaggi; Philip J. B. Jackson; Wenwu Wang, “Modeling the Comb Filter Effect and Interaural Coherence for Binaural Source Separation”, IEEE/ACM Transactions on Audio, Speech, and Language Processing, Volume: 27, Issue: 12, pp. 2263 – 2277, Dec. 2019.

[10] Christopher Davies, Matthew Martinez, Catalina Fernández, Ana Flores, Anders Pedersen. Machine Learning Approaches for Predicting Student Performance. Kuwait Journal of Machine Learning, 2(1). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/174

[11] Sonia Tabibi, Andrea Kegel, Wai Kong Lai, Norbert Dillier, “Investigating the use of a Gammatone filterbank for a cochlear implant coding strategy”, Elsevier Journal of Neuroscience Methods, Volume 277, Pages 63-74, 2017.

[12] L. Lightburn, E. D. Sena, A. Moore, P. A. Naylor, M. Brookes, "Improving the perceptual quality of ideal binary masked speech", Proc. Int. Conf. Acoust. Speech Signal Process, pp. 661-665, 2017.

[13] Deshpande, V. (2021). Layered Intrusion Detection System Model for The Attack Detection with The Multi-Class Ensemble Classifier . Machine Learning Applications in Engineering Education and Management, 1(2), 01–06. Retrieved from http://yashikajournals.com/index.php/mlaeem/article/view/10

[14] Yuma Koizumi, Kenta Niwa, Yusuke Hioka, Kazunori Kobayashi, Yoichi Haneda, "DNN-Based Source Enhancement to Increase Objective Sound Quality Assessment Score", IEEE Transactions on Audio Speech and Language Processing, vol. 26, no. 10, pp. 1780-1792, 2018.

[15] D. S. Williamson, D. L. Wang, "Time-frequency masking in the complex domain for speech dereverberation and denoising", IEEE/ACM Transaction on Audio Speech Lang. Process., vol. 25, no. 7, pp. 1492-1501, Jul. 2017.

[16] M. S. A. Chowdhury, "Linear predictor coefficient, power spectral analysis and two-layer feed forward network for bangla speech recognition," 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, pp. 1-6, 2019.

[17] Hussein, A.B., Lasheen, R.M., Emara, A.A. et al., “Listening effort in patients with sensorineural hearing loss with and without hearing aids”, Springer Journal of Otolaryngology, Vol. 38 (99), 2022.

[18] Sable, N.P., Wanve, O., Singh, A., Wable, S., Hanabar, Y. (2023). Pressure Prediction System in Lung Circuit Using Deep Learning. In: Choudrie, J., Mahalle, P., Perumal, T., Joshi, A. (eds) ICT with Intelligent Applications. Smart Innovation, Systems and Technologies, vol 311. Springer, Singapore. https://doi.org/10.1007/978-981-19-3571-8_56.

[19] Abdul Rahman, Artificial Intelligence in Drug Discovery and Personalized Medicine , Machine Learning Applications Conference Proceedings, Vol 1 2021.

[20] N. P. Sable, V. U. Rathod, P. N. Mahalle and D. R. Birari, "A Multiple Stage Deep Learning Model for NID in MANETs," 2022 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2022, pp. 1-6, doi: 10.1109/ESCI53509.2022.9758191.