Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 4 (2023)

Designing a Fault-Tolerant System for Hearing Aids: Echo Detection and Rectification Using SVM Classifier

[1*]Vijay Kumar Reddy I, [2]Dr. Aravinda H. S.
[1]Research Scholar, Dept. of ECE, JSSATE,
[2]Professor, Dept. of ECE, JSSATE,
[1,2]Affiliated to Visvesvaraya Technological University, Belagavi – 590018

Abstract: **-**Hearing aids, akin to various mechanical instruments, emit specific auditory signals indicative of operational health, accurate parameter settings, or maintenance requirements. Analogous to automobile users discerning vehicle health via auditory cues without formal instruction, such auditory recognition emerges rapidly. Yet, embodying this auditory discernment within an artificial framework introduces intricate challenges. The scope of artificial structures capable of pre-emptive defect or maintenance identification through auditory signals remains expansive. In hearing aids, commonly observed complications encompass inadequate amplification, acoustic reverberations presenting as echoes, and diminished acoustic integrity leading to auditory deformities, frequently labelled as Non-linear harmonic distortion. The present study predominantly concentrates on the pervasive issue of auditory reverberations in hearing aids, often leading to user discomfort. The primary goal revolves around developing a resilient system harnessing an SVM classifier to detect echoes. Notably, the outlined architecture has showcased an accuracy metric of 95.4%.

***Keywords**: Hearing Aids, Auditory Signals, Artificial Framework, Acoustic Reverberations, Echoes, Acoustic Integrity, Auditory Deformities, Non-Linear Harmonic Distortion, SVM Classifier.*

I. Introduction

Mechanical devices, encompassing combustion engines, pipes transporting gaseous or liquid chemicals, and drills, may manifest flaws during operation. While a flaw doesn't render a system inoperable, neglect can lead to system failure. Such flaws arise from wear, part failure, or process outside specified tolerances and conditions. Ensuring proper settings is crucial, as incorrect configurations might result in malfunctions. System failures pose risks to safety and the environment, leading to financial implications. This section elucidates faults, their types, and detection methods in hearing aids.

1. Faults in Hearing Aids:

Faults in hearing aids refer to introducing undesirable audible components resulting from the processed signal's interaction with an internal non-linear mechanism. If these introduced components remain minor relative to the overall signal strength, they might not cause interference. However, when these components overshadow the desired sound, they can detract from the listening experience, rendering sounds annoying or incomprehensible.

2. Types of Faults in Hearing Aids:

Faults can be broadly classified into three categories. Immediate faults lead to an instant system disruption, causing significant components, like a car's crankshaft, to fail. Gradual faults permit the system to operate, albeit with accumulating damage until a total system failure. Misconfigured systems fall under a distinct fault category; they work but influence the system's performance and the wear experienced during operation. Common faults in hearing aids include inadequate gain, acoustic feedback (echoes), and subpar acoustic quality marked by sound distortion, also known as Non-linear harmonic distortion. Addressing these issues requires effective fault detection and the design of fault-tolerant systems. Figure 1 presents the general block diagram of a hearing aid.

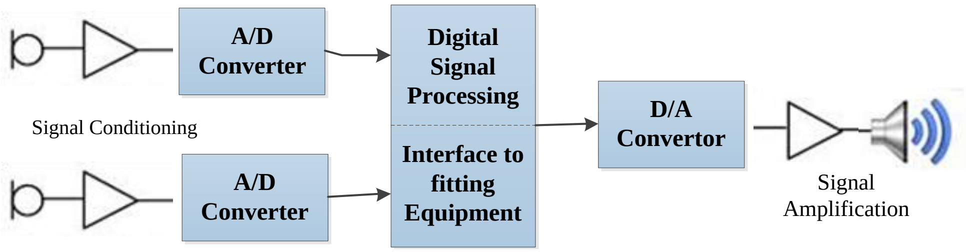


Figure 1: General Block Diagram of the Digital Hearing Aid.

The structure and operation of hearing devices within the external ear canal lead to several challenges and faults. These issues are outlined below:

• **Insufficient Gain:** Sound amplification is critical for individuals with pronounced hearing loss. Although hearing devices do not generate sound, they amplify incoming sounds detected by microphones, a process called gain. The resulting sound pressure level produced by the hearing device combines the initial incoming sound with the added gain from the device. These devices adjust the gain applied to incoming sounds across different frequencies, enhancing frequencies where hearing is impaired and reducing gain where hearing is adequate. However, excessive amplification can lead to uncomfortably loud or even harmful sound levels.

• **Compromised Acoustic Quality and Sound Distortion (Non-linear Harmonic Distortion):** Conventional hearing aids often have limited frequency ranges they can amplify effectively. Most are optimized for the speech frequency range (500–2000 Hz). However, they might struggle to consistently amplify lower frequencies (as seen in Meniere disease) or higher frequencies (as in cases of presbyacusis or autotoxicity). This inconsistency can introduce distortions, especially when transitioning between frequency ranges, leading to phase shifts.

1. Acoustic Feedback: Some acoustic waves produced by the hearing device's speaker reflect off the eardrum and leak through the space between the machine and the external ear canal wall. These reflected waves re-enter the microphone, get re-amplified, and produce a high-pitched feedback sound. Devices where the microphone and speaker are nearby, or those used by individuals with larger mastoid cavities, are especially prone to this feedback phenomenon.

2. Fault Detection Methods: Techniques for fault detection aim to identify deteriorating operational conditions or pinpoint specific faults during device operation. These techniques rely on signals produced by the operating device. Some methods are based on known platform-specific issues. There are two primary strategies for fault detection:

- The first method leverages known fault signals to identify specific problems. This approach relies on a dataset of pre-recorded fault signals representing a range of fault conditions and routine operations. The technique employs classifiers to categorize real-time operational signals based on this pre-established dataset. These methods excel in identifying specific, known faults, mainly if these faults manifest in consistent patterns. They are often robust against external noise because they focus on detecting established fault patterns. However, they have limitations, such as their inability to detect new or previously unknown faults. Acquiring a comprehensive set of fault signals can also be challenging, given the unpredictable nature of faults.

- The second method centers on identifying anomalies in signals that wouldn't appear during regular operation. The difference between a standard operational signal and an anomalous one forms the basis for this categorization. This versatile technique can adapt to various scenarios without requiring prior data collection. It can also recognize previously undetected or new faults. Notably, it excels in identifying ambiguous issues, such as squeaking sounds [9]. However, this method comes with its set of challenges. It must account for the variable signals machines produce throughout the day or during different operational phases. Additionally, this method can be more susceptible to noise interference from external sources.

II. Literature Survey

Understanding signal-based defect detection methodologies is crucial for informed research. A comprehensive literature review offers insights into this domain's current state of the art.

- Kurunakar et al. [01]: This research delved into detecting defects in internal combustion engines using a discrete wavelet transform (DWT) at level five. The DWT's analysis window size is tailored to the frequency component, facilitating multi-resolution signal analysis. An artificial neural network (ANN) of unspecified configuration was used for signal classification. The team tested a Hyundai I20 combustion engine for potential defects, targeting issues like timing belts and fuel malfunctions. Although the study was executed on a Windows Mobile platform, the results did not prove the system's accuracy. However, the time analysis on this platform took 1 minute and 35 seconds.

- Maniak et al. [02]: The research revolved around using sound analysis for quality assurance in the production line of sound signaling devices. By analyzing the signal, 26 Mel frequency cepstral coefficients were derived. These coefficients were inputs for an ANN with 50 hidden neurons, trained on 40 faulty and 160 non-faulty samples. Their methodology achieved an impressive accuracy rate of up to 99.7%.

- Hayashi et al. [03]: The team targeted the detection of electromagnetic valve defects using differences in the power spectrum. A feature vector, 240 units in length, was crafted using frequencies spanning from 55 Hz to 1250 Hz at 5-Hz intervals. An ANN comprising 30 hidden neurons undertook the task of classifying faults. Although the network was trained on standard valve sounds and noises emanating from a 2mm diameter hole or a 5mm crack, the research did not offer explicit numerical results.

- Benko et al. [04]: This study honed in on the sound module of a system and adopted a multi-domain approach to defect identification. The defect detection system, employed for quality control of vacuum cleaner motors, incorporated various signals - sound, vibration, rotational speed, voltage, current, and brush voltage. Post-bandpass filtering the signal within a frequency range of 2.5 to 3.5 kHz, a Hilbert transform was applied. By altering the phase of all frequency components by $pi/2$ rad, this transform enables the computation of amplitudes and instantaneous frequencies. Subsequently, the signal was filtered to extract four frequency components, which were then integrated into a single feature using RMS. Another feature was the utilization of two weighted histograms of the signal to identify intermittent brushing noises. The two types of faults studied were rubbing sounds and periodic brushing, and the two features were used for their detection. However, the research didn't extend to presenting classification results beyond visualization.

In conclusion, these studies showcase the diversity and depth of approaches available for signal-based defect detection, each with strengths, methodologies, and challenges.

III. Methodology

In this section, a structured methodology for identifying faults in hearing aids is presented. This methodology relies on sound signals generated by testing apparatuses. The illustrated process encompasses six pivotal stages, as delineated in Figure 2:

1. Data Acquisition: This stage involves capturing the auditory output of the hearing aid using high-sensitivity sensors.

2. Pre-processing: The acquired sound signal undergoes denoising procedures to eliminate extraneous signals, ensuring the primary speech signal remains undistorted.

3. Feature Extraction: Essential and relevant features are extracted from the pre-processed signal to aid the subsequent classification process.

4. Fault Diagnosis with SVM Classifier: A Support Vector Machine (SVM) classifier is trained using the extracted features to identify and diagnose potential faults within the hearing aid proficiently.

3.1 Sound Acquisition

Before a computer-based system can effectively extract salient features from a sound signal, the signal must be accurately measured, converted into a digital format, and subjected to initial processing. For the measurement phase, a directional microphone serves as the primary instrument. The analog sound signal is then transformed into its digital counterpart using an analog-to-digital converter. This results in the sound being represented in a discrete sample format, making it amenable to digital processing.

An integral parameter in this conversion process is the sampling frequency or rate. This denotes the number of samples procured per unit time. Each sample's amplitude is quantified as a floating-point value or an integer. With the sound signal now in a discrete format, it becomes feasible for digital devices to process and analyze it further.

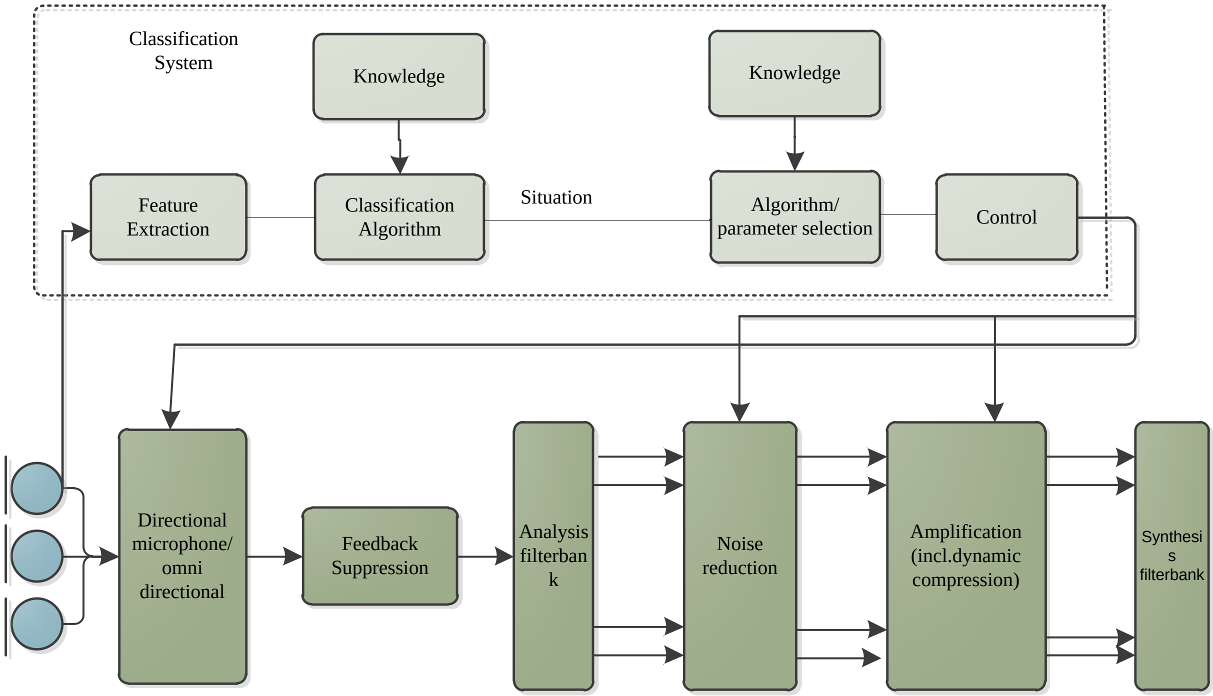


Figure 2: Architecture of Proposed System

3.2 Pre-Processing

In the pre-processing phase, the sound signal undergoes modifications to eradicate external disturbances and noise, ensuring the retention of the primary auditory components.

• Long-term Smoothed Modulation Frequency-based Noise Reduction:

This method segregates the noisy signal into multiple subbands, subsequently applying long-term smoothed attenuation to the subbands exhibiting the lowest average Signal-to-Noise Ratio (SNR). This noise reduction mechanism, a staple in contemporary digital hearing aids, diminishes frequency components with a notably low SNR. Through modulation frequency analysis, the method discerns between subbands encompassing desired signal components and those primarily filled with noise.

The underlying processing algorithm leverages a spectral subtraction technique akin to the mechanisms found in modern hearing aids. The noise level is estimated by harnessing the input from a singular microphone and considering the long-term stimulus average. Components of the input signal that surpass this long-term average are identified as the primary signal.

Various frequency bands are analyzed to calculate the noise's average power level, and based on this, the attenuation rate is determined. Concurrently, the power level of the composite input signal (which amalgamates both the primary signal and the noise) is estimated for analogous frequency bands. By juxtaposing this level with the average level of a noise-only stimulus, an instantaneous running SNR is computed for each frequency channel. This SNR then aids in ascertaining the running attenuation pertinent to each channel. The gain function for this mechanism can be depicted by Equation 1.

$$Attenuation = \begin{cases} -1.55\,SNR + 20.91; & 9dB < SNR < 18\,dB \\ -0.78\,SNR + 14; & -2dB < SNR < 9dB \\ 24; & SNR < -2dB \end{cases} \quad (1)$$

The algorithm's attenuation was capped at a maximum of 24 dB, employing the highest noise reduction level possible. This ensures that while the noise is significantly reduced, the primary auditory components of the signal are not unduly compromised.

Figure 4 provides a comparative representation of time-domain signals before and after the de-noising procedure. This visualization underscores the efficacy of the pre-processing technique, clearly delineating the reduction of noise while preserving the integral components of the auditory signal. Such visual comparisons are instrumental in validating the effectiveness of noise reduction methodologies, demonstrating their utility in real-world applications, especially in the context of hearing aids.

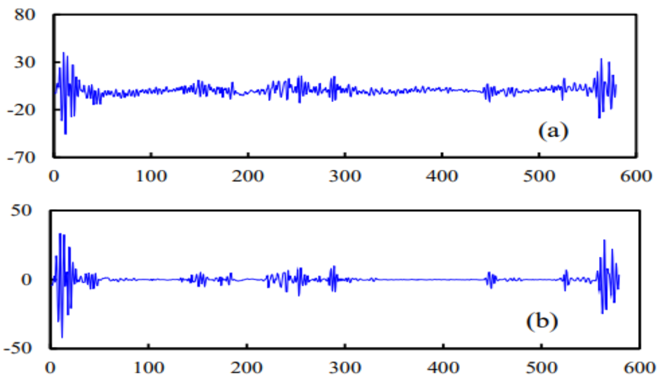


Figure 4. Comparison of time-domain signals before and after signal de-noising (a) Original signal, (b) Noise reduction signal

3.3 Frequency Domain Conversion and Feature Extraction

To glean frequency-associated attributes, the windowed signal transforms a frequency domain representation. A widely recognized technique for procuring the frequency spectrum from a discrete-time signal is the Fast Fourier Transform (FFT). However, executing FFT can be computationally intensive and demanding in embedded systems.

Given the challenges associated with FFT in such contexts, an alternative approach is employed: the filter bank method. This technique facilitates the extraction of the frequency spectrum more efficiently and is suitable for embedded systems. Figure 5 provides a graphical representation of the filters used in the filter bank, illustrating the frequency ranges each filter targets and how they work in tandem to provide a comprehensive frequency domain representation of the signal.

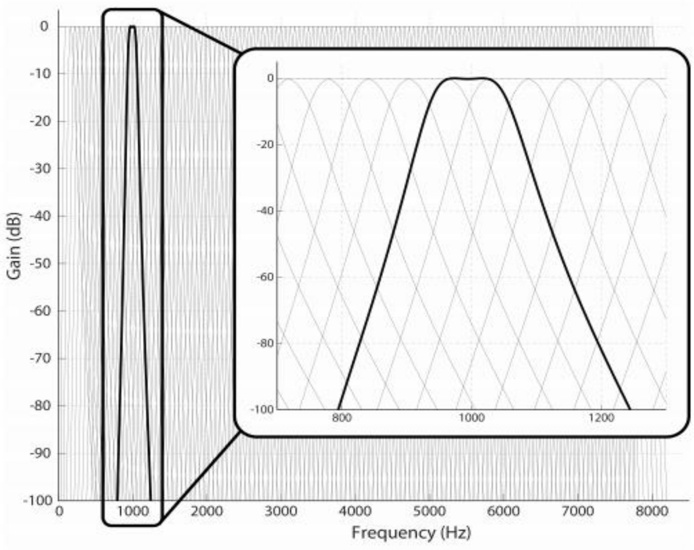


Figure 5: Filter Bank

3.4 Filter Bank Configuration

The filter bank employed in this methodology is composed of 256 bandpass filters. These filters are systematically distributed across 100 Hz to 16 kHz. Each filter within this assembly is characterized as a 4-stage Infinite Impulse Response (IIR) elliptical filter, distinguished by its narrow 4th order.

To ensure efficiency and cost-effectiveness, the filters have been pre-designed and optimized for signals exhibiting a sampling rate of 48 kHz. This preemptive calculation eliminates the need for real-time filter design during signal processing, enhancing the speed and reducing computational overhead.

For feature extraction, each band's root mean square (RMS) is computed once the input signal has been processed through the bandpass filters. This yields the frequency spectrum representation for every segment of the signal. The RMS values serve as a reliable measure of the signal's magnitude within each frequency band, enabling a detailed spectral analysis and facilitating subsequent stages of fault detection.

Time-domain and frequency-domain representations are considered to comprehensively analyze a speech signal, which is replete with information and potential faults. The time domain offers insights into how the signal varies with time, while the frequency domain sheds light on the signal's spectral composition.

In the context of the time domain, specific features can be extracted to provide valuable information about the nature and characteristics of the signal. These features can be instrumental in detecting and diagnosing faults or anomalies within the signal.

1. *Short-time energy* is a low-level temporal feature of a sound explained in equation (2) for continuous signals, where $x(t)$ is the signal's function.

$$E = \int_{-\infty}^{\infty} |x(t)|^2 \quad (2)$$

2. *Root mean squareamplitude* A signal's amplitude has both +ve and -ve values; therefore, utilizing the average of the instantaneous samples to find the signal is problematic using the equation below.

$$E_{rms} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} x_t^2} \quad (3)$$

3. *Amplitude variance* is found by describing the difference in a group of values as differences. The variance for the discrete set is obtained in the equation.

$$E_i(X) = \sum_{1}^{N_w} (x(i) - \mu_x)^2 \quad (4)$$

4. *Zero crossing rate* is explained as the total times that an amplitude amount of a signal switches signs. Zero crossing rate can be expressed using the below equation.

$$zcr = \frac{1}{N} \sum_{t=2}^{N-1} \| sign(x(t) - sign(x(t-1)) \| \quad (5)$$

$$sign\big(x(k)\big) = \begin{bmatrix} 1, if x(i) > 0 \\ 0, if x(i) = 0 \\ -1, if x(i) < 0 \end{bmatrix} \quad (6)$$

Time-domain and frequency-domain features can comprehensively understand a speech signal and its anomalies.

The time-domain features often capture the temporal characteristics of the signal, such as its amplitude, energy, and duration. For instance, fault sounds or anomalies might manifest as sudden spikes, dips in amplitude, or unexpected patterns over time.

On the other hand, frequency-domain features focus on the signal's spectral content, revealing insights about energy distribution across different frequency bands. Specific faults or anomalies might alter the shape or distribution of the frequency spectrum in distinct ways.

1. *A spectral peak* is a scale that finds the frequency range with the maximum energy value. The spectral peak can be calculated efficiently by measuring the index of the maximum value in the spectrum.
2. *The spectral centroid measures* the center of mass of the spectrum. The spectral centroid is expressed as (4.11), where $x\ (I)$ is the weighted value of the frequency band.

$$f_{centroid} = \frac{\sum_{i=1}^{N} f(i)x(i)}{\sum_{i=1}^{N} x(i)} \quad (7)$$

3. *Spectral kurtosis* calculates how relatively the spectrum is the same as the Gaussian distribution. Therefore, it finds the peakedness of the spectra.

$$v(i) = \frac{2\sum_{i=1}^{N} (|x(i,n)| - \mu_x)^4}{f_{max}\, \sigma_x^4} - 3 \quad (8)$$

4. *Spectral spread* calculates how centered the spectra are surrounding the spectral centroid.

$$v_{spread} = \sqrt{\frac{\sum_{i=1}^{N} (i - f_{centroid}\ (x)^2 |x(i)|^2}{\sum_{i=1}^{N} |x(i)|^2}} \quad (9)$$

To discover abnormalities in frequency spectra that do not contain enough energy to modify the entire distribution, we measure the spectral peak, the frequency with the most energy. To identify substantial changes in a frequency distribution, we use the spectral centroid to quantify the spectrum's center of mass. Spectral kurtosis spread and flatness, on the other hand, define the overall form of the spectrum. The characteristics are normalized to a range of zero to one. Normalization guarantees that all features are given equal weight.

3.3 Fault Detection Model

Acoustic Feedback in Hearing Aids

Acoustic feedback is a predominant issue in hearing devices and remains a primary factor contributing to dissatisfaction among hearing aid users.

Acoustic Feedback Mechanism: As depicted in Figure 6, acoustic feedback occurs due to the unintended acoustic coupling between the loudspeaker (or speaker) and the hearing aid microphone. This undesired coupling arises primarily from acoustic leakages. Specifically, when sound emitted by the loudspeaker leaks back to the device's microphone, it can be inadvertently re-amplified.

Consequences: Such feedback disrupts and diminishes the incoming speech signals intended for the hearing aid user. When re-amplified multiple times, this looping of the sound between the speaker and the microphone manifests as a highly unpleasant and often sharp "whistling" or "screeching" noise. This disruptive sound phenomenon is commonly called the "howling effect."

The howling effect not only diminishes the quality of sound perceived by the user but can also lead to discomfort and, in some cases, potentially harm the user's residual hearing. Addressing acoustic feedback is, therefore, paramount in ensuring the optimal performance of hearing aids and enhancing user satisfaction and comfort.

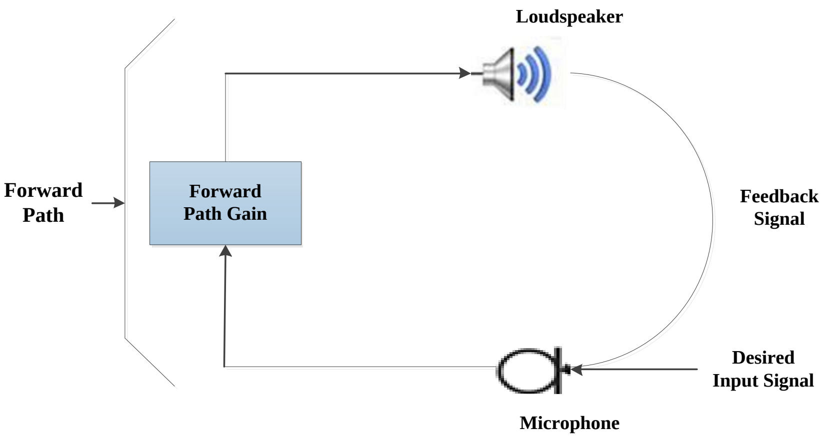


Figure 6: Acoustic coupling between Loudspeaker and Microphone of a Hearing aid device.

Acoustic Feedback Challenges and Solutions in Hearing Aids

The repercussions of acoustic feedback, particularly the howling sound (HS), have profoundly affected the user experience and the technical design of hearing aids. The presence of HS often restricts the application of higher gain values, which might be essential for certain hearing aid users.

As cited from references [4], statistical insights indicate that 10% to 15% of in-the-ear hearing aids are returned to manufacturers due to feedback-related complications within the initial 90 days post-production. This not only inflates the overall cost of hearing aids but also results in significant inconvenience and dissatisfaction among users.

Despite the advancements in hearing aid technology, acoustic feedback remains a complex and pressing research area. The challenge lies in accelerating the solution's convergence rate while ensuring the system's reliability remains uncompromised. The central problem mandates an economical solution that preserves the quality and intelligibility of speech.

Historically, several strategies have been formulated to mitigate the acoustic feedback dilemma. These strategies span across various categories:

1. Phase Modulation Methods: These techniques involve altering the phase of the feedback signal to suppress its effects.

2. Gain Reduction Methods: These methods dynamically adjust the hearing aid's gain to minimize feedback.

3. Spatial Filtering Methods: These techniques spatially filter out the feedback sound by leveraging multiple microphones.

4. Room Modelling Methods: These approaches model the acoustic environment and predict potential feedback paths, allowing for proactive feedback suppression.

In the context of this paper, the focus is primarily on **Filtering Methods** to counteract the detrimental effects of acoustic feedback. One inherent challenge in handling feedback in hearing aids is the correlation between the

Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 4 (2023)

feedback signal and the desired speech signal. This correlation stems from the closed-loop configuration prevalent in these devices.

The filter-based approach in this discourse employs a Finite Impulse Response (FIR) filter to approximate and process the acoustic feedback path. Placing the estimated FIR filter parallel to the actual feedback path negates or " cancels" the feedback, thus restoring the desired speech signal's clarity. This mechanism of feedback cancellation is visualized in Figure 7.

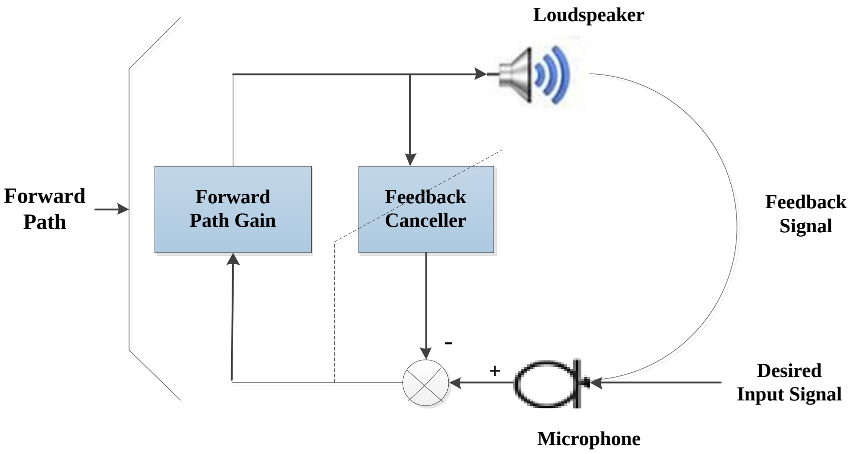


Figure 7: Filter-based Method for feedback Cancellation.

3.3.1 Fault Identification using SVM

The **Support Vector Machine (SVM)** is renowned for its ability to discern patterns in complex datasets, making it an ideal choice for tasks such as fault identification in hearing aids.

Basics of SVM:

SVM operates on the principle of determining the optimal hyperplane that best differentiates between classes in a dataset. This optimal separation is achieved by maximizing the margin between the classes. The architecture and phases of the SVM Classifier, as depicted in Figure 8, underscores its structured approach to learning and classification.

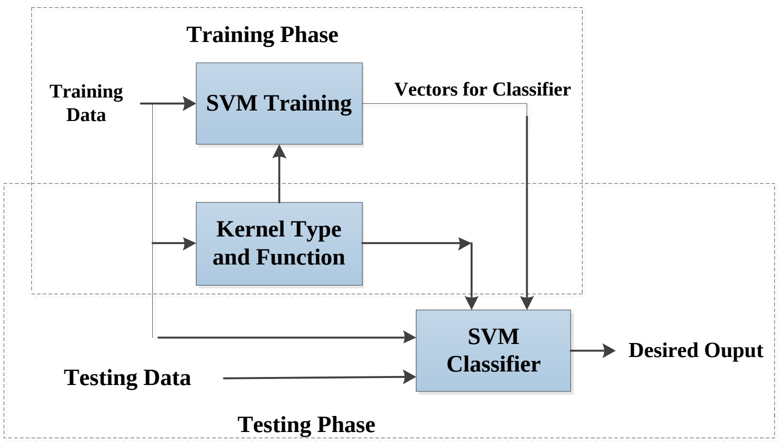


Figure 8: Architecture of SVM Classifier.

Critical Components of SVM:

1. Maximal Margin: Reducing the VC (Vapnik–Chervonenkis) dimension minimizes the SVM's upper bound, enhancing its generalization capability.

2. Kernel Trick: This strategy allows the SVM to transform the problem's dimensionality, making the estimation function more adaptive to the data. The empirical error risk diminishes by elevating the issue from a lower to a higher dimension.

3. Sparseness: Fewer support vectors (SVs) enhance the system's generalization capability. The sparsity of SVs also ensures computational efficiency since the feature set comprises these SVs.

4. Convex Optimization: The optimal solution of SVM is derived from a quadratic optimization model. The convex nature of this formulation ensures a unique and global solution.

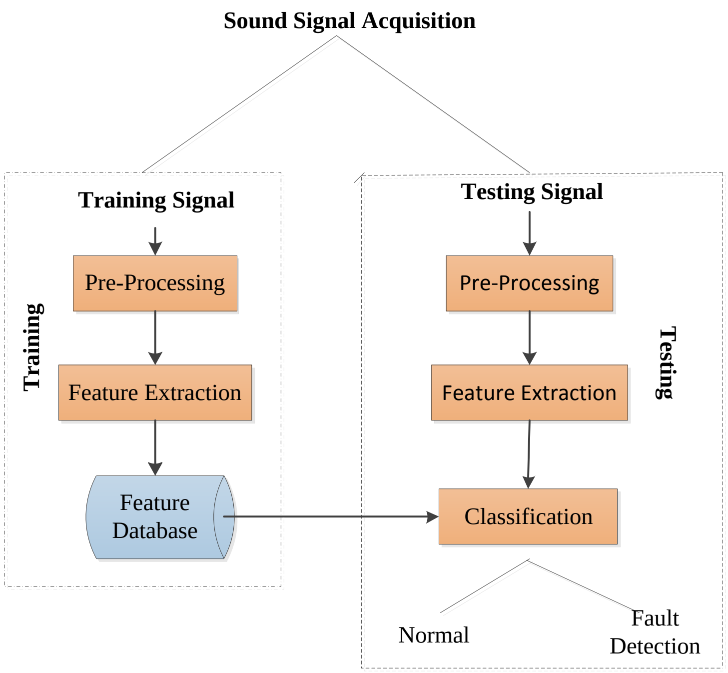


Figure 9: Flowchart detailing the training and testing procedure of SVM.

Fault Detection Process:

The features extracted from the speech signal mirror the operational state of the hearing aid. The SVM classifier is trained using these features to distinguish between standard and faulty hearing aids.

The SVM's decision parameter, pivotal in this differentiation process, is optimized based on the training data. Once trained, the classifier assesses new data: the hearing aid is deemed functional if the decision parameter falls within a predetermined boundary. Conversely, a fault is likely present if it lies outside this boundary.

IV. Experimental Results And Analysis

In this section, the hardware device in focus receives a detailed overview, followed by an elucidation of the implementation specifics. Hearing devices undergo classification based on shape and functionality into behind-the-ear, receiver-in-canal, in-the-ear, in-the-canal, completely-in-canal, and invisible-in-canal—figure 10 displays images representing various formats of available hearing aid devices. For this research, the experiment utilizes Shamanics Axon K 80 ITE and Shamanic Hearing Aid AXON V 185 Ear Machines, with sample images presented in Figure 11.

In-the-ear hearing aids are designed to fit entirely within the outer ear and address a range from minimal to profound hearing loss. Two formats or sizes exist: one encompassing the outer ear (entire shell) and another covering only the lower portion of the outer ear (half shell).

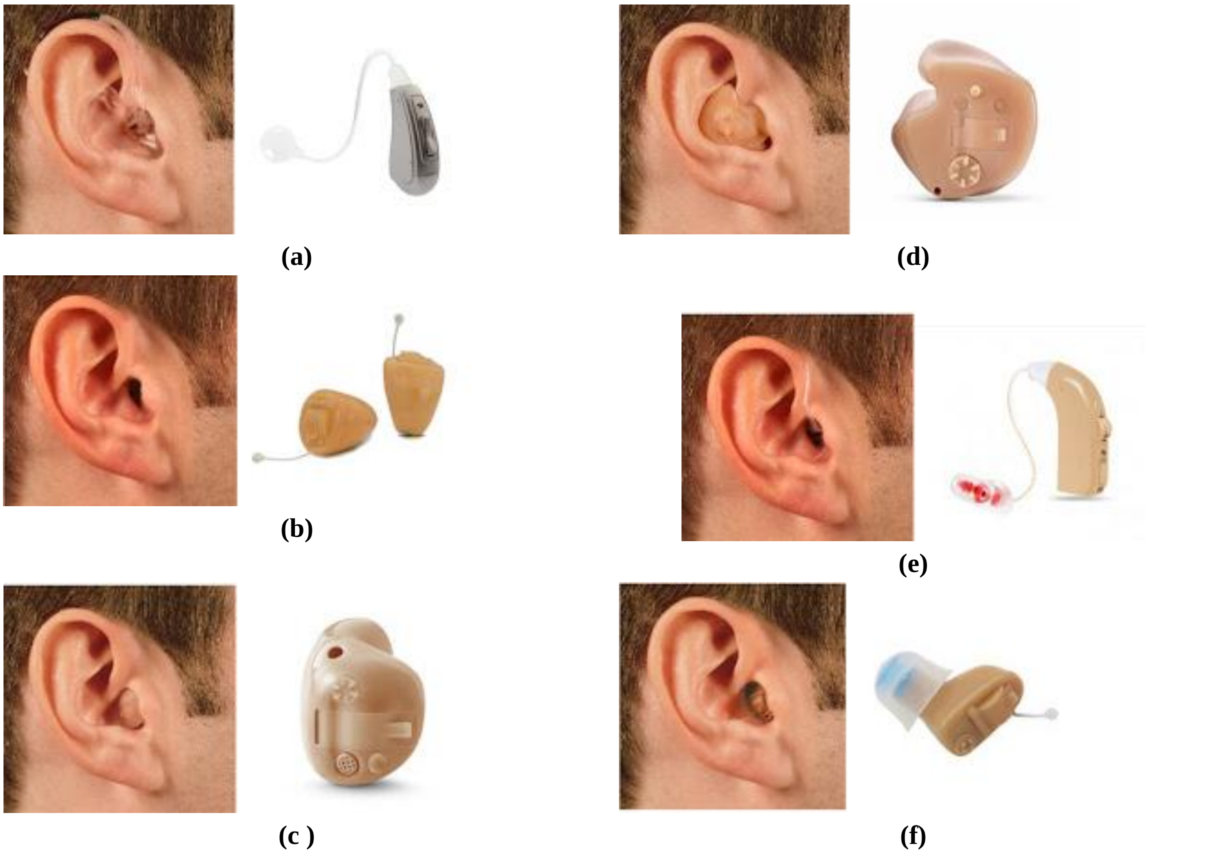


Figure 10: (a) behind-the-ear (BTE) (b) invisible-in-canal (IIC) (c) in-the-canal (ITC) (d) in-the-ear (ITE) (e) receiver-in-canal (RIC) (f) completely-in-canal (CIC)

Figure 11: In-the-ear hearing aids used in the proposed Work

Signal data procured from the hearing aid device undergoes initial preprocessing before signal classification. Noise elimination occurs within this signal data. Subsequently, the SVM classifier aids in determining the categorization of the input signal, distinguishing between non-faulty and faulty classifications.

4.1 Dataset description

A database formed from speech signals sampled by 10 individuals at a rate of 16 kHz, translating to 1600 samples per second. Each of these samples boasts a resolution of 16 bits. Near-end and far-end speakers were selected independently, culminating in aggregating 50 speech signal files. To achieve a total of 50 sets, three distinct sound sources, specifically far-end speaker, near-end speech, and noise, were generated at random positions with distances of 1.5 m, 1 m, and 2 m from the microphone array. Thirty of these sets served the learning process, with the remaining twenty designated for testing.

For near-end speech, signal-to-echo ratio levels, encompassing [-6 dB, -3 dB, 0 dB, 3 dB, 6 dB], were randomly selected and combined with the adjusted acoustic echo sign. Figure 12 illustrates a speech representation plot derived from the database's speech signal.

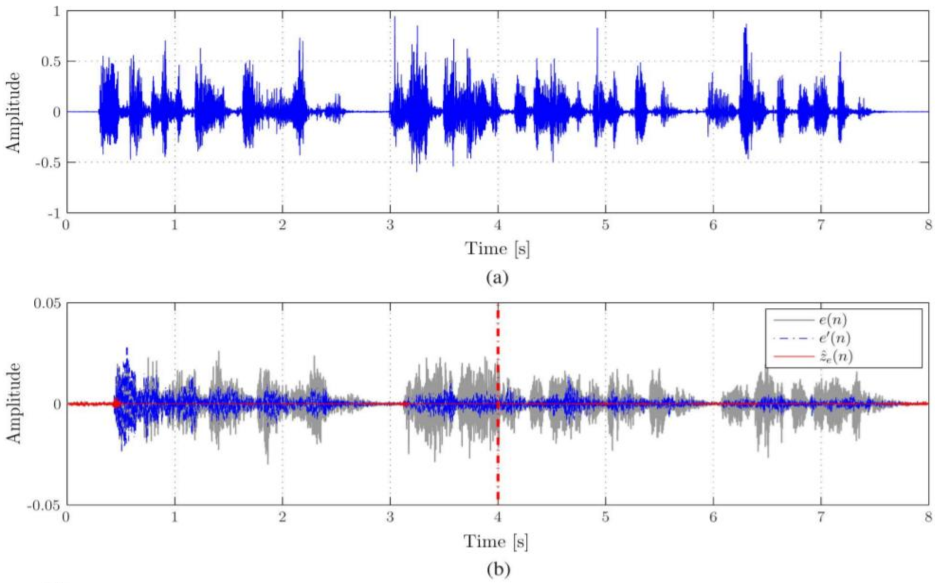


Figure 12: Sample Input Speech Signal (a) Speech Signal (b) Speech Signal after Acoustic Echo.

The results derived from the system's implementation receive a step-by-step elucidation. Figure 13 showcases a sample speech signal subjected to the proposed methodology. Initially, the pre-processing technique, represented in Figure 14, depicts the example speech signal before applying the pre-processing measure. Following this, Figure 15 reveals the outcome of the speech signal post-denoising.

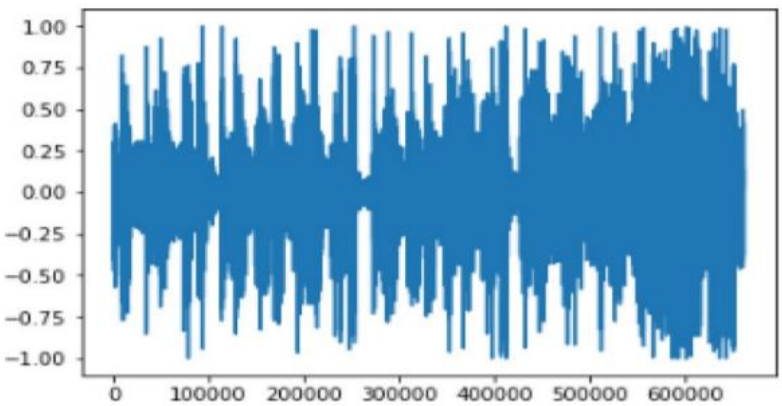


Figure 13: Sample of a Speech Signal

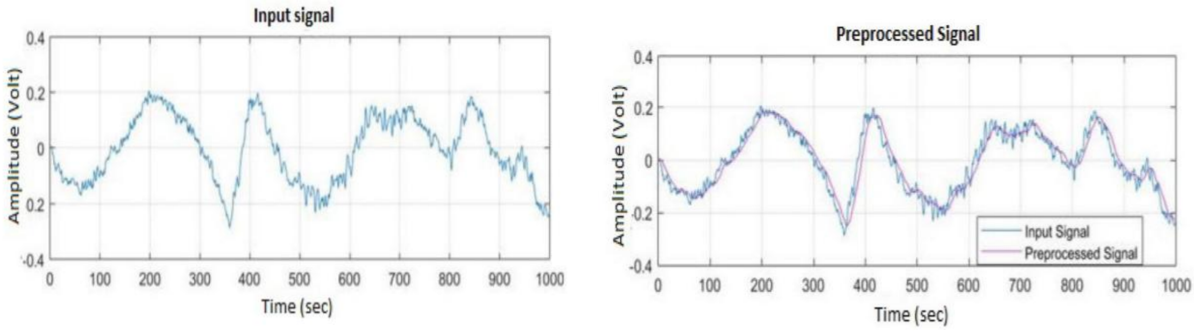


Figure 14: (a) Input Speech Signal (b) Pre-Processed Speech Signal

In Figure 14 b, the speech signal undergoes pre-processing and background noise is segregated from the speech signal data. The subsequent feature extraction from the time domain is illustrated as follows: Figure 15 represents the extracted ZCR, STE, spectrogram, and bandwidth from the pre-processed speech signal, achieved by employing the proposed approach.

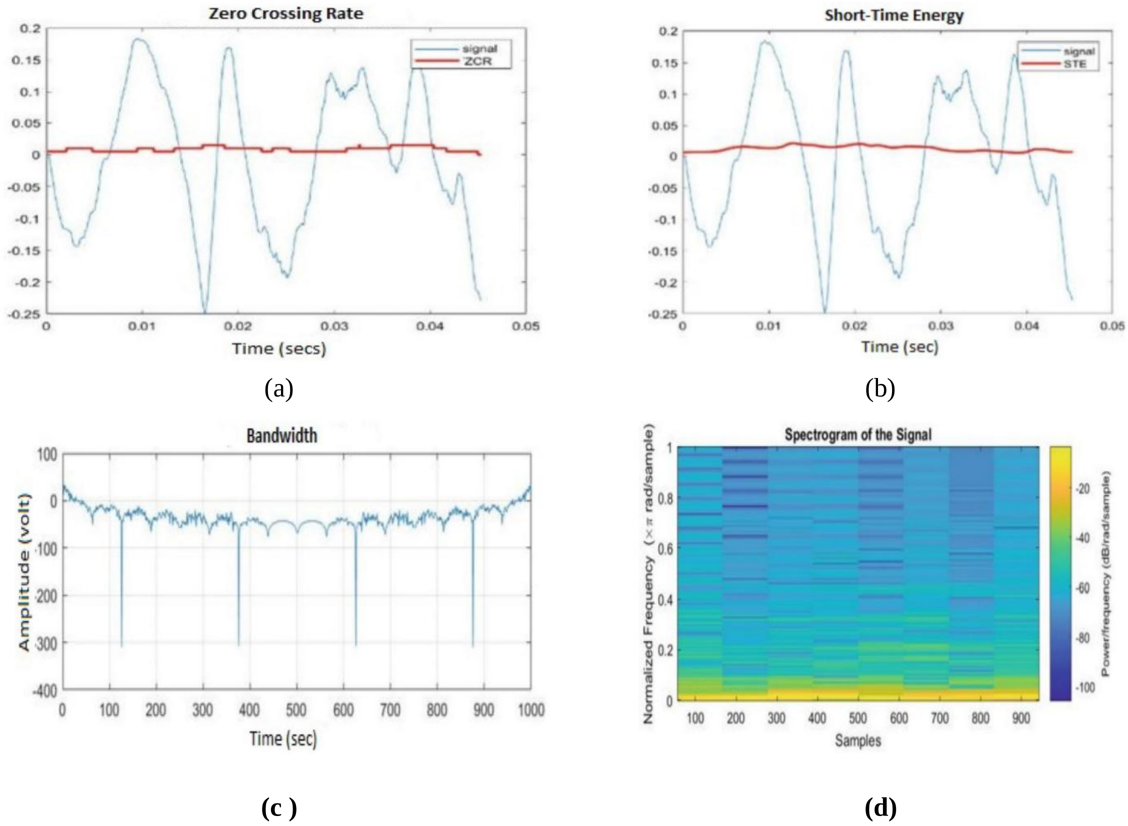


Figure 15: Results of Feature Extraction (a) Zero Crossing Rate (b) Short-Time Energy (c) Bandwidth (d) Spectrogram.

Following the feature extraction process, the SVM characterization undergoes an update. This SVM computation works to segregate the hyperplane based on designated constraints. Within this context, the faulty speech signal receives a label of 0, while the non-faulty speech signal is designated with a label of 1. Figure 16 presents the hyperplane plan diagram after the finalization of the proposed descriptive analysis.

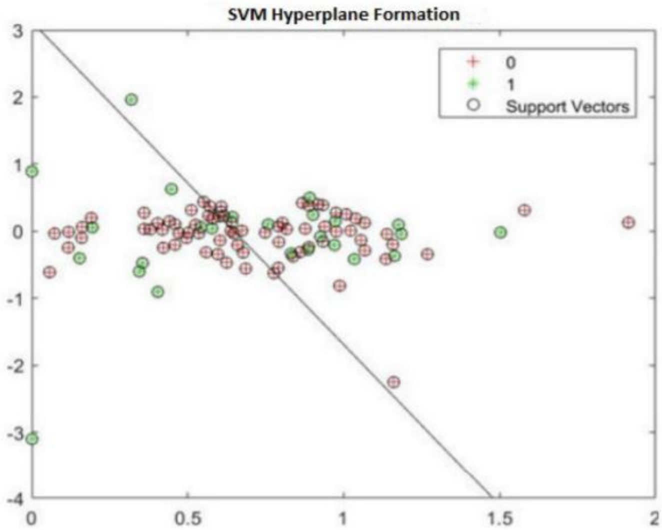
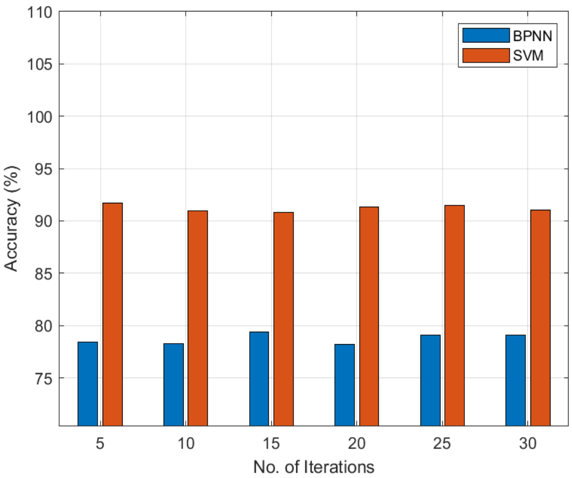


Figure 16: Proposed SVM Hyperplane formation graph.

4.2 Performance Analysis

Several metrics, such as accuracy, precision, sensitivity, and specificity, were employed to evaluate the presented model's performance. Table 1 outlines the results. These results highlight the peak value of quantitative metrics up to a specific epoch number. Subsequent illustrations compare Accuracy (Figure 17), Precision (Figure 18), Recall (Figure 19), and F-measure (Figure 20) between the proposed SVM classifier and BPNN classifiers.



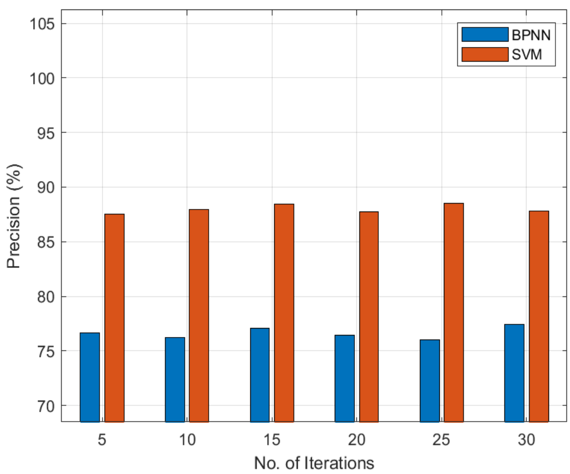
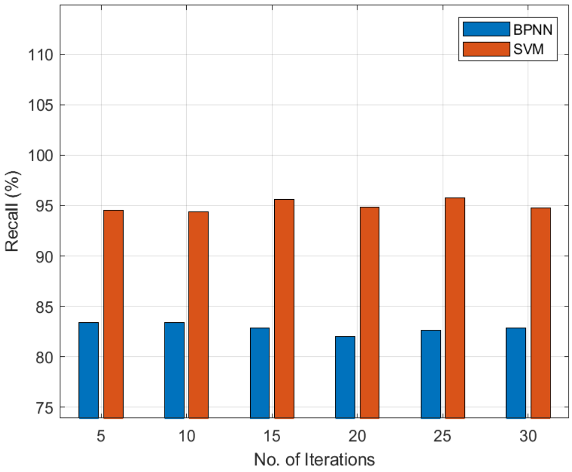


Figure 17: A Graph of Accuracy with Comparison Figure 18: A Precision graph with a comparison



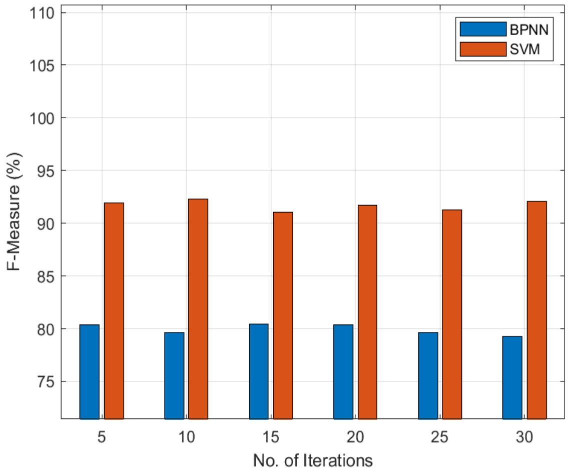


Figure 19: Recall graph with comparison. **Figure 20: A Graph of F-measure with comparison**

Table 1. Comparison of different Classification Algorithms

Parameter	BPNN Classifier	SVM Classifier
Accuracy	78.106	90.023
Recall	82.030	94.064
Precision	75.927	87.038
F-Measure	78.847	90.386

Figure 21 illustrates the model's convergence speed by plotting training and testing efficiency against losses over epochs.

The confusion matrix, sometimes called the error matrix, provides a visual representation of the performance of a machine learning or statistical classification algorithm. In this matrix, rows symbolize predicted ground truths, while columns indicate actual ground truths. Figure 18 depicts the recognition rate graph for the SVM Classifier, and Figure 13 presents the confusion matrix for the discussed model, focusing on the two classes.

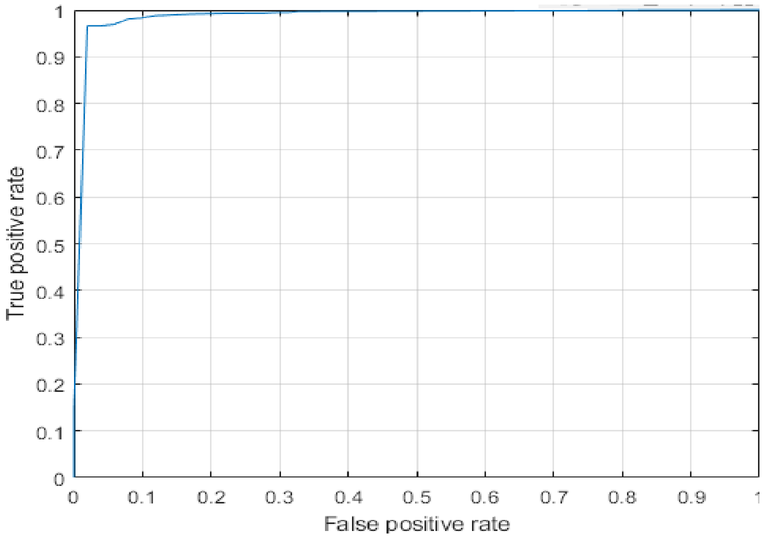


Figure 21: ROC for SVM Classification

V. Conclusion

A preferable scenario involves predicting machine malfunctions and initiating preventative maintenance to ensure machines operate safely and deliver optimal performance. The suggested fault detection approach incorporates both time and frequency domains. The filter bank comprises 256 bandpass filters from 100 Hz to 16 kHz. Each filter constitutes a 4-stage Infinite Impulse Response (IIR) elliptical filter with an expansive 4th order. These filters were pre-computed for signals at a 48 kHz sampling rate to enhance efficiency and cost-effectiveness. After bandpass filtering, the root mean square for each band is computed, producing the frequency band for each segment in the feature vector. This vector characterizes sound in diverse manners and proves instrumental in identifying various anomalies. An experimental evaluation of the proposed strategy was conducted, and based on the outcomes, the method demonstrates success in pinpointing defects in hearing aid devices.

References

[1] Karunakar, Y., Kuwadekar, A., and Al Begin, K. A mobile-based application for detecting fault by sound analysis in car engines using triangular window and wavelet transform. In Computational Intelligence and Communication Networks (CICN), 2010 International Conference on (Nov 2010), pp. 523–528.

[2] Maniak, T., Iqbal, R., Doctor, F., and Jayne, C. Automated sound signaling device quality assurance tool for embedded industrial control applications. In Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on (Oct 2013), pp. 4812–4818.

[3] Hayashi, S., Asakura, T., and Zhang, S. Study of machine fault diagnosis system using neural networks. In Neural Networks, 2002. IJCNN ’02. Proceedings of the 2002 International Joint Conference on (2002), vol. 1, pp. 956–961.

[4] An approach to fault diagnosis of vacuum cleaner motors based on sound analysis. Mechanical Systems and Signal Processing 19, 2 (2005), 427 – 445.

[5] Sound-based induction motor fault diagnosis using Kohonen self-organizing map. Mechanical Systems and Signal Processing 46, 1 (2014), 45 – 58.

[6] Ai, C., Zhao, H., Ma, R., and Dong, X. Pipeline damage and leak detection based on sound spectrum lpcc and hmm. In Intelligent Systems Design and Applications, 2006. ISDA ’06. Sixth International Conference on (Oct 2006), vol. 1, pp. 829–833.

[7] Alameh, K., Cite, N., Hoblos, G., and Barakat, G. Feature extraction for vibration-based fault detection in permanent magnet synchronous motors. In Technological Advances in Electrical, Electronics, and Computer Engineering (TAEECE), 2015 Third International Conference on (April 2015), pp. 163–168.

[8] Anami, B., and Pagi, V. Acoustic signal-based detection and localization of faults in motorcycles. Intelligent Transport Systems, IET 8, 4 (June 2014), 345–351.

[9] Bin Mansoor, A., Ahmed, H., and Mahmood, Z. Design and develop an automated acoustic-based jet engine performance evaluator. In Industrial Technology, 2009. ICIT 2009. IEEE International Conference on (Feb 2009), pp. 1–4.

[10] Ferraz, I., and Garcia, A. Turbomachinery failure prognostics. In Modern Advances in Applied Intelligence, M. Ali, J.-S. Pan, S.-M. Chen, and M.-F. Horng, Eds., vol. 8481 of Lecture Notes in Computer Science. Springer International Publishing, 2014, pp. 349–358.

[11] Gong, X., and Qiao, W. Current-based automatic fault detection for direct-drive wind turbines via synchronous sampling and impulse detection. Industrial Electronics, IEEE Transactions on 62, 3 (March 2015), 1693–1702.

[12] Guan, X., Shi, Z., Gu, F., and Ball, A. Cabin noise fault detection using deviation energy method in cross time and frequency domain. In Image and Signal Processing (CISP), 2010 3rd International Congress on (Oct 2010), vol. 8, pp. 3765–3769.

[13] Zubair S, Yan F, Wang W (2013) Dictionary learning-based sparse coefficients for audio classification with max and average pooling. Dig Signal Process 23(3):960–970

[14] Lavner Y, Ruinskiy D (2009) A decision-tree-based algorithm for speech/music classification and segmentation. EURASIP J Audio Speech Music Process 1:239892

[15] Zahid S, Hussain F, Rashid M, Yousaf MH, Habib HA (2015) Optimized audio classification and segmentation algorithm using ensemble methods. Math Probl Eng 1–11

[16] Bhaskar J, Sruthi K, Nedungadi P (2015) Hybrid approach for emotion classification of audio conversation based on text and speech mining. Proc Comput Sci 46:635–643

[17] Dhanalakshmi P, Palanivel S, Ramalingam V (2011) Pattern classification models for classifying and indexing audio signals. Eng Appl ArtifIntell 24(2):350–357

[18] Lee H, Pham P, Largman Y, Ng AY (2009) Unsupervised feature learning for audio classification using convolutional deep belief networks. In: Advances in neural information processing systems, pp 1096–1104

[19] Chen L-T, Wang M-J, Wang C-J, Tai H-M (2006) Audio signal classification using support vector machines. In: International symposium on neural networks, pp 188–193

[20] Wang J-C, Wang J-F, Lin C-B, Jian K-T, Kuok W (2006) Content-based audio classification using support vector machines and independent component analysis. In: 18th International conference on pattern recognition (ICPR'06), no. 4 pp 157–160

[21] Temko A, Nadeu C (2006) Classification of acoustic events using SVM-based clustering schemes. Pattern Recognit 39(4):682–694

[22] Scardapane S, Uncini A (2017) Semi-supervised echo state networks for audio classification. CognComput 9(1):125–135

Bibliography

Vijay Kumar Reddy I, Research Scholar in the Department of E&CE at JSSATE, Bangalore, received a B.Tech degree in Electronics and Communication Engineering from B.I.T. Institute of Technology, Hindupur. An M.Tech degree in VLSI Design was obtained from Sreenivasa Institute of Technology and Management Studies, Chittoor, affiliated to JNTUA, Anatapur. Previous positions include a role as Assistant Professor in the Department of Electronics & Communication Engineering at Acharya Institute of Technology, Bangalore. Authored books include "Digital Electronics" and "Linear ICs and Applications". Research interests encompass Fault Tolerance in Hearing Aids, VLSI, and Signal Processing.

Dr. Aravinda H. S. holds the position of Professor in the Department of Electronics & Communication Engineering at JSS Academy of Technical Education, Bangalore. This scholar completed a B.E. Degree in Electronics & Communication Engineering from Bangalore University in 1996. Post-graduation was achieved from Mysore University in 1999. A Ph.D. on the subject of "Fault tolerance Measurements Techniques of Acoustic Spectrum" was awarded by VTU during 2012-2013. Primary research interests include Fault Tolerance, Signal Processing, and Image Processing. With 24 years of experience in teaching engineering for UG & PG courses, more than 50 papers have been published in both international and national journals and conferences. Additionally, Dr. Aravinda H. S. reviews for six national and international journals.