CHAPTER 2
REVIEW OF LITERATURE

The literature review presents a survey of the acoustic analysis for human voice disorder classification using optimization and machine learning techniques with importance to pathology detection and its classification. The voices input signal is given to various classifications, in such a way the classification produces the result against the pathology voices and normal voices with respect to the male and female human voice. In preprocessing, the Noise Removal and Silence Removal were taken by Electro Glotto Graph (EGG) with the help of Wiener Filter and DWT Filters. For this voice filtration, this research used the Hybrid Wiener Filter Discrete Wavelet Transforms (HWFDWT). This HWFDWT consist of two processes such as "Wiener Filter Minimization" as well as "Discrete Wavelets Sample". This literature survey shows the previous techniques taken against the voice filtration are relevant to the Hybrid Wiener Filter Discrete Wavelet Transforms (HWFDWT). In feature selection and extraction process the InfoGain, Correlation, and Principal Components Analysis are evaluated by using Wavelet Thresholding Algorithm. This helps to reduce Dimensionality.

The Feature Extraction is done by Cat Swarm Optimization Mel Frequency Cepstrum Coefficients called (CSOMFCC). It is used to extract the best features from the disorder voice signal. This research survey shows previous work done in contrast to the Mel Frequency Cepstral Coefficients (MFCC) and Cat Swarm Optimization (CSO). This helps to improvise the research outcome on the classification of the voice signal. The classification is the process of grouping the normal voices and pathological voices of the input voice signal. For classification, the proposed new system called Modified Optimized Back Propagation Network Disorder Voice Classification (MOBPNDVC) were used to classify the input voice signal with the help of Support Vector Machine (SVM) and Back Propagation Neural Network (BPNN). This survey shows that voice disorder detection is previously done by the SVM and BPNN. Moreover, in this research, the Specificity, Sensitivity, Accuracy and the Execution Time of each process were examined optimized with respect to the proposed Modified Optimized Back Propagation Network Disorder

Voice Classification (MOBPNDVC). Finally, the Receiver Operating Characteristic (ROC) Curve is generated to compare diagnostic tests and also a plot of the true positive rate against the false positive rate. A ROC plot shows the relationship between sensitivity and specificity. The following discussed the acoustic analysis for human voice disorder classification and optimization techniques on the Saarbruecken data set and the own Real-time dataset by various researchers.

2.1 Parametric Representation by preprocessing

Harar et al (2017) discussed the voice disorders and used Deep Neural Networks (DNN) for voice pathology detection. The research uses the German corpus Saarbruecken Voice Database with recurrent Long-Short-Term-Memory (LSTM) layers for voice pathology detection. It is found there is high accuracy with sensitivity and specificity on testing files and training data set. This system achieved an efficient time to find healthy persons and pathological patients. Finally, the development of Voice Pathology Detection system is a hard problem in the whole dataset with accuracy. This research motivates us to use Saarbruecken Voice Database for pathology detection with the various QoS parameters.

Zhang et al (2008) investigated the acoustic characteristics of pathologies. It concentrates on the vowels in terms of sustain as well as running. The Perturbation methods, SNR and nonlinear dynamic methods are used to investigate characteristics of pathologies. And classify as low-dimensional rate as the laryngeal pathologies and high-dimensional rate as normal voice. Finally, the research is performed by nonlinear dynamic analysis for sustain and running vowels in processing voice analysis. This helped to manage pathology voice detection with respect to the perturbation analysis.

Parsa et al (2001) explored the Pathological Voice investigation. It differentiates the normal and pathological voices with respect to the acoustic measures. The sustained vowels and continuous speech for acoustic measurements are taken. the speech signals are extracted from sustained vowels and continuous speech, with respect to the fundamental frequency, amplitude perturbation, spectral measures, and glottal noise. The result shows, normal voice, and pathological voice classifications are done using the acoustic measures with the higher isolation and classification on continuous speech. This research motivated us for the Pathological Voice investigation on the acoustic measures.

Pribuisiene et al (2006) proposed the Acoustic Characteristics and Perceptual Characteristics of Voice Changes in Reflux Laryngitis Patients. The study was to concentrate the multidimensional perceptual, subjective and instrumental acoustic measures of voice data. The working on the detection of laryngitis (RL) patients alone is carried out. The comparative analysis is measured with hoarseness scale, the patients having less grade value with respect to the Jitter, Shimmer, Noise Energy, Voice Handicap Index, and Phonautogram are analyzed. This research results demonstrated us that multidimensional voice assessments are a very useful method for deteriorated voice quality capabilities checking of RL patients.

Uloza et al (2005) examine the acoustic voice discrimination of normal and abnormal voices with respect to the acoustic pathological voice measurements. The pitch and glottal noise for detecting the patients as the pathological or normal group are also taken. The measurement of Normalized Noise Energy (NNE) is high then the patients are with pathological, whereas measurement of Normalized Noise Energy (NNE) is low the patients are normal. Finally, the optimum parameter of perceptual and acoustic analysis shows that the normal speakers detected with respect to the documentation of the patient's pathology voice. This research helps to detect the pathological or normal group with respect to the acoustic and perceptual measures.

Hirano et al (1988) analyzed the acoustic parameters of pathological voice. The different types of patients were acoustically analyzed with respect to the vocal fold and nerve paralysis. The acoustic analysis was conducted with respect to the pitch perturbation, amplitude perturbation, and noise energy. This research analyses the diseases such as carcinoma, hoarseness, and stroboscopy. This analysis focuses on the Pitch, Amplitude Perturbation, and Noise Energy. It allows to study the acoustical measures such as hoarse voices, vocal fold vibration, turbulent noise pathologies. This research motivates us to study the production of the acoustic parameters against the pathology disease.

Verde et al (2018) proposed the m-Health system for the voice health state classification, these mobile-based health systems have a huge reach to make patient care faster, better and cheaper many pathological conditions by using Support Vector Machine algorithm or the Decision Tree. The voice pathology detection was also done on the dataset of Saarbruecken Voice Database. The higher Accuracy, Sensitivity, Specificity and ROC area features are evaluated and achieved. This research helps to perform the efficient feature selection and feature extraction process on the dataset.

Silva et al (2009) proposed Jitter Estimation Algorithms for Detection of Pathological Voices detection. The evaluation of jitter value in quasi-periodic signal for pathologies in the vocal fold nodules or a vocal fold polyp is focused. The jitter estimation algorithm to detect pathological voices on the MEEI database is used. Finally, a percentage of the glottal cycle is displayed, but with absolute jitter values are measured. Also, the new Short Time Jitter Estimation (STJE) algorithm measure for jitter better than the commonly used tools.

Martens et al (2007) worked on the perceptual evaluation of pathological voice quality. A simple test for improving consistency speech (Spectrogram) is proposed. Also, a rating of Grade, Roughness, and Breathiness affect the correlation acoustical speech and perceptual speech parameters are considered. Also, the improvement in the Reliability, Feasibility, and Availability of acoustical analysis on the voice quality is evaluated.

Niedzielska et al (2001) proposed the acoustic analysis of voice with respect to the acoustic tests. The Acoustic parameters of voice pathology detection are taken. The result is evaluated by the acoustic analysis by using jitter and shimmer. This jitter and the shimmer was justifying pathological changes instead of normal values. Finally, the acoustic analysis of voice might be considered for the research as an input voice signal to the pathology detection.

Moon. J et al (2018) proposed the modern feature extraction method using voice signal. The proposed extraction method comprises the Higher-Order Statistics (HOS) and higher-order Differential Energy Operator (DEO). This system used to improve the medical diagnosis system. To accommodate the healthy voice and pathological voice on the Saarbruecken Voice Database (SVD) are taken. The report produces the pathological voice of Cysts, Paralysis, and Polyp is detected. Finally, this research compares the accuracy features extraction and the possibility of pathological voice occurrences efficiently on the neural networks. This work helps to improve the feature extraction accuracy of the system.

Borovikova et al (2018) classify the voice quality Acoustic analysis with the help of laryngeal pathology recognition and classification to detect the pathology by an abnormality of the vocal tract. The abnormality of the vocal tract is directly proportional to the voice quality. This proportionality is compared with the individual acoustic characteristics of people voice quality. The determination result shows the efficient pathology voices deduction by using Normalized First Harmonic Energy (NFHE), Voice Sonority Coefficient (VSC), Voice Harmonization Coefficient (VHC), Harmonic to Noise Ratio (HNR), Shimmer and Jitter.

Mesallam et al (2017) discussed voice pathology evaluation by using speech features and machine learning algorithms. The reduction of Ethnicity is concentrated. Also, the research focuses on voice disorder detection and classification with the help of characteristics measures. The results show the voice disorders measures in terms of Accuracy, Reliability, and Perceptual Severity. The detection and classification of voice disorders on the basis of vowel and the running speech are also compared. This research motives us for pathology evaluation efficiency calculation with respect to Accuracy and Reliability.

Saenz-Lechon et al (2006) discussed the issues in the development of automatic systems for voice pathology detection. The automatic detection of voice pathology uses the cross-validation strategy for final confusion matrix creation. The Detector Error Trade-off (DET) and Receiver Operating Characteristic (ROC) plots are used to measure the efficiency of detection of voice pathology. The averaged classification ratio used to detect the different folds of tests to improve the pathological voice detection and classification using the automatic systems for voice pathology detection. This system helps to develop the Automatic Voice Pathological Identification System.

Godino-Llorente et al (2006) concentrate on the Acoustic analysis for larynx cancer with respect to the voice diseases like unhealthy voice. The detected voice signal is rectified by using voice alterations. The proposed method contains the Gaussian Mixture Models and Short-Term Mel-cepstral parameters analysis. In this model, the frame energy plays a vital role in the detection of laryngeal pathology voice. This research helps to make input voice signal-dependent to running speech performance maintenance.

De Oliveira et al (2000) estimate the Residue Signal for Voice Pathology Diagnosis. The proposed accurate diagnose of diseases used the noninvasive method to estimate the larynx and vocal tract. The different pathologies voices are removed using time-invariant inverse filtering. The time-invariant inverse filtering extracts the different features relevant to laryngeal diseases. Finally, the pathologies characteristics from the Mann–Whitney test are discriminated. This research motivates to voice pathology diagnosis with respect to the normal and abnormal speech signal.

Bhuta et al (2004) determined the noise parameters of the Multi-Dimensional Voice Program (MDVP) acoustical analysis system with the perceptual rating system. This system is used to reduce the noise quality of the affected input voice signal. The output produces the standard, reliable, valid and consistent measure against the voice pathology. The Voice Turbulence Index, Noise Harmonic Ratio and Soft Phonation Index for Roughness detection and reduction of Breathiness detection improvement are used.

2.2 Feature Selection and Extraction

Saeedi et al (2011) proposed the Digital analysis of pathological voices. The abnormalities of the vocal system are investigated. The non-invasive tool is used to deduct the equality of the voice by using the wavelet-based method. These methods are producing the final result of normal and disordered voices. The Support Vector Machine algorithm is used for feature extraction and classification on Orthogonal filter banks. The final classification rate is producing the normal and pathological voices on the databases. Also, the genetic algorithm is implemented to search for the best classification rate. Finally, the proposed algorithm can classify the disordered signals with an excellent choice for pathological voices.

Fezari et al (2014) discussed the acoustic voice analysis for unhealthy voice diseases voice abuse deduction. And also investigate the voice pathologies detection voice diagnostic. The healthy and pathological voices, nonneurological from the German database which contains many diseases are classified. Also, the supervised algorithm is used for voice pathologies detection. To accomplish this task, Mel Frequency Cepstral Coefficients are modeled by a weighted Gaussian Mixture Model (GMM) and plotted for features extraction. This pathological voices detection uses the acoustic voice analysis, with automatic speaker recognition techniques. This helps to find the MFCC, frequencies, and amplitudes. This helps to classify the accurate pathological voices. Finally, this research helps to detect normal and abnormal voices with respect to the MFCC and Energy-Optimal system.

Villa-Canas et al (2012) discussed the functional disorders and laryngeal pathologies. Analyzed those disorders by using vibrational patterns of the vocal folds. The non-parametric cepstral coefficients in Mel filter bank scales were used for automatic feature selection. The relevant features are automatically selected using Principal Components Analysis (PCA) and Sequential Floating Features Selection (SFFS). Voice is healthy or pathological was decided, with the help of linear and quadratic Bayesian, K nearest neighbors and Parzen methods.

Verde et al (2018) studied the detection, monitoring, and treatment of pathological health condition. And proposed the new mobile health system for pathological conditions specially dysphonia Voice disorders. This application also supports voice pathology detection and identification of pathological and healthy voices accurately. Also, the machine learning techniques for voice pathology detection on Saarbruecken Voice Database is investigated. The system returns better Accuracy, Sensitivity, Specificity and ROC area by Support Vector Machine algorithm. This study gives support for voice pathology identification by using machine learning techniques. The Support Vector Machine, Decision Tree, Bayesian Classification, Logistic Model Tree and Instance-based

Learning are used to focus on identifying appropriate voice signal on the Saarbruecken Voice Database. This results accuracy in voice pathology detection and classification of voice pathology are compared with the proposed research system accuracy parameters.

Souza et al (2015) discussed the detection of vocal fold pathologies through speech signal analysis in a pathological voice detection system. The quality of the voice against the laryngeal pathologies is estimated. The researcher proposed the binary Particle Swarm Optimization (PSO) using Multilayer Perceptron (MLP) neural network for features selection. This proposed technique investigates the healthy signals and pathological signals classification by using the two-dimensional wavelet coefficients of the speech signals.

Salhi et al (2013) suggest a new method Mel Frequency Cepstrum Coefficients using Fisher discriminant ratio to improve the performance features selection. These methods are combinedly used in the classification of pathological and normal voices. The Multi-Layer Perceptron (MLP) classifier is with Feed Forward Back Propagation training algorithm (FFBP). The mixed voices database was used on the feature selected and the classification rate. Finally, the proposed system helps in feature selection of acoustic features and classification of pathological and normal voices in an efficient manner. MFCC, PLP, and RASTA-PLP results the identification of pathological voices are compared with the proposed pathological classification method, the average improvement of the deduction and classification of pathological were higher in proposed research on voice pathological identification and classification.

Al Mojaly et al (2014) achieve the Automatic Speech Recognition (ASR) mechanism for voice extracted with high accuracy. The different acoustic features for feature extraction are used. The Mel Frequency Cepstral Coefficient (MFCC), Linear Prediction Cepstral Coefficients (LPCC), and Relative Spectra are modified and created the Relative Spectra-Perceptual Linear Predictive (RASTA-PLP). These techniques are applied to MFCC by using the t-test, Kruskal-Wallis test, or Genetic Algorithm (GA). This classification was done using the Support Vector Machine (SVM) Gaussian Mixture Model (GMM). This classification mechanism is also called the multi-class pathology classification. The experimental results on the MEEI subset database give ideas about the high accuracies mechanism to the research work.

Markaki et al (2009) proposed the pathology detection and classification system. The research focuses on the Modulation Spectra. And also the Higher Order Singular Value Decomposition (SVD) and SVM with a Radial Basis Function (RBF) for feature selection are used. The kernel as a classifier of SVM with a radial basis function (RBF) is used for the classification of healthy and pathological voices with pathology classification detection rate against keratosis leukoplakia, adductor spasmodic dysphonia, and vocal nodules.

Markaki et al (2011) explored the pathological voice detection and pathology classification. For voice disorders classification the modulation frequency and modulation spectrum for feature selection are used. The Higher Order Singular Value Decomposition (HOSVD) for dimensionality reduction is proposed. The proposed approach is used to classify healthy and pathological voices, using Support Vector Machines (SVMs). Finally, the voice pathology detection classification is achieved with high accuracy in the research, the HOSVD is comparably less. However, for voice pathology classification the suggested approach significantly outperformed the performance of cepstral-based features.

Firdos et al (2016) discussed the classification of normal and two disordered voices. The Support Vector Machine (SVM) for classification of normal and two disordered voices are used. This classification is characterized by Hoarseness, Vocal Fatigue, Periodic Loss of Voice, Inappropriate Pitch, and Loudness. The Mel Frequency Cepstral Coefficients (MFCC) is used for the feature extraction from the pathological voice signal and used the Genetic Algorithm (GA) for efficient feature selection. In terms of classification, the proposed performance result is compared with the proposed research result.

Ezzine et al (2018) investigated voice pathology detection. The pattern recognition process is used to find relevant glottal flow features for detecting voice disorders. The pattern recognition process is applied to the “MEEI” and “SVD” databases. The proposed method detects the normal and pathological voice signal which was pronounced by male and female speakers. This research used the Artificial Neural Network (ANN) and Support Vector Machines (SVM) for classification of normal and pathological voice signals. The system accuracies of SVM classifier compared with the research classifier accuracy, but it remains less with the proposed system Modified Optimized Back Propagation Network Disorder Voice Classification.

Delgado-Trejos, .E et al (2008) proposed the multivariate statistical analysis and features heuristic search methods called hybrid methodologies for reduce dimensionality and identify pathologies effectively. The statistical and geometrical relevancy for accurate pathologies detection is used. Since it doesn't have distinctness in the characterization, the research used the MANOVA progressive algorithm to analyze the subsets of data. Also, the data preprocessing, to identify pathologies is performed. For this reason, the classification performance and feature selection were increased without any complexity. Also, Principal Component Analysis (PCA) was employed for data variance reduction. Finally, the Multivariate Analysis of Variance (MANOVA) technique is used for classifying data input with respect to the normal voice and pathology voices.

Azadi et al (2015) proposed the Feature Selection, Classification scheme named as Gaussian mixture model (GMM). This model helps to detect voice disorders for early diagnosis. The feature selection with respect to the healthy and Patient with Parkinson (PWP) is achieved. This was worked based on the Gaussian Mixture Model (GMM). The accurate classification was achieved for the classification of male voice disorders and female voice disorders. The kernel support vector machine classifier is applied for gender voice disorders classification and features extraction.

2.3 Classification of Voiced Data

Zhang et al (2017) proposed the automatic and rapid computer-aided diagnosis system for pathological brain images detection. Magnetic Resonance Imaging (MRI) is used for simplification and classification. The classification result might be pathological or normal. For feature extraction, the Hue Moment Invariants (HMI), and for classification the used of Twin Support Vector Machine (TSVM) and Generalized Eigenvalue Proximal SVM (GEPSVM) are used. Finally, the methods “HMI + GEPSVM” and “HMI + TSVM” achieved higher classification accuracy on pathological brain detection and are proposed. These contributions are used to concentrate the better classification performance.

Zhang et al (2017) proposed the MDNet to direct multimodal mapping. This helps to map medical images and diagnostic reports. The diagnostic reports are done by visualizing the attention. The medical images mapping is done by using the symptom descriptions. That technique analysis is used to read images, generate diagnostic reports and retrieve images. This research is used to enhance the feature extraction and efficiency in an optimized manner. This research helps to make the multimodal mapping concept to manage voice in classification in different criteria.

Wang et al (2017) proposed the Optimization techniques for Pathological Brain Detection. The viaWavelet Packet Tsallis Entropy for Detection and Real-Coded Biogeography for Optimization are used. This research is detecting pathological brains with the help of Novel Pathological Brain Detection (PBD) with respect to the Feed Forward Neural Network (FNN). This research shows the proposed WPTE + FNN + RCBBO produced accuracy better than the existing approach. This research proposes the Feed Forward Neural Network (FNN) approach for final optimization in the voice pathology detection.

Wang et al (2017) evaluated the backpropagation neural network algorithm on the therapeutic mechanism of Amomum compactum gentamicin. Acute Kidney Injury (AKI) on the Back Propagation Neural Network (BPNN) is taken. These models were established for classifying data from the control, model, and AC-treated groups. Finally, the BP neural network algorithm was used for the classification of data, and the accuracy rates for classification were good in both positive and negative spectra. This investigation has helped us to obtain a better understanding of protective mechanisms of classification as well as the Backpropagation Neural Network algorithm.

Wang et al (2017) proposed the computer-aided diagnosis system for hearing loss. The artificial intelligence is applied for detecting unilateral hearing loss and used the magnetic resonance image scanning for efficient preprocessing. For features extraction, the Discrete Wavelet Packet Entropy (DWPE) on brain images is used. Finally, the Single-Hidden Layer Neural Network (SLNN) is used for classification. The three processes are used to detect the pathological left-sided and right-sided sensor neural hearing loss with higher accuracy on the diagnosis system.

Wang et al (2018) proposed the new pathological brain detection approach for healthy or pathological brain image classification. The novel Stationary Wavelet Entropy (SWE) for brain image extraction is proposed. The Wavelet Entropy (WE), Wavelet Energy (WN), and Discrete Wavelet Transform (DWT) comprised of proposed Stationary Wavelet Entropy (SWE). The classification performance was improved for the proposed system. This contribution feature of Stationary Wavelet Entropy (SWE) for pathological brain detection had been used in this research. Discrete Wavelet Transform for optimization in the feature extraction, a Feed-Forward Neural Network for classification, and Principal Component Analysis to track voices on neural network.

Shia et al (2017) proposed the voice disorder detection system using Discrete Wavelet Transform (DWT). The Feed Forward Neural Network (FFNN) is used to classify the normal and abnormal utterances taken from Saarbruecken Voice Database (SVD). The Discrete Wavelet Decomposition is used to distinguish the pathological voices from normal voices with higher accuracy. The maximum accuracy obtained is a reliable feature for detection of voice pathology. Normal and pathological voice classification is used in this research are efficient in the classification of normal and abnormal input voice signal with respect to the Discrete Wavelet Transform (DWT) and Feed Forward Neural Network (FFNN) on Saarbruecken Voice Database (SVD).

Pedraza et al (2017) examined the Glomerulus Classification with Convolutional Neural Networks. This proposed system used diseases diagnosis in Neuropathology. Deep learning framework is also used for automatic glomerulus classification. This form the Convolutional Neural Networks (CNN) classification to deduct and classify the Glomerulus and Non-Glomerulus image segments on the Glomerulus regions. This is suitable for Glomerulus classification with robust results. This helps the research for usage in a large quantity of voice classification with lesser time tasks on the voice classification.

Paiva et al (2018) proposed the Multiclass Support Vector Machine Recursive Feature Elimination (SVM RFE) method, Support Vector Machine (SVM) and Artificial Neural Network (ANN). The SVM RFE is used for relevant features extraction. The SVM + ANN is used for classification. This supervised learning uses the automatic learning method to base on pathologic Arterial Pulse Wave (APW). The APW is used in noisy-relevant segments of the signal reduction. The APW dataset is used for analyzing the healthy and non-healthy patients using the cross-correlation features, morphologic, time domain statistics, and wavelet features. The comparison results of SVM and NN shows, the higher performance of SVM are used to make the research with the efficient classification of healthy, pathologic and noise diagnosis measure. The SVM performs significantly better to distinguish between “noisy” from “pathological” and “non-pathological” with higher Accuracy on the original classifier test with simplicity and low complexity to remove noise signals.

Nayak et al (2017) proposed the efficient CAD system and analyzed the brain MR images for pathological brains and healthy brains classification. The Simplified Pulse-Coupled Neural Network (SPCNN) for preprocessing is used. Meanwhile, the Region of Interest (ROI) segmentation and Fast Discrete Curvelet Transform (FDCT) are applied for feature extraction. PCA+LDA approach is used for dimensionality reduction. Probabilistic Neural Network (PNN) is applied for classification. Similar to this research in the proposed system, for preprocessing phase the input speech signal, silence removal and noise removal are performed by combining Wiener and Discrete wavelet transformation filters to produce preprocessed speech. The proposed methodology for feature selection and Extraction process is performed in extracting only the relevant features. This helps us in dimensionality and time reduction and Cat Swarm Optimization technique for improved optimization. The proposed technique CSOMFCC has extracted the features in reduced dimensionality and in less time. The proposed classification process used to extract features and classify voices depending on the pathological and normal classification.

Kolachalama et al (2018) deal with the deep neural networks of Pathological Fibrosis with renal Survival for kidney disease severity deduction and classification. The Convolutional Neural Network (CNN) models were used for image input and output classifications. The Pathologist-Estimated Fibrosis Score (PEFS) is outperformed better in classification study.

Jia et al (2017) proposed the pathological brain detecting system for classifying the pathological or healthy or other different voices automatically and accurately with the help of Magnetic resonance imaging (MRI) development with respect to the cerebrovascular disease, degenerative disease, inflammatory disease, and neoplastic disease. The data augmentation for unbalanced distribution of the dataset is performed, deep stacked sparse autoencoder to train the network and softmax layer for classification.

Fang et al (2018) proposed the voice disorders classification by Mel frequency cepstral coefficients. The detection of voice disorders used for a screening method for voice diseases. The Deep Neural Network (DNN), Support Vector Machine, and Gaussian Mixture Model are used for cross-validation in classification. The detecting voice pathologies in male and female based on three Mel Frequency Cepstral Coefficient acoustic features is carried out efficiently to differentiate between normal and pathological voice samples.

Salehinejad et al (2018) proposed the simulation of pathology for imbalanced with over-representation of voice problems. The Generative Adversarial Network (GAN) for pathology classification is used. The Deep Convolutional Neural Network (DCNN) is used to deduct the pathology on the dataset using GAN generated images.

Pan et al (2017) proposed the pathology image analysis. The automated cell segmentation for access to complexity, variability, and individual texture is used. An automated nuclei segmentation method for Pathological image calculation with respect to the Deep Convolutional Networks (DCN) with efficient segmentation performance and efficient error reduction is carried out.

2.4 Results of the Diagnostic Method

Kukharchik et al (2007) proposed the vocal pathology diagnostic method. The acoustic analysis consists of wavelets and pseudo-wavelets for feature extraction.

Moreover, continuous wavelet and wavelet-like transform are used for vocal pathology diagnostic methods. The direct vocal fold observation classifies the speech signal with the help of SVMs and Neural Networks. In the classification method, Support vector machine is used in fold pathology detection. Silence Removal and Vowel Extraction were also used to improve classification results with higher accuracy.

Saeedi et al (2011) examined the pathological voice analysis process. Non-invasive tool for the detection is used as the main goal of pathological voice analysis. The normal and disordered voices are classified based on the wavelet-based method. The Support Vector Machines are combined with Wavelet filter banks. This combination is used in feature extractors and classifiers. The Support Vector Machines are used for classification and Wavelet filter for feature extraction. The Orthogonal filter (lattice) with a genetic algorithm with a fitness function is proposed to fine-tune the classification process in an efficient manner. This proposed method rectifies the normal and pathological voice signals with a higher classification rate on databases.

Saidi et al (2015) proposed the non-invasive method. This process is used to classify normal voice signals and disorder voice signals. The wavelet decomposition is used for feature extraction with better frequency. In addition to this, the M-band wavelet decomposition is also applied for feature extraction. After applying this algorithm, a genetic algorithm is employed for optimal wavelet filtering. Support Vector Machine is used in the final classification of normal and pathological voices in the database. To manage the classification technique, the noninvasive method namely M-band wavelet system, five-band wavelets, and genetic algorithm are implied for improved classification performance.

Peng et al (2007) discussed the classification of pathological voice. The 30 acoustic features are considered to classify healthy and abnormal voices. The PCA comprised of the feature space transformation and data dimension reduction. The Support Vector Machine (SVM) algorithm is used for classification of healthy and pathological voices. Finally, the detection and classification rates are measured with respect to Sensitivity and Specificity and Feasibility of voice signal classification. The feature selection is done by using LSSVM classifier with the RBF kernel function.

Hosseini et al (2008) proposed a Local Discriminant Bases (LDB) with respect to the Wavelet Packet tree. This helps in pathological voice diagnosis on voice signal classification. In the speech technology, the researcher introduces a Genetic Algorithm for feature selection and Support Vector Machines as a classifier for finding the pathologies such as keratosis leukoplakia, adductor spasmodic dysphonia and so on. Combined all those systems and it is called the Discriminant Based System. This system helps in separating polyp and nodule and produces a higher performance against wavelet packet features. The result shows that the feature selection using GA is an optimization procedure to find the best feature classification against the pathology.

Fonseca et al (2007) discussed Wavelet time-frequency analysis. This voice digital analysis focuses on laryngeal pathology identification. And also discussed the discrete wavelet transform (DWT-db), linear prediction coefficients (LPC) and Least Squares Support Vector Machine (LS-SVM). And identify nodules in vocal folds by using larynx pathology classifier with higher classification accuracy and low complexity. Finally, to classify the normal and pathologically affected voices by DWT-db20. The proposed LS-SVM classifier linear, RBF and MLP also used to larynx pathology classification.

Fonseca et al (2005) discussed pathological voice signals identification methods. The Discrete Wavelet Transform and Support Vector Machine are used in the research. This system is used to classify nodules in vocal folds, with different ages for both male and female. Based on Double Discrete Wavelet Transform (DWT–db) and Support Vector Machines (SVM) classifier is employed to produce the output. The Linear prediction coefficients (LPC) filter returns the non-linear Least Square Support Vector Machines (LS-SVM) for pathology classification with higher classification accuracy. The pathological voices and normal voice signals (low mean square values) are differentiated. Base Function (BF) is used as classification generalization to distinguish between normal and pathological voices.

Arjmandi et al (2012) studied the identification of voice disorders in the vocal folds. The preprocessing is done based on the amalgamation of Short-time Fourier Transform (STFT) and Continuous Wavelet Transform (CWT). This will help to classify the disordered voices and normal voices. Wavelet Packet Transforms (WPT) is also used to optimize the preprocessing process. Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) combine together and the dimension reduction is done. The proposed concepts return the best Accuracy, Sensitivity, and Specificity, to optimize the results based on the LDA and a Support Vector Machine-Based Classification method. This proposed system improves the accuracy and high-speed diagnosis procedure. So, these processes rectify the abnormal noise from the vocal folds of the voice signals in pathology.

Arjmandi et al (2011) discussed voice disorder identification. This Acoustic analysis deals with the diseases diagnosing at early stages. To compare and identify voice disorders with laryngoscopy. The statistical pattern recognition techniques are applied to the proposed feature reduction and pattern recognition methods. This helps to classify voice disorders with feature vectors. The feature vectors originate by Principal Component Analysis (PCA) or linear discriminant analysis (LDA). The LDA contains the individual LDA, Forward LDA, backward LDA, and branch-and-bound LDA methods are used as the feature selection. These methods are used to identify voice disorders with higher Accuracy, Efficient Sensitivity, Specificity, and Area under the Receiver Operating Characteristic Curve (AUC). Finally, the LDA with SVM returns the best performance with the lowest complexity.

Dundar et al (2010) explored the optimal classification of pathology. The research deals in a computer-assisted classification of digitized optimal classification of pathology. This research produces the optimized classification accuracy by using traditionally supervised training techniques with ROI level, and the multiple-instance learning approach under the ROC to obtain better classification.

Wu et al (2019) investigated the Voice analysis process in a Non-invasive way. The research proposed a novel model Joint Learning based on Label Relaxed low-Rank Ridge

Regression. This method is used for voice-based disease detection. The preprocessing is done over the 15 models by using regression losses and transformation matrix. the e-dragging technique is used to classify the classes. This tow process helps to make the correlation structure on the classified signal. The proposed algorithm consists of multiple input signal and weight. The e-dragging technique achieved high performance on the detection.

Amin et al (2019) proposed cognitive healthcare framework with an Internet of Things cloud technologies. The proposed system helps in intelligent decision-making with accurate, timely, and high-quality healthcare services. The EEG pathology classification technique uses deep learning for classifying the pathologic or normal signals. The patient is observed and monitoring the EEG processes the results to the healthcare providers for the patient's condition and provides emergency help in a critical state.

Walton et al (2019) discussed the voice quality and performance analysis. The glottal closure for phonation is used to improve vocal function by Unilateral Vocal Fold Paralysis (UVFP). The voice features are extracted with respect to multidimensionality, timing, selection rationale, validity, reliability, and responsiveness. The proposed system performs the selection rationale, validity, reliability, and responsiveness selection in a successful manner.

Gómez-García et al (2019) examined the automatic voice condition analysis of voice pathologies. And discussed the robustness of voice pathologies with respect to the input voice. To improve the speech task analysis and extralinguistic classification using the acoustic features in classifiers by the classification methodologies under cross-dataset scenarios. The feature selection is used to identify a reduced subset of relevant features in the hierarchical scenario at Saarbruecken voice dataset. The ROC curve of the system plotted for techniques to be transferred to the clinical setting of voice pathologies.

Garnavi et al (2016) discussed the pathological condition analysis from the computer-implemented method. The proposed system classified the severity image in an efficient manner. The discriminative pathology histogram and generative pathology histogram are merged together and formed as the hybrid image representation model. This classifier is used to classify the severity of the pathological condition by using baseline convolutional neural network.

Kuresan et al (2019) discussed the neurological disorder and its progression. This research helps to diagnosis and monitoring system in the pathology environments. To identify the signal as the Parkinson or not with the help of Wavelet Packets, MFCC and a fusion of MFCC and WPT. This feature extraction might be applied to the classifiers HMM and SVM. The study gives the MFCC, WPT with HMM performance in extraction and classification.

Cui et al (2019) discussed the neurodegenerative disorder. With the help of Magnetic Resonance Images (MRI) the diagnosis and monitoring the disease progression are maintained by the longitudinal analysis. The feature extraction and classification were compared with the characteristics abnormalities and the time points are also calculated to speed up the extraction and classification. The convolutional and recurrent neural networks are used for feature selection and then the classification is done by the Recurrent Neural Networks (RNN). In addition to this, the feature extraction and classifier are made by the proposed method for achieving optimal performance improvement.

Rose et al (2014) discussed the speech-language pathologists at Australia. It concentrates on the aphasia rehabilitation practices on the aphasia management are identified in Low levels of knowledge and confidence, culturally and linguistically diverse, clients and discourse approach. The Group and intensive services implement communicative access for pathologists in an efficient manner.

Boyanov et al (1997) analyzed the software system for pathological voice analysis and screening of laryngeal diseases. And also use the analysis in a non-invasive way for laryngoscopes. To propose and build standard sound Blaster with a microphone. With, the graphics-driven and user-friendly training additionally.

Dibazar et al (2006) discussed the assessment of pathological voices. Depending on the perceptual judgments, the pathological voices are labeled and also assigned to the given voice recognition system. This proposed research focuses on the five specific pathologies. These were classified based on the single label classification and originally labels assignment over the voice samples. The results show that the pathological voice assessment performance by labels assignment and improved with respect to the single label.

Saenz-Lechon et al (2008) discussed the automatic detection of voice pathology. The Massachusetts Eye & Ear Infirmary (MEEI) Voice Disorders Database, cross-validation is used with the help of final confusion matrix, confidence intervals for all measures are considered. Detector performance curves such as Detector Error Trade-off (DET) and Receiver Operating Characteristic (ROC) plots are also considered for short-term parameters and multi-layer perceptron.

Meanwhile, the Voice Pathology Detection and Classification are performed by the Al-Nasheri. A. et al (2018) with the help of Auto-correlation. The early detection of voice pathologies and the diagnosis are concentrated. The robust feature extraction by frequency bands using autocorrelation and entropy is used. Finally, the detection and classification voice pathologies depending on the bands, method, and database are achieved.

To overcome detection accuracy Alarifi. A., et al (2018) proposed the multiband approach based on a three-level Discrete Wavelet Transformation (DWT). Here the Fractal Dimension (FD) was calculated for power spectrum analysis. From this spectrum range, the normal and pathological subject was classified with a lower detection rate.

Baishya et al (2018) proposed the Speech De-noising using Wavelet Thresholding. The classification of the signals is done in three regions such as Voiced, Unvoiced and Silence Regions. The proposed method was to extract the noise from the input SPEAR database effectively.

S´aenz-Lech´on et al (2008) proposed the Automatic Detection of Voice Pathologies process on Audio Compression. The compressed audio input file is taken and this file is characterized by the voice signals. Finally, the classification is performed using Gaussian mixtures models and support vector machines.

Zhang et al (2018) suggest the Cat Swarm Optimization for identification of alcohol use disorder by computer-vision based technique. The proposed system is a combination of three criteria such as Wavelet Entropy, Feed Forward Neural Network, and Cat Swarm Optimization (CSO) called the proposed BWE+TFNN+CSO approach. The usage of HMI-SVM, FRFT, LMCoP, and PNNFA methods proposed alcoholism identification methods. The proposed BWE+TFNN+CSO approach returns the best result in terms of wavelet transform techniques.

Sahidullah et al (2013) proposed the novel family of windowing technique to compute Mel Frequency Cepstral Coefficient (MFCC) for automatic speaker recognition from speech based on a fundamental property of Discrete-Time Fourier Transform (DTFT). The Hamming window is modified for the creation of power spectrum and phase information. It is used in speaker recognition by window functions.

Deshpande et al (2018) proposed the Electro glotto graphic Parameter Extraction for Voice Pathology Assessment in voice pathology analysis. The new method is used for detection of glottal instants and EGG parameters from voiced and non-voice segments. The proposed Adaptive Variational Mode Decomposition (AVMD) is used for suppressing low and high-frequency noises. Finally, the non-glottal instants are retrieved from the noise-free and noisy EGG signals.

2.5 Chapter Summary

The strategies available in voice pathology detection process for feature extraction, feature selection, and voice pathology classification are explored. However, the normal and pathology classification is a challenging task against the two different data set namely Saarbruecken data set and own Real-Time dataset, particularly in Saarbruecken data set, the voice pathology detection cannot be accurately classified with the normal voice signal. Related works are discussed about the novel methods which classify all data with respect to the distribution over a pathology classification. Moreover, this research includes Parametric Representation by preprocessing. It retrieves the necessary signal from the overall input signal using the preprocessing techniques Noise Removal, and Silence removal. This review concentrates on the preprocessing parametric representation of the speech signal in order to increase the efficiency of the subsequent steps like feature extraction and classification.

In this research, the Wiener and wavelet transform filters are combined together as Hybrid Wiener Filter Discrete Wavelet Transforms (HWFDWT) and are proposed to reduce the noisy speech signals. Speech signals are divided into blocks and windowing is done by taking a small subset of a larger dataset and the samples are taken. Those signals are applied to the one-dimensional Wavelet Transforms (WT) and Wiener filters, the minimum speech values are considered to perform noise filtering, in the proposed method WT along with Double density wavelet is applied to get the resulting output close to the original signals. Then finally Wavelet Transforms (WT) and Inverse windowing are performed subsequently to get enhanced speech as output. The process of retaining useful information from the signal and discarding the unwanted information is feature extraction. In this section, the relevant feature selection and Extraction process are discussed in order to collect only the relevant features. This helps in dimensionality and time reduction. In the feature Selection phase, From the denoised signals, the number of specific wavelets and threshold type is applied in order to select the most dominant features; the increasing amount of research is discussed which focus on in selecting the right threshold method and calculation.

The classification of Voiced input Data are discussed with respect to the various researchers' proposals. The BPNN algorithm is modified and given as a proposed work. In the proposed methodology Modified Optimized Back Propagation Network Disorder Voice Classification (MOBPNDVC) learning algorithm, functional constraints are involved (which defines the range of alternative weights). This is achieved by the cost function which is designed for output layer and hidden layer separately. The classification of normal and pathological voices feature vectors is used to adjust the parameters of a Modified Optimized Back Propagation Network Disorder Voice Classification (MOBPNDVC) for each input and to train a Support Vector Machine (SVM) efficiently by optimizing the Classification vector by fine-tuning the weights, as the result by adjusting the weights the speed will increase significantly. The voice optimization of classified voices is performed in reducing complexity. Finally, the Results and Discussions were concluded about the pathology diagnostic method. The conventional method available for, these steps along with the enhancement operations used to improve these methods, are outlined in the following chapter, Chapter 3, **Research Methodology and Approach**.