Voice-Disorder Identification of Laryngeal Cancer Patients

Mr.G.B.Gour[1]

Department of Electronics and Communication Engineering
BLDEAs V.P.Dr.P.G.Halakatti College of Engineering and
Technology, Vijayapur, Karnataka, India

Dr.V.Udayashankara[2]

Department of Electronics and Instrumentation Engineering
Sri Jayachamarajendra College of Engineering, Mysuru,
Karnataka, India

Dr.Dinesh K. Badakh[3]

Department of Radiation Oncology
Sri Sri Siddhivinayak Ganapati Cancer Hospital, Miraj,
Maharastra, India

Dr.Yogesh A Kulkarni[4]

Department of Medicine
Nargis Dutt Memorial Cancer Hospital, Barshi
Makarastra, India

Abstract—**This Previous studies have shown that much of laryngeal cancer-based work was carried out with a minimal set of linear features. Much of the work was focused on the study of larynx preservation, quality of life around radiotherapy, or surgery. The voice disorder database was not solely limited to laryngeal cancer. In the context of this, the paper proposes a non-invasive voice disorder detection of laryngeal cancer patients. The sustained vowel /a/ was recorded with 55 laryngeal cases and 55 healthy cases. Owing to the non-linearity property of the vocal cords, seven non-linear parameters along with biologically inspired 39 Mel-Frequency Cepstral Coefficients (MFCC) are extracted. This forms a laryngeal dataset of size 110X46. The wrapper method is used for better feature selection and to enhance the discriminating ability of the present work. The classification is carried out using a tuned support vector machine (SVM) with grid search and random forest (RF). The present work has shown an improved accuracy of 76.56% with SVM and 80% in the case of random forest. The forward selection of features along with the involvement of non-linear features has played a significant role in the better performance of the present system.**

Keywords—*Support Vector Machine (SVM); random forest; Mel Frequency Cepstral Coefficients (MFCC); voice disorder detection; laryngeal cancer; non-linear features*

I. INTRODUCTION

One of the main categories of head and neck cancer is laryngeal cancer (LC). The oncologists can treat the LC by subjective evaluations and invasive diagnostic methods. LC patients are found to be reluctant in many situations. The stroboscope does not record the status of vocal function with cycle-to-cycle information. Moreover, the parameters and rating information provided by stroboscope are visual perceptual ratings and subject to reliability and validity errors. The video laryngoscope (VLS) has limitations of irritating patients and a need for topical anesthesia [1]. The voice analysis provides clinicians an alternative, non-invasive, and objective analysis tool in this regard. So far the LC based works were carried out using statistical and machine learning approaches with linear features only. The experimental study was conducted by [2,3] with 80-1925 LC cases. The statistical analysis was carried out with only few linear features. Jitter, shimmer, noise-to-harmonic ratio (NHR) and maximum phonation time (MPT) have shown a significant difference between cancer and dysfunctional groups in the results (p<0.05). The perturbation measures, NHR and MPT were verified with laryngoscopic evaluations for which LC cases are reluctant. To carry out the LC classification, a prototype distribution type map (PDM) was proposed by [4]. But, this PDM modeling was done using neural maps which was quite complex and time consuming with increased number of iterations. Moreover, efforts have been made to preserve the larynx after surgery, radiotherapy, and rehabilitation [10]. It is clear from the previous studies that,

- Much of the LC database used was found to contain LC cases along with dysphonia cases.

- Only linear based voice acoustic features were used.

- Much of the work was focused on voice analysis of LC cases to study the quality of life and preservation of larynx w.r.t surgery or radiotherapy.

- Much of the study was carried out on LC were lacking voice disorder analysis using machine learning approaches.

The main objectives of the present study are:

- To create and use a specific LC database concerned with the pathologies related to adjacent regions of the larynx.

- To assess the performance of linear and non-linear descriptors to discriminate between the patients with LC pathologies and healthy cases.

- To develop a cost-effective (using open-source platform Ubuntu 15.0 and GNU Octave 4.0) and non-invasive voice acoustic classification tool.

Hence, the paper presents a non-invasive voice disorder detection of LC cases with optimized feature selection using forward selection method and, classification by tuned SVM

and random forest. Rest of this paper is organized as follows. Section II describes the background study with reviews in the existing research, Section III illustrates materials and proposed methodology for LC detection, Section IV depicts extensive experimental evaluation for the proposed method and finally Section V presents concluding remarks with future scope.

II. BACKGROUND STUDY

Researchers have adopted many clinical and acoustic methods in classifying the LC cases but with little attention to detection of LC. The clinical methods have been suffering from patients comfortless to laryngoscopy and variations in measurements. A more diversified LC database was adopted in the following studies. Data dependent random forest was proposed by [5] in fusing knowledge with increased classification accuracy. Here, 110 LC cases with different pathologies (tumors, polyps, cysts, papillomata, keratosis, carcinoma and paralysis) were included in the study. The auto-associative neural network was used to differentiate LC cases (139-laryngitis cases, 211-hyper-functional dysphonia and 212-recurrent laryngeal nerve paralysis) from healthy group. Frequency-based features like MFCC, Cepstral-based features, HNR, and LP-based parameters were used to form a 14-dimensional vector for each subgroup. Over 37 linear features were used to train neural network with 87.5% accuracy [6]. The authors believes that with more protuberant features describing dynamics of vocals cords along with better feature selection methods can enhance the accuracy of the LC detection system.

However, much of the experimental work was carried out to understand the impact of radiotherapy on LC cases at particular periodic follow up. The retrospective study was conducted on 115 early-stage (in-situ, T1-T2) LC cases to assess the improvements in visual, acoustic, and patient-reported findings [7]. A similar study was conducted on laryngopharyngeal cancer cases [8]. To determine the effect of supra-clavicular RT on the physiology and functioning of the vocal fold, an experimental study was performed on 29 female patients diagnosed with breast cancer who underwent supra-clavicular RT, reported at intervals of 1 and 6 months before and after treatment. [9]. In all these radiation-based studies, a limited number of linear voice acoustic parameters were used with an average significant variation. These variations were used in the assessment of the early stage of LC or analysis of voice quality. The work was carried out to know the impact of voice rehabilitation on laryngeal cancer patients after radiotherapy using jitter, shimmer, quality of life scores (QOL), and voice handicap-index (VHI). These studies have provided a better road map towards multi-classification among LC cases. But, no specific voice rehabilitation was found to be necessary for laryngeal cancer patients after radiotherapy as there was no statistical significance found with these parameters [10].

III. MATERIALS AND METHODS

A. Database

The present study involves cases with laryngeal pathology. The voice samples of laryngeal cancer patients came for radiotherapy, were collected after taking the written consent from each case. The cases have given sufficient information about the non-invasive procedure being followed while recording the voice samples. The ethical approval was taken in advance from the Sri Siddhivinayak Ganapati Cancer Hospital, Miraj, and Nargis Dutt Memorial Cancer Hospital, Barshi, (Maharastra), India. The sustained vowel /a/ was recorded using an Omni-directional microphone for 1-3 seconds. The recording has performed at a sampling frequency of 44.1 kHz with a 16-bit resolution using the Praat software. The operated laryngeal cases were not included in the recording. A total of 55 laryngeal cases in the age group of 62.8±10.8, having 49 males and 6 female cases are included as shown in the Table I. The voice samples for the control group of 55 cases are collected in the age group of 63.1±10.9. This forms the voice dis-order database for LC having a total of 110 cases.

B. Proposed Algorithm

The methodology adopted in the present work is as shown in Fig. 1. Pre-processing is optionally adopted in the present work based on the type of classifier used. As shown in Fig. 1, the speech frames of 25 ms with 50% overlap are used throughout work. A pre-emphasizing filter with α = 0.97 is used to boost the high-frequency components.

$$y_p(n) = y(n) - \alpha y(n-1) \tag{1}$$

Then, the speech enhancement is carried out using a two-step noise reduction method (TSNR) using a Wiener filter [11]. This stage is followed by the extraction of the MFCC and non-linear features as discussed in the section C. To validate the irrelevance to the present function, the features were investigated. By utilizing forward function selection, the identification of voice dysfunction is optimized and comparative analysis is established between optimized and non-optimized features with a strongly validated mix. In classification process, SVM with grid search and random forest methods are adopted owing to their ability in dealing with low dimensional features space.

TABLE I.   NON-OPERATED LC CASES INCLUDED IN THE STUDY

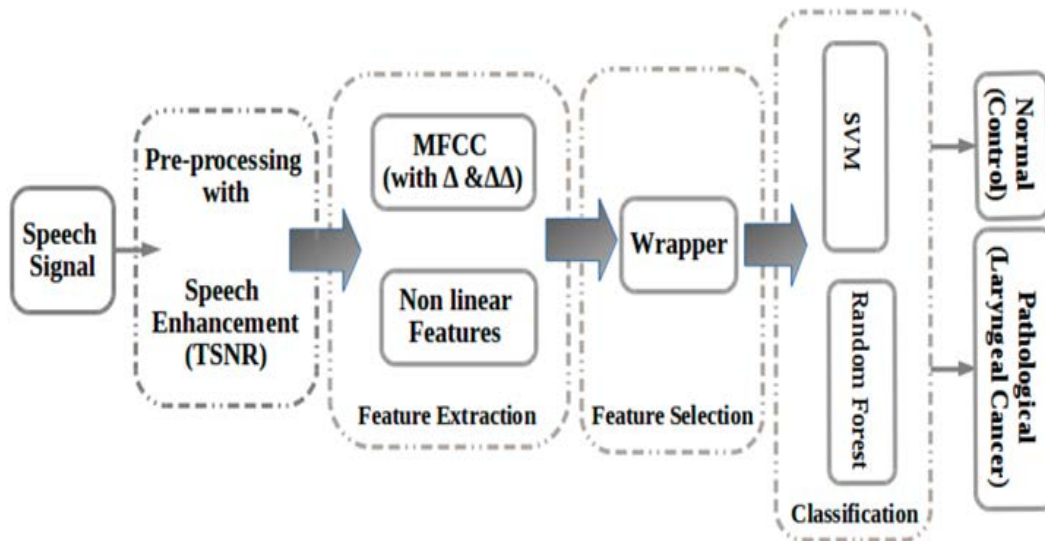| Voice Pathologies	Number of Cases
Ca-larynx	24
Ca-laryngo-pharynx	1
Ca-Supraglottis	7
Ca-epiglottis	12
Ca-Pharynx	4
Ca-Hypopharynx	6
Ca-Cricopharynx	1
Total	**55**

Fig. 1. Methodology of the Present Work.

C. Feature Extraction

1) MFCC: In the feature extraction, linear and biologically inspired 39-dimensional MFCC features are extracted per frame as shown in Fig. 2. Here, the 13 MFCC is derived from each of the speech frames by taking the discrete Fourier transform (DFT). To obtain its power spectrum, it is then passed through a triangular filter bank with 24 filters, uniformly positioned on the Mel scale. Log energy is computed for each of the banks, which is found to be sensitive to small variations in the articulatory movements. Then the MFCC of 13 coefficients is obtained by using DCT. A total of the 39-dimensional vector of MFCC are obtained with its derivatives. The 39-dimensional vectors are fed to the supervised SVM for the classification of LC pathology or healthy [12].

2) Non-linear features: It is clear from the previous studies that, the linear time-invariant property is no longer holding good for the non-linear structure of vocal cords and their dynamic behavior. The non-linear features are capable of producing the trajectories in phase space from the vibrations produced by a dynamical system like the vocal folds. The 7-dimensional non-linear data subset was obtained by the extraction following features [13,14,15];

A. Mutual information (MI): It provides a flexible approach in evaluating dynamical variables by using the method of delays. It refers to the reconstruction of phase space of voice signal using with minimum delay, called the first minimum of mutual information. The MI is computed by using the equation (2).

$$S = \sum_{ij} p_{ij}(\tau) \ln\left(\frac{p_{ij}(\tau)}{p_i p_j}\right) \tag{2}$$

3) False nearest neighborhood (FNN): If there are two nearest neighbours $\vec{s_i}$ and $\vec{s_j}$ in dimension m with distance between them is $\|\vec{s_i} - \vec{s_j}\|$ then, $\vec{s_{i+1}}$ and $\vec{s_{j+1}}$ are the maps of the respective $\vec{s_i}$ and $\vec{s_j}$ in m+1 dimension. The divergence rate of these points while travelling from dimension m to m+1 is shown by (3),

$$R_i = \frac{|s_{i+1} - s_{j+1}|}{\|\vec{s_i} - \vec{s_j}\|} \tag{3}$$

These two points become false neighbors if the distance between them exceeds a certain threshold. Then the condition for maximum embedding dimension is the fraction of points for $R_i > R_t$ threshold gives the estimation for embedding dimension m. Here, p_i is the probability of finding a time series value in the i-the interval, and $p_i(\tau)$ is the joint probability of p_i and p_j corresponding two points. MI measures the mutual dependence of the points p_i and p_j.

4) Correlation Sum and Dimension (CD): The correlation dimension quantitatively describes the complexity or irregularity of the trajectory in phase space. This irregularity is confined to the correlation of two points on the trajectory known as the correlation dimension. The convergence of finite correlation can be obtained by straight-line fitting of the log-log plot of the correlation sum which is given by equation (4).

$$C_i^m(r) = \sum_{i=1}^{N}\left(\frac{2}{N(N-1)}\sum_{j=i+1}^{N}\theta\left(r - \|\vec{s_i} - \vec{s_j}\|\right)\right) \tag{4}$$
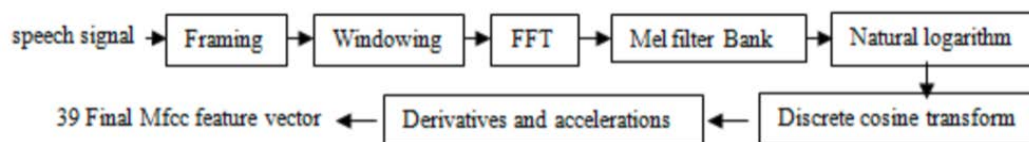

Fig. 2. Estimation of MFCC.

The correlation dimension quantitatively describes the complexity or irregularity of the trajectory in phase space. This irregularity is confined to the correlation of two points on the trajectory known as the correlation dimension. The convergence of finite correlation can be obtained by straight-line fitting of the log-log plot of the correlation sum which is given by equation (5),

$$D_{\text{TT}}(r) = \frac{C(r)}{\int_0^r \left(\frac{C(r')}{r'}\right) dr'} \tag{5}$$

Information dimension or correlation dimension of order 2 (CE_2): Information dimension CE_2 can be estimated from the modified correlation sum.

5) Largest Laypunov Exponent (LLE): This feature depicts the average divergence rate of the neighbor trajectories as calculated by Rosenstein's method. The separation or the average divergence of points in a trajectory is d (t) = C $e^{\lambda_1 t}$. Here λ_1 is the LLE and C is a constant.

6) Renyi Entropy (RE_1, RE_2): In a dynamic system like vocal cords, Renyi entropies describe the loss of information in time. There will be an evolution of a nearby point in phase space to far points. In random systems, Renyi entropies tend to infinity.

At the end of this section, 39-dimensional MFCC and 7-dimensional non-linear parameters provide a 46-dimensional feature vector.

D. Feature Selection

The feature selection methods help in revising the model with reduced complexity and optimized accuracy. The commonly categorized feature selection methods are filtering, wrappers, and embedding methods [16]. The wrapper method is based on a predictive model. The forward selection (FS) method is used as it is less prone to overfitting and having low computational cost, than backward selection. The FS method provides a reduced, un-correlated significant feature sub-set with average computational cost [17,18]. The steps involved in the forward selection are:

- Begins with null model M_0 with zero predictors.

- Pre assuming the significant levels S_{add} and S_{drop} to add and drop features respectively.

- For k = 0, ..p-1 (0 to 9 features), means considering all (p-k) models that augument with the predictors in M_k.

- Augumenting of significant model for S_{add} < cross-validation error (CV) and droping the previously added predictor for S_{drop} > CV error.

- The last step is replicated before a final optimum set of characteristics is obtained among M_1 ,M_2 , . . .M_p.

E. Classification Algorithms

Two supervised classification algorithms, SVM and random forest are used in this work owing to the size of the present dataset and number of features. Random forest is sensitive to the small changes in training data (bagging) that enable us to include in the present work. Both algorithms are fast in training and testing data-sets. In order to infer the most productive classifier for the detection of voice disorders, the comparative analysis is compiled between their success rates.

1) Support Vector Machine with a grid search: SVM is a supervised binary classifier using kernel tricks to trace the best hyperplane with the maximum margin between two classes. In a higher-dimensional space, the non-linear data is mapped, where the kernel becomes linear. Commonly used kernels with SVM are linear, polynomial, radial, and sigmoid. The SVM is used to classify the LC cases due to low computational costs and less prone to over-fitting. The data comprising of n-dimension feature vectors are first labeled by using the Audacity tool, scaled and normalized before feeding to the SVM. In all the cases, SVM-kernel tricking is employed with the help of a grid search. SVM is tuned with hyper-plane parameters with different kernels, gamma, and C values [19, 20]. The grid search will build and evaluate the model for each combination of hyper-plane parameters. Then, best hyper-plane parameters with good cross-validation (CV) accuracy. In the present work 10-fold, cross-validation is used. The C (1 to 500), gamma (0.00003 to 0.002), and the kernels are traced to find the best fit of these hyper-parameters. The grid search is separately performed in using different data-sets during experimental investigations as discussed in section III.

2) Random forest: Random forest is an ensemble of a set of more number of individual and uncorrelated decision trees. Because of RF's sensitivity to changes in training data, it allows each tree to randomly sample from the main data-set with replacement known as bagging. Hence, the bagging process provides us with a sample size of N, which is less in size than the actual whole data-set. But with replacement, each of the attributes may be repeated in each chunk. In the model, this forces much more variance among the trees and eventually results in less correlation between trees. Therefore, trees that are not only learned on different sets of data (bagged) but also use different features will make decisions eventually [20,21]. The typical steps to be followed for the core working of RF are;

- For a given training data of $\{(a_1,b_1), (a_2,b_2)…,(a_n,b_n)\}$ with X_i predictors which corresponds to the root node.

- Each non-terminal node splits up into two descedent nodes based on the value of one of the predictor variables.

- Categorical predictor variable creates partitions on either side of the split point with different sub-set of categories.

- 4. This process proceeds recursively before the conditions for stopping are satisfied.

- 5. For all events in the terminal nodes, an estimated value occurred by averaging the computation.

- Response of the most frequent class for classification problems.

An in-depth tree is kept at 5 to address the over-fitting of data.

F. Evaluation Process

By using precision, sensitivity, specificity, precision, and the Area Under Curve (AUC), the efficacy of the proposed algorithm is evaluated. The sensitivity evaluates the likelihood of pathological samples to be detected by the algorithm. In turn. The specificity assesses the algorithm's ability to classify typical samples. Precision reflects the percentage of pathological samples from the pathological class that is well identified. Besides, accuracy measures the correct classification rate of the algorithm. AUC assesses the capacity to differentiate between the normal and abnormal samples. The AUC offers an alternative means of calculating the performance of the method suggested. These measures are based on the following notions:

$$\text{Accuracy} = ((TP+TN)/total) * 100 \qquad (6)$$

$$\text{Sensitivity(Recall)} = (TP./(TP+FN)) * 100 \qquad (7)$$

$$\text{Specificity} = TN/(TN+FP) * 100 \qquad (8)$$

$$\text{Precision} = (TP/(TP+FP)) * 100 \qquad (9)$$

$$\text{AUC} = 0.5x(\text{Specificity} + \text{Sensitivity}) \qquad (10)$$

IV. EXPERIMENTAL RESULTS WITH DISCUSSION

According to Fig. 1, the speech signal related to sustained vowel /a/ is framed using a 25 ms window and 50% overlapping. Then, for each 25 ms speech frame, a 49-dimensional feature vector is extracted as discussed in section. Now with 110 cases, gives 110X46 sized main dataset. In this research, the dataset was divided a 70% of the training data and 30% for validation. All simulations were conducted in GNU-Octave using python.

The forward selection has been applied to a 46-dimensional data-set (MFCC & NL) and also on 39-dimensional MFCC separately. FS will select a sub-set of maximally performing 10 features, which improves accuracy with minimal computational time. In the present work, FS iteration is supported with 10 features due to more computational time, as shown in Table II. The best score of 0.80 is achieved. MFCC along with two non-linear features, CD and LLE are found to be significant. Where in the above Table II, Δn, and ΔΔn denote n[th] coefficient of first and second derivatives of MFCC respectively. Similarly,

10-MFCC significant parameters are selected using the wrapper method.

After performing forward feature selection, the experiments are conducted in two ways, with MFCC and Non-linear parameters and with only linear parameters MFCC to assess the impact of having non-linear features in the prediction of laryngeal cancer.

A. Evaluation based on SVM Performance

The 110 by 46 sized data-set is applied to SVM (with grid search). Table III shows the SVM performance with tuned hyper-parameter (C, gamma values) before and after applying the wrapper. It is observed that, hyper-parameters selected with gamma = 0.0001 with radial basis (RB) kernel for the whole dataset (C=30) and linear kernel (C=100) in case of MFCC data.

Table III shows performance of SVM in terms of accuracy (%), sensitivity (%), specificity (%), area under curve (AUC) and precision (%). The forward selection has shown a significant impact of involving non-linear features in the experimental study. The SVM with (RB) kernel has shown an improved accuracy from 72.85% to 75.66% for the whole dataset and accuracy enhancement from 62.82% to 66.66% in the case of the MFCC dataset. Thus wrapper method along with (RB) kernel of SVM has played a significant role in optimizing the dataset. This resulted in good discrimination of laryngeal cancer. Fig. 3 shows an increase in accuracy with feature selection (optimized) related to MFCC dataset from 62.82% to 66.66%. The maximum accuracy of 75.66% is achieved with a dataset having 2.81% optimization. Moreover, an average sensitivity of 74% has shown better discrimination ability of SVM with a complete dataset along with a better AUC rate of 76.56% as shown in Fig. 4.

B. Process Evaluation based on Random Forest Performance

Table IV shows the performance of random forest presented in terms of accuracy (%), sensitivity (%), specificity (%), the area under the curve (AUC), and precision (%). It is clear from Fig. 5 and 6, that with the whole dataset, 3.44% of optimization is achieved. This is responsible for the maximum enhancement in the accuracy from 76.56% to 80% along with a maximum AUC rate of 79.80%. Hence, from the experimental observations, it is evident that random forest is showing better accuracy (80%) and discriminating ability (79.80%) with a complete dataset.

TABLE II. AVERAGE SCORES OF FEATURES USING FORWARD SELECTION APPLIES ON THE COMPLETE DATASET

Average Score	Feature Names
0.72	ΔΔ8
0.70	MFCC10,ΔΔ8
0.76	MFCC10, MFCC11, ΔΔ8
0.74	MFCC2, MFCC10, MFCC11, ΔΔ8
0.74	MFCC2, MFCC4, MFCC10, MFCC11, ΔΔ8
0.76	MFCC2, MFCC4, MFCC10, MFCC11, ΔΔ8, ΔΔ-log-energy
0.77	MFCC4, MFCC10, MFCC11, ΔΔ8,ΔΔ12, ΔΔ-log-energy,CD
0.78	MFCC10, MFCC11, Δ8, ΔΔ8,ΔΔ12, ΔΔ-log-energy,CD,LLE
0.77	MFCC8,MFCC10, MFCC11, Δ8, ΔΔ8,ΔΔ12, ΔΔ-log-energy,CD,LLE
0.80	MFCC8,MFCC10, MFCC11, Δ6, Δ8, ΔΔ8,ΔΔ12, ΔΔ-log-energy,CD,LLE

TABLE III. SVM PERFORMANCE BEFORE AND AFTER APPLYING FORWARD SELECTION METHOD

Wrapper	Features	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)	Precision (%)	C	Kernel
Without Feature selection	MFCC + Non-linear parameters	72.85	73.33	72.00	72.66	82.5	30	Radial Basis
With Feature selection	MFCC parameters	62.82	61.82	65.21	63.52	80.95	100	Linear
Without Feature selection	MFCC + Non-linear parameters	75.66	74.00	79.17	76.56	83.33	30	Radial Basis
With Feature selection	MFCC parameters	66.66	64.81	70.83	67.82	83.33	50	Linear

TABLE IV. PERFORMANCE OF RANDOM FOREST BEFORE AD AFTER APPLYING FORWARD SELECTION METHOD

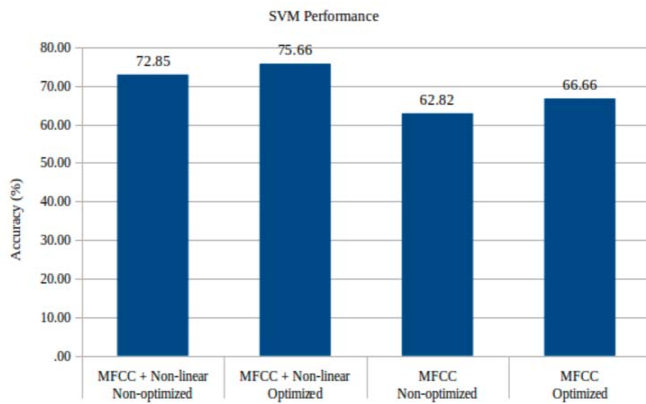| Wrapper	Features	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)	Precision (%)
Without Feature selection	MFCC + Non-linear parameters	76.56	68.75	72.72	70.74	84.61
With Feature selection	MFCC parameters	68.57	60.71	63.63	62.17	80.95
Without Feature selection	MFCC + Non-linear parameters	80.00	80.43	79.17	79.80	88.09
With Feature selection	MFCC parameters	69.33	68.62	70.83	69.72	83.33

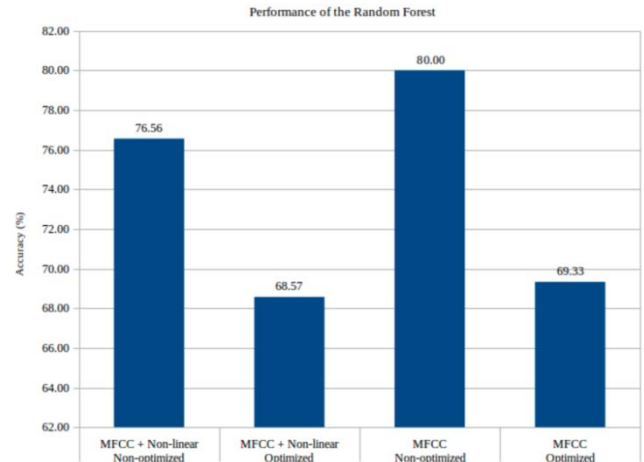Fig. 3. SVM Performance in Terms of Accuracy(%).

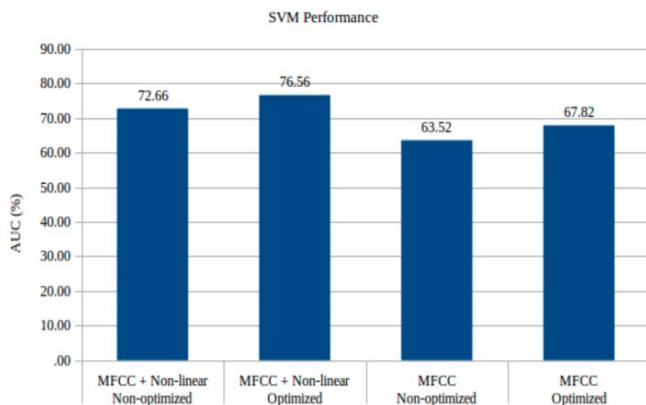Fig. 5. Performance of Random Forest in Terms of Accuuracy(%).
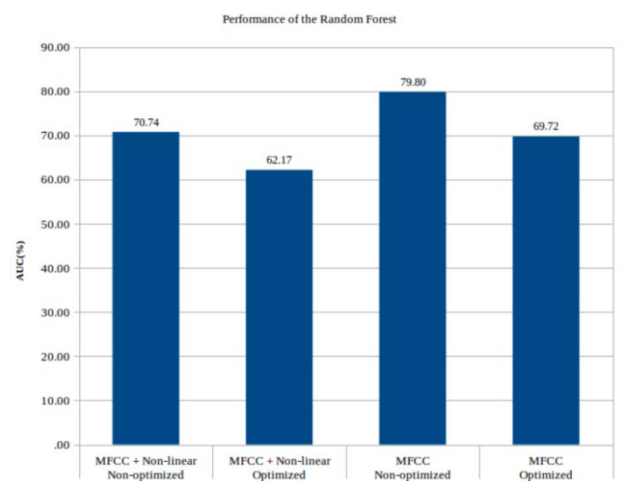

Fig. 4. SVM Performance in Terms of AUC(%).

Fig. 6. Performance of Random Forest in Terms of AUC(%).

V. Conclusion and Future Scope

The paper presents a non-invasive laryngeal cancer detection platform. Both linear and non-linear features are tested over 110 LC cases. Features are optimized with the forward selection method. The system's performance is evaluated with SVM and random forest. The main findings are:

- Hyper-parameter tuning using grid search, helped in the identification of significant features.

- Better optimization of features resulted in improved accuracy of 76.56% with SVM and 80% in the case of random forest.

- Better discriminating abilities are observed with 75.66% with SVM and 79.80% with random forest.

The observations suggest that the SVM with RB kernel and forward selection of features involving non-linear parameters can be used for the development of a more enhanced non-invasive diagnostic tool for laryngeal cancer. With such findings along with more significant features and deep learning techniques, a better diagnostic tool can be developed for the detection of laryngeal cancer.

Acknowledgment

References

[1] G. Cummings: Otolaryngology:" Head & Neck Surgery", 4th ed. Mosby.Inc,2005.

[2] Om Karlsen, Lorentz Sandvik, John-Helge Heimdal, Hans Jorgen Aarstad, " Acoustic Voice Analysis and Maximum Phonation Time in Relation to Voice Handicap Index Score and Larynx Disease", Journal of Voice, Vol. 34, No. 1, 2020.

[3] Young Ae Kang,Jae Won Chang,Ho-Ryun Won,Bon Seok Koo, " Comparison Between Early Glottic Carcinoma and Epithelial Dysplastic Lesions of the Vocal Fold Via Voice Analysis", 24 April 2020, https://doi.org/10.1016/j.jvoice.2020.03.005.

[4] Stefan Hadjitodorov, Boyan Boyanov, and Bernard Teston,"Laryngeal Pathology Detection by Means of Class-Specific Neural Maps", IEEE Transactions On Information Technology in Biomedicine, Vol. 4, No. 1, March 2000.

[5] Adas Gelzinis, Antanas Verikas, Evaldas Vaiciukynas, Marija Bacauskiene, "Exploring sustained phonation recorded with acoustic and contact microphones to screen for laryngeal disorders", 978-1-4799-4527-6/14/2014 IEEE.

[6] Daria Hemmerling, "Voice Pathology Distinction Using Autoassociative Neural Networks", 2017 25th European Signal Processing Conference (EUSIPCO).

[7] Matthew R. Naunheim, Jonathan Garneau, Chris Park, Linda Carroll, Leanne Goldberg, Peak Woo, "Voice Outcomes After Radiation for Early-Stage Laryngeal Cancer", Journal of Voice, Volume 34, Issue 3, May 2020, Pages 460-464.

[8] Priya Dubey, Jacqueline B. Fernandes, Mahesh Bhat,"Acoustic Analysis of Voice in Laryngopharyngeal Cancers Pre and Post Radiotherapy", Indian Journal of Otolaryngology and Head & Neck Surgery, July 2020.

[9] Saime Sagiroglu, Neslihan Kurtul, "The Effect of Supraclavicular Radiotherapy on Acoustic Voice Quality Index (AVQI), Spectral Amplitude and Perturbation Values", Journal of Voice, Volume 34, Issue 4, July 2020, Pages 649.e7-649.e13.

[10] Taito, M., Taito, S., Banno, M. et al. "Voice rehabilitation for laryngeal cancer after radiotherapy: a systematic review and meta-analysis", Eur Arch Otorhinolaryngol 276, 1573–1583 (2019). https://doi.org/ 10.1007/s00405-019-05452-2.

[11] Cyril Plapous, Member, IEEE, Claude Marro, and Pascal Scalart," Improved Signal-to-Noise Ratio Estimation for Speech Enhancement", IEEE Transactions on Audio, Speech and Language Processing, Institute of Electrical and Electronics Engineers, 2006.

[12] HyunBum Kim, Juhyeong Jeon, Yeon Jae Han , YoungHoon Joo , Jonghwan Lee ,Seungchul Lee and Sun Im, ,"Convolutional Neural Network Classifies Pathological Voice Change in Laryngeal Cancer with High Accuracy", J. Clin. Med. 2020, 9, 3415; doi:10.3390/jcm9113415.

[13] Jack J. Jiang, Yu Zhang, and Clancy McGilligan," Chaos in Voice, From Modeling to Measurement", Journal of Voice, Vol. 20, No. 1, pp. 2–17, 0892-1997/$32.00, 2006 The Voice Foundation, doi:10.1016/j.jvoice. 2005.01.001.

[14] Patricia Henríquez, Jesús B. Alonso, Miguel A. Ferrer, Carlos M. Travieso, Juan I. Godino L lorente, and Fernando Díaz-de-María, " Characterization of Healthy and Pathological Voice Through Measures Based on Nonlinear Dynamics", IEEE Transactions On Audio, Speech, And Language Processing, Vol. 17, No. 6, August 2009.

[15] Iasonas Kokkinos, Student Member, IEEE, and Petros Maragos, Fellow, IEEE," Nonlinear Speech Analysis Using Models for Chaotic Systems", IEEE Transactions On Speech And Audio Processing, Vol. 13, No. 6, November 2005.

[16] Girish Chandrashekar,Ferat Sahin,"A survey on feature selection methods", Computers & Electrical Engineering,Volume 40, Issue 1, January 2014, Pages 16-28Brant, Rollin. Forward Selection. MDSC 643.02 Lecture Materials. https://www.stat.ubc.ca/~rollin/teach/643w04/lec/node41.html on July 7, 2018.

[17] Gabriel Solana-Lavalle, Juan-Carlos Galan-Hernandez, Roberto Rosas-Romero " Automatic Parkinson disease detection at early stages as a pre-diagnosis tool by using classifiers and a small set of vocal features",Biocybernetics and Biomedical Engineering, 40 (2020) 505-516.

[18] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin," A Practical Guide to Support Vector Classification", Department of Computer Science, National Taiwan University, Taipei 106, Taiwan, May 19, 2016, http://www.csie.ntu.edu.tw/~cjlin.

[19] Iwan Syarif , Adam Prugel-Bennett, Gary Wills, " SVM Parameter Optimization Using Grid Search and Genetic Algorithm to Improve Classification Performance", TELKOMNIKA, Vol.14, No.4, December 2016, pp. 1502~1509, ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v14i4.3956.

[20] Adele Cutler, D. Richard Cutler and John R. Stevens,"Random Forests", Chapter in Machine Learning, January 2011, DOI: 10.1007/978-1-4419-9326-7_5.

[21] Zuzana Dankovi cova, David Sovak, Peter Drotár and Liberios Vokorokos, "Machine Learning Approach to Dysphonia Detection", Appl. Sci. 2018, 8, 1927; doi:10.3390/app8101927.