Uncertainty handling using Improvised Intuitionistic fuzzy ANN based Voice Disorder Detection

P. Kokila, G. M. Nasira

Abstract: The voice pathology detection is one of the essential process which has to be determined in the earlier stages because it is a sign for raising health related problems. The aim of this paper is to handle the uncertainty in voice dataset due to inconsistency in extracting potential features and vagueness in dealing voice signals. The raw voice signals are preprocessed by feature extraction using meyer wavelet and potential features involved in voice disorder detection are done using sequential forward feature selection methods as voice preprocessing. This research work introduced an improvised intuitionistic fuzzy artificial neural network which enhances the process of voice disorder detection is SVD database by using analytical hierarchical processing for assigning weights and thus the complete neural network performance was fine tuned instead of assigning the weights randomly. The simulation results proved the performance of the proposed model as best by producing more promising result while comparing with ANN, PANN and Fuzzy ANN models.

Keywords: Voice pathology, Artificial Neural Network, Feature extraction, Classification, Fuzzy, Intuitionistic Fuzzy, uncertainty

I. INTRODUCTION

In recent days 25 percentage of world population are under the risk of pathological voice issues, specifically persons whose profession lies on speech, they are excessively louder than normal level at the victims of voice disorder. For example, teachers, instructors, lawyers, actors and singers are also considered to create substantial difficulties in voice. As a result, in speech signals using digital processing has been determined to offer a noninvasive investigative method that is taken in to the account as a prominent assisting tool to medical experts during discovery of voice disorders in earlier stages. Voice pathologies create issues on vocal folds during the process of phonation. It makes vocal folds generating abnormal vibrations owing to faulty results of various factors offering to vocal ambiances. Vocal folds are invariably exaggerated by vocal fold abnormal consequences in different vocal folds vibratory series, this is due to the capability to be locked correctly is reduced. The shape of the vocal tract is also affected by voice disorders and yieldindiscretions in spectral characteristic [1]. It is also recognized that there is no intralaryngeal affect on vocal tract through generation of vowel which is considered that foundation of voice has immeasurable confrontation.

As indicated about, voice pathologic is a sign of health-related issues. However, discovering dysphonic voice at earlier stages is not an easy process. Generally, a trained expert is needed and a serious of speech exercise is essential. Automatic speech assessment can permit for cost effective and time consumption for analyzing speech and used for screening of large volume of population. The foremost unbiased realization of this exploration work is utilizing machine learning for the diagnostiction of voice disorder persons. The definitive goal is to devise an improvised intuitionistic fuzzy back prorogation enabled neural network for classification of pathological voice and normal voice.

II. RELATED WORK

There are many existing works on speech pathology detection, researchers are mainly focusing in deploying methods on feature extraction and classification of pattern by developing models.

The capability of acoustic features such as formants, pitch, jitter and shimmer in the discernment of normal voices from the pathological voices are conversed by Srinivasan et al [2].

The recognition of voice pathology from voice database is deployed using support vector machine and prior to that feature selection is done using genetic algorithm which produces best feature subset in order to progress the accuracy of classification.

Pravena et al [3] in their work introduced feature extraction using Mel frequency cepstral coefficient (MFCC) and Gaussian Mixture Model (GMM). The acoustic features of speech signals are greatly involved in discovery of vocal fold pathology using the help of speech signals recorded from the voice disorder and normal peoples.

Vikram and Umarni [4] investigated an approach which integrates the MFCC based feature extraction and classification of voice disorders using SVM for diagnosing neurological disorders in the speech signals.

Vahid Majidnezhad and Igor Kheidorov [5] developed a voice disorder detection model using Hidden Markov Model which categorizes speeches into two different classes namely normal voice and pathological voice. Two hidden markov models were involved in training the two different classes of speech and then using the trained models the testing dataset of voice is applied to classify them.

Cheolwoo Jo [6] analyzed the vocal quality of a particular group of women and men with presence or absence of voice disorders. It is done based on the assessment of a group of experts who are experienced in the area of vocal rehabilitation.
Mandar Gilke [7] used modulation spectra to detect pathological voice and classify them either pathology or normal voice. The voice source parameters are analyzed to discover the influence on voice source from the database of pathological voice.

El Emary et al. [8] in their work classified the speech signal with MFCC, jitter and shimmer as important parameters to measure the voice disorder. The discovery of voice issues from neurological disorders was done by applying GMM method on a very small set of SVD database which comprised of 63 healthy voice and 38 pathological voices.

Fonseca et al [9] developed a Daubechies discrete wavelet transform, linear prediction coefficients and last squares support vector machine (LS-SVM) was used to recognize laryngeal pathologies. The experiments were passed out using a private database.

Laura Verde et al [10] investigated on voice pathology detection by applying various machine learning approaches. The dataset is collected from Saarbruecken Voice Database. The results are validated using support vector machine and decision tree which depends on the feature selection methods.

Kim et al [11] introduced a Long short term memory (LSTM) which is special kind of recurrent neural network which is extensively cast-off for long term needs. LSTM for voice activity detection separates the incoming speech with noise.


III. PROPOSED METHODOLOGY OF UHIFANN

This research work handles the issues of uncertainty in detection of pathological voices in presence of vague information about the patients. The dataset used in this work is collected from Saarbrucken voice database (SVD) [15]. The collection of voice recording used in this work is 1040. The length of the audio clip with sustained vowels /a/, /i/, /u/ is 2 seconds. Initially, the collected raw dataset is preprocessed by noise removal technique to enhance the quality of the voice and improve the accuracy of classification. To determine whether the selected voice is a healthy or pathological voice its features are extracted using Meyer wavelet transform which generated the important features of the voice signals. In order to handle the voluminous feature set, sequential forward feature subset selection is applied to choose the potential features which mainly contributes in detection of pathological voices more precisely. To handle the uncertainty in classifying vague instances this work introduced an improved intuitionistic fuzzy artificial neural network which is comprised of five different layers for processing the input of the voice dataset and produces the classification of input as healthy voice or pathological voice with more accuracy.

![Overall Framework of Uncertainty handling using Improvised Intuitionistic fuzzy ANN based Voice Disorder Detection](image1)

**Improved Intuitionistic fuzzy ANN for Voice Disorder detection**

In this section the process and principle of intuitionistic fuzzy artificial neural network is explained in detail for detection of voice disorder detection. This model encompassed of four different layers to achieve its functionality. The figure 2portrays the construction of the IFANN.

![Architecture of Intuitionistic fuzzy artificial neural network](image2)

In this work IIIFANN consist of four different layers. The input layer is the first layer, receives the input of the voice dataset after performing feature subset selection the potential attributes selected are nine features and they are involved in the process of pathology voice detection. The second layer, performs conversion of given input values which are in crisp format are converted into intuitionistic fuzzy value in terms of degree of membership and non-membership values as follows.
In third layer for a single neuron with input i of IFANN it is represented in intervals \([0,1]\) \([13,14]\). The weight \(w_t\) assigned to the link between previous layer node and current layer node is multiplied with the input \(i\) using the following formula:

\[
\mu_0(y) = \begin{cases} 
0; & y \leq e \\
\frac{y-e}{f-e} & e < y \leq f \\
\frac{g-e}{g-f} & f \leq y < g \\
0; & y \geq g 
\end{cases}
\]

\[
\nu_0(y) = \begin{cases} 
1 - \varepsilon; & y \leq e \\
1 - \frac{y-e}{f-e} & e < y \leq f \\
1 - \frac{g-e}{g-f} & f \leq y < g \\
1 - \varepsilon; & y \geq g 
\end{cases}
\]

The weight assignments are done under trail and error basis in back propagation method to improve the performance of the network this work introduced Analytical hierarchical processing for assigning weights among hidden layers.

IV. RESULTS AND DISCUSSIONS

The proposed model is deployed using MATLAB tool. The dataset for voice dataset feature subset selection is done using SVD \([15]\). The number of instances used in this work is 1040. The meyer wavelet transform is used for extracting the features from voice signals and from the extracted features, only potential attributes are selected using Sequential Forward Subset Selection. Among 26 features extracted by meyer wavelet only 9 features are used by sequential forward Subset selection method. The input for the Improved Intuitionistic fuzzy ANN is those 9 features namely Jitter (ppq5), Shimmer (localdb), Shimmer (apq5), Shimmer (apq11), HTN, Standard deviation, Maximum pitch, Number of pulses, Mean period are used for performing the classification process.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>87.60%</td>
<td>84.30%</td>
<td>82.10%</td>
<td>83.60%</td>
</tr>
<tr>
<td>PANN</td>
<td>90.10%</td>
<td>89.60%</td>
<td>88.50%</td>
<td>89.20%</td>
</tr>
<tr>
<td>FANN</td>
<td>91.70%</td>
<td>90.30%</td>
<td>91.60%</td>
<td>90.80%</td>
</tr>
<tr>
<td>IIFANN</td>
<td>98.20%</td>
<td>97.60%</td>
<td>98.50%</td>
<td>97.80%</td>
</tr>
</tbody>
</table>
Uncertainty handling using Improvised Intuitionistic fuzzy ANN based Voice Disorder Detection

From the table it is examined that the involvement of Improvised Intuitionistic Fuzzy Artificial neural network performs better than the other three existing approaches. Because the weight assignment on each hidden nodes of hidden layer are not performed in optimized manner in conventional neural network, probabilistic neural network and fuzzy neural network. The accuracy of proposed IFANN is 98.2%, while fuzzy artificial neural network is 91.7% this is because the use of degree of membership, non-membership and hesitation are the prominent reason which handles both inconsistency and vagueness of the voice dataset where the fuzzy can able to define them only with degree of membership.

Table 2: Performance Comparison based on Time Taken

<table>
<thead>
<tr>
<th>Methods</th>
<th>Time Taken(seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>40</td>
</tr>
<tr>
<td>PANN</td>
<td>35</td>
</tr>
<tr>
<td>FANN</td>
<td>25</td>
</tr>
<tr>
<td>IFANN</td>
<td>12</td>
</tr>
</tbody>
</table>

The ultimate motivation of this research work is to handle the vagueness and inconsistency which prevails in the real time dataset of voice signals. This work introduces improvised intutionistic fuzzy artificial neural network whose aim is to produce more accurate results for pathological voice detection even in case of uncertainty. The use of analytical hierarchical processing improves the IFANN to assign the optimal weights on the hidden nodes so the entire network produces better accuracy while comparing the existing model and it is proved with the obtained simulation results.

V. CONCLUSION

The ultimate motivation of this research work is to handle the vagueness and inconsistency which prevails in the real time dataset of voice signals. This work introduces improvised intutionistic fuzzy artificial neural network whose aim is to produce more accurate results for pathological voice detection even in case of uncertainty. The use of analytical hierarchical processing improves the IFANN to assign the optimal weights on the hidden nodes so the entire network produces better accuracy while comparing the existing model and it is proved with the obtained simulation results.

REFERENCES


Figure 6: Performance Analysis of Four different classification Models for voice pathology Detection

Figure 7 Performance Analysis based on Time taken

From the table and the figure, it is observed that the time taken by this research methodology Improvised Intuitionistic fuzzy ANN is very less while comparing the other three methods. This is because the optimization of the weights assigned on the hidden nodes are done using analytical hierarchical processing so that the number of iterations is greatly avoided to correct the observed and expected results differences.

AUTHORS PROFILE

P. Kokila, pursuing PhD in Chikkanna Government Arts College Tirupur. I have completed my master degree in MSC IT in Kongu Engineering college at Perundurai and Master of Philosophy in GRG Krishnammamal college at Coimbatore. I have 2 years 6 months working experience as Assistant professor in department of Computer science in Park’s College, Tirupur. My article is published in reputed journals such as UGC and Springer and interested to publish about neural networks and speech recognition. Currently I am doing my research in voice pathology detection.

Dr. G. M. Nasira graduated in Computer Science in Madras University. She completed her post graduate M.C.A. in Bharthidasan University and M.Phil degree in Bharathiar University. Finally she awarded Doctoral degree in Computer Science from Mother Teresa University. At present she is working as Assistant Professor & Head of Computer Applications in Chikkanna Government Arts College, Tiruppur. She has more than 20 years of teaching experience. Her interesting area of research is Fuzzy logic and Data mining. Besides, she has published more than 100 referred journals and acted as a potential reviewer to her credit. She has delivered invited lectures in the various conferences and seminars. Under her Research Supervision more than 3 Ph.D Awarded. She got awards for best teaching also.