International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-2, July 2019

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 faultyresults of various factors offering to vocal ambiances. Vocal folds are invariantly exaggerated by vocal fold abnormal consequences in different vocal folds vibratory series, this is due to the capability to be lockedcorrectly 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 foremostunbiased realization of this exploration work is utilizing machine learning for the diagenesis 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.

Revised Manuscript Received on 30 July 2019.

* Correspondence Author

P. Kokila*, Research Scholar, Department of Computer Science, Chikkanna government Arts college, Tirupur-641602, Tamil Nadu, India

Dr. G. M. Nasira, Professor & Head, Department of Computer Applications, Chikkanna Government Arts College, Tirupur-641602, Tamil Nadu, India.

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 Umarani [4] investigated an approach which integrates the MFCC based feature extraction and classification 0f 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 Daubechiesdiscrete wavelet transform, linear prediction coefficients andlast squares support vector machine (LS-SVM) was used torecognize laryngeal pathologies. The experimentations 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 extensivelycast-off for long term needs. LSTM for voice activity detection separates the incoming speech with noise.

Kim et al [12] in their work devised a Convolutional neural network-based LSTM to discover dysarthric speech disorder. Recurrent neural networks are widely used for speech recognition, music genre classification, natural language processing and sequence prediction problem.

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.

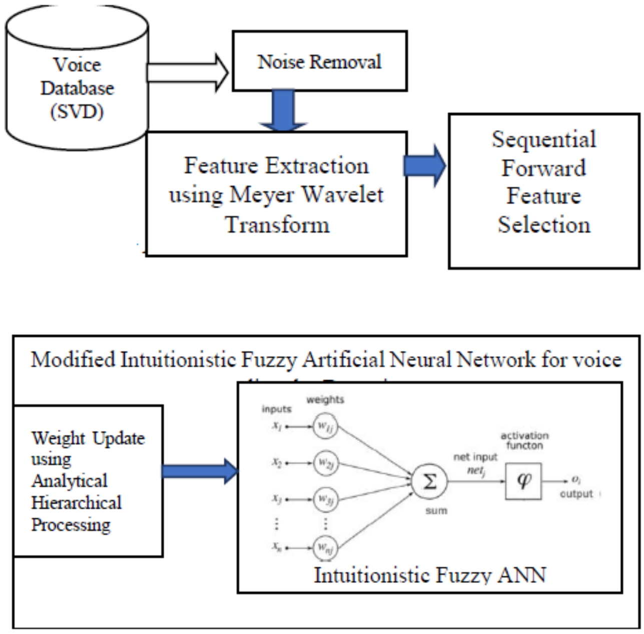


Figure 1: Overall Framework of Uncertainty handling using Improvised Intuitionistic fuzzy ANN based Voice Disorder Detection

Improvised 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.

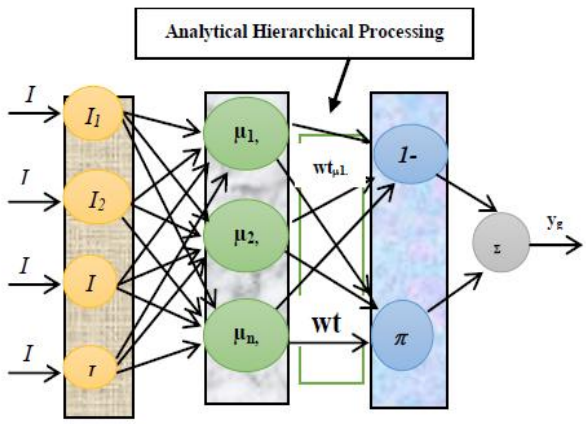
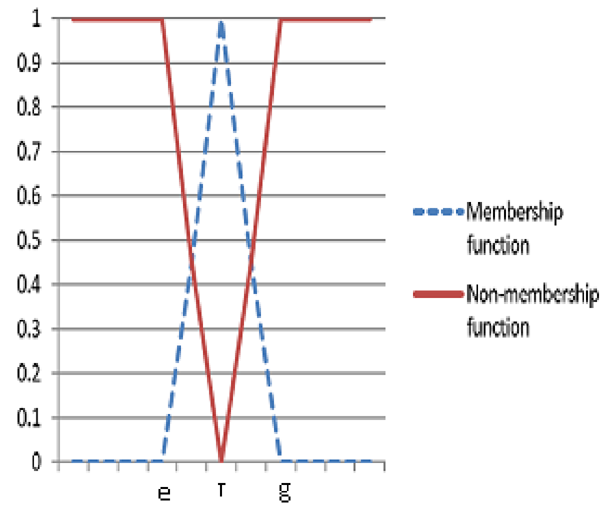


Figure 2: Architecture of Intuitionistic fuzzy artificial neural network

In this work IIFANN 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.

$$\mu_D(y) = \begin{cases} 0; & y \le e \\ \left(\frac{y-e}{f-e}\right) & e < y \le f \\ \left(\frac{g-e}{g-f}\right) & f \le y < g \\ 0; & y \ge g \end{cases}$$

$$\nu_D(y) = \begin{cases} 1-\in; & y \le e \\ 1-\left(\frac{y-e}{f-e}\right) & e < y \le f \\ 1-\left(\frac{g-e}{g-f}\right) & f \le y < g \\ 1-\in; & y \ge g \end{cases}$$



In third layer for a single neuron with input i of IFANN it is represented in intervals [0,1] [13, 14]. The weight wt assigned to the link between previous layer node and current layer node is multiplied with the input i using the following formula:

$wt_i \equiv \langle \mu_i, \nu_i \rangle = \langle \max(\mu_w, \mu_i), \min(\nu_w, \nu_i) \rangle$;

The fourth layer is a summator Σ. The other element that is passed to the summation, is $\mu_b + \nu_b \le 1$

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.

Weight Optimization using Analytical Hierarchical Processing (AHP)

AHP splits the problems into sub problems of order set. This technique handles the issue by handling sub problems very easily and computation is also done with less complexity. Depending on the input vector's relative intensity normalized eigen vector is computed[16]. Weights with input vector is known as normalized eigen vector. Comparison among Relative Index (RI) and Consistency Index (CI) is done. The Consistency Ratio (CR) is obtained by dividing CI with RI. While CR doesn't match with need level which is greater than 0.1 then pairwise comparison matrix will be substituted [17]. Input constraints priorities are assigned by AHP. Pairwise comparison is easier and more straightforward. It is mainly used for handling inconsistency in input dataset. AHP is used in decision making problems like discovering best combination, finding relative advantage of a set of alternatives, depending on variety of limits.

Procedure of Analytical Hierarchical Processing

```
Input Values: B[i] [j]
Output: weight Wt[i]
Start
D[i] = E[i][j] = S[i]= CNI=0
For p= 1 to j
 For q=1 to i
D[n] = B[p][q]+ D[q]
 For p = 1 to i
For q= 1 to j
E[p]= E[p]+ (B[p][q] / D[q])
For p= 1 to i
For q= 1 to j
E[m]= E[m]/j
For p= 1 to i
For q= 1 to j
 Wt [p]= E[p] * B[p][q]
For p = 1 to i
S = wt[p] +S
CNI = (S – i) / (i-1)
If CNI<0.1
Assign obtained weights to hidden nodes
Else
Change Value of Weights
end
```

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 form 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 1 : Performance Comparison of four different models based on the evaluation metrics

Methods	Accuracy	Precision	Recall	F-measure
ANN	87.60%	84.30%	82.10%	83.60%
PANN	90.10%	89.60%	88.50%	89.20%
FANN	91.70%	90.30%	91.60%	90.80%
IIFANN	98.20%	97.60%	98.50%	97.80%

Retrieval Number: A1424058119/19©BEIESP
DOI: 10.35940/ijrte.A1424.078219
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

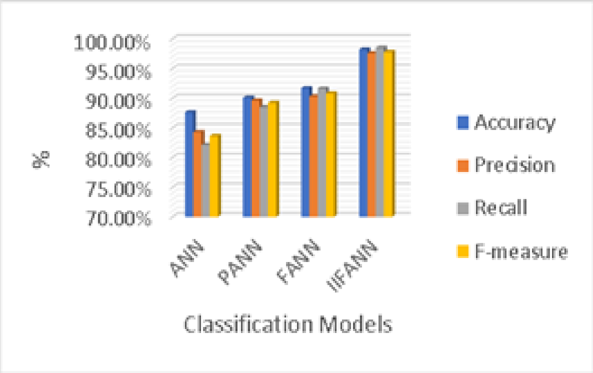


Figure 6: Performance Analysis of Four different classification Models for voice pathology Detection

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

Methods	Time Taken(seconds)
ANN	40
PANN	35
FANN	25
IIFANN	12

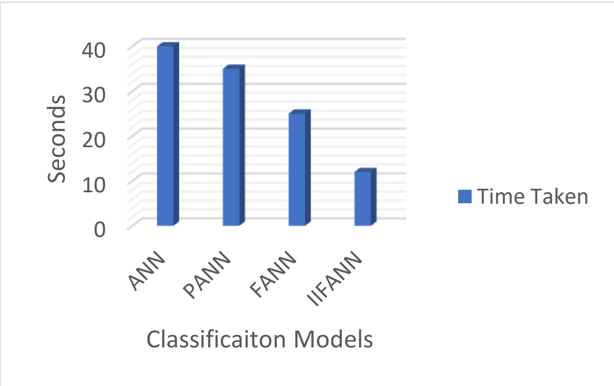


Figure 7 Performance Analysis based on Time taken

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.

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 improved intuitionistic 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

1. G. Muhammad, G. Altuwaijri, M. Alsulaiman, Z. Ali, T. A. Mesallam, M. Farahat, K. H. Malki, and A. Al-nasheri, "Automatic Voice Pathology Detection and Classification Using Vocal Tract Area Irregularity," Biocybernetics and Biomedical Engineering, vol. 36, pp. 309-317, 2016.
2. V. Srinivasan, V. Ramalingam, V. Sellam, "Classification of Normal and Pathological Voice using GA and SVM", International Journal of Computer Applications, Volume 60 – No.3, December 2012.
3. D. Pravena, S. Dhivya, A. Durga Devi, "Pathological Voice Recognition for Vocal Fold Disease", International Journal of Computer Applications, Vol-47, No.13, 2012.
4. C.M. Vikram and K. Umarani, "Pathological Voice Analysis to Detect Neurological Disorders Using MFCC & SVM", International Journal of Advanced Electrical and Electronics Engineering, Volume-2, Issue-4, 2013.
5. Vahid Majidnezhad, Igor Kheidorov, "A HMM-Based Method for Vocal Fold Pathology Diagnosis", International Journal of Computer Science, Vol. 9, Issue 6, No 2, November 2012.
6. Cheolwoo Jo, "Source Analysis of Pathological Voice", Proceedings of the international Multi conference of Engineers and Computer Scientists, Vol II, IMECS March 2010. [6] Alireza A, Dibazar, Theodore W. Berger, and Shrikanth S. Narayanan,"Pathological Voice Assessment", IEEE EMBS 2006,New York.
7. Mandar Gilke, Pramod Kachare, Rohit Kothalikar,Varun Pius Rodrigues, Madhavi Pednekar, "MFCC-Based Vocal Emotion Recognition Using ANN", International Conference on Electronics Engineering and Informatics (ICEEI), vol 49, 2012, Singapore.
8. M. El Emary, M. Fezari, and F. Amara, "Towards developing a voice pathologies detection system," Journal of Communications Technology & Electronics, vol. 59, no. 11, p. 1280, 2014.
9. E. S. Fonseca, R. C. Guido, P. R. Scalassara, C. D. Maciel, and J. C. Pereira, "Wavelet time-frequency analysis and least squares support vector machines for the identification of voice disorders," Computers in Biology and Medicine, vol. 37, no. 4, pp. 571–578, 2007
10. Laura Verde, Giuseppe De Pietro, Giovanna Sannino, Voice Disorder Identification by using Machine Learning Techniques, IEEE Access, 1-11, 2169-3536, 2018
11. Kim, J., Kim, J., Lee, S., Park, J., & Hahn, M. (2016, November). Vowel based voice activity detection with LSTM recurrent neural network. In Proceedings of the 8th International Conference on Signal Processing Systems (pp. 134-137).
12. Kim, M., Cao, B., An, K., & Wang, J. (2018). Dysarthric Speech Recognition Using Convolutional LSTM Neural Network. Proc. Interspeech 2018, 2948-2952.
13. Krassimir T. Atanassov, Fuzzy Sets and Systems, North-Holland, Volume 20 (1986), pages 87-96, ISSN 0165-0114
14. Krassimir T. Atanassov, Series "Studies in Fuzziness and Soft Computing", Volume 35, Springer Physica-Verlag, 1999, ISBN 3-7908-1228-5
15. Barry, W.J., Pützer, M. Saarbrücken Voice Database, Institute of Phonetics, Univ. of Saarland, http://www.stimmdatenbank.coli.unisaarland.de/
16. Geoff Coyle: Practical Strategy (2004) Open Access Material. AHP © Pearson Education Limited

17. M.L. Caliusco and G. Stegmayer, "Semantic Web Technologies and Artificial Neural Networks for Intelligent Web Knowledge Source Discovery", Advanced Information and Knowledge Processing, Springer-Verlag London, 2010.

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 Krishnammal 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 Fuzzuy 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.

Retrieval Number: A1424058119/19©BEIESP
DOI: 10.35940/ijrte.A1424.078219
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication