Journal of Theoretical and Applied Information Technology
31st January 2019. Vol.97. No 2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

TWO SCOPES OF ACOUSTIC SIGNAL AND FUZZY-RELIEF ALGORITHM FOR IMPROVING AUTOMATIC SPEAKER RECOGNITION

[1]EMAN S. AL-SHAMERY, [2]WAFAA M. AL-HAMEED

[1,2]Department of Software, Faculty of Information Technology, University of Babylon, Iraq
Email: [1]emanalshamery@itnet.uobabylon.edu.iq, [2]it.wafa.m.s@uobabylon.edu.iq

ABSTRACT

Speaker recognition is one of the important biometrics methods that have entered into many applications such as security, marketing service, and bank transfers. The main aim of this paper is to identify the speaker with high accuracy through his or her voice. All previous research deal with recorded files for speakers as only sound signals. This research introduces a new idea for dealing with the recorded sounds in two scopes, one of which is dealing with the signal as an acoustic file and the other as an image, which is picked from the sound signal. These files are analyzed using Discrete Wavelet Transform (DWT), specifically (Daubechies) db1 at three levels. Additionally, another contribution to the development of the Relief method for feature selection has been proposed by including a fuzzy inference system. The proposed Fuzzy-Relief method divides the features into three groups that are ordered according to the importance degree from the best up to the worst depending on the new membership function. The first two groups are taken into account in the recognition process and neglected the third group. Furthermore, The Logistic Regression (LR) method and Multi-Layer Perceptron (MLP) neural network are employed with 3-cross-validation for tspeaker recognition. The proposed system has been applied to the twenty speakers, ten females and ten males, ten recordings and two different sentences for each speaker, in normal room circumstances. The system is evaluated using accuracy and some other measures resulting from the confusion matrix. After a comparison between the two scopes, the recognition accuracy of acoustic is varied from 78% as a worst to 96% as a best with a reduction percentage of features reaches 62.5%. While, for image files, the recognition accuracy is ranged from 92% to 100% with the reduction percentage reaches to 78.9%. In general, the results of LR are better than MLP, and the results of Image files are much better than acoustic files.

Keywords: *Discrete Wavelet Transform, Fuzzy-Relief Algorithm, Image of Sound, Logistic Regression, Multi-Layer Perceptron.*

1. INTRODUCTION

There are many important areas for automatic speaker recognition such as financial services, transaction, biometric authentication, Forensics on intercepted calls, and access control for computers and data networks [1], [2], [3].

Speaker recognition means identification or verification the speaker identity through his or her voice [4]. Speaker identification has been implemented by matching the sound to all stored sounds to give a decision while in speaker verification one matching with only one sound will be implemented and the result acceptance or rejection of this identity [3]. Both need to be determined whether the text is dependent or independent. In the first approach, the speaker is restricted by pronouncing specific words or phrases that are adopted in training and testing, the second approach does not restrict the speaker in specific words or phrases. Thus, the first is easier and more secure where these words or sentences can be considered as passwords. While the second approach is harder but generalizable and more flexible [5].

Many of the difficulties that face us in this field, including the change of sound as a result of aging,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

emotional and psychological state, sick and others [2]. To overcome these difficulties, there is a need to be careful in extracting appropriate parameters in the analysis process.

There are two types of characteristics of speech analysis, one of which reflects the physical structure of the vocal tract and the other depends on the behavioral of the speaker [5]. Accordingly, linear predictive coding, filter-bank, cepstral, pitch periodicity and other coefficients of speech analysis. All of these features are considered low-level features, which are related to the limits of the brain's perception of speech [6].

In this study, the Discrete Wavelet Transform (DWT) coefficients are exploited for speech analysis because they have a locality characteristic in the time domain and scale characteristic in the frequency domain [7].

On the other hand, it is important to select an important classifier in the recognition process. Hence, machine learning techniques have improved decision-making in many areas, including the field of automatic speaker recognition [2]. Thus here, two methods are adopted for speaker recognition Logistic Regression (LR) and Multi-Layer Perceptron (MLP) artificial neural network.

This paper is organized as follows:
Section2: reviews related work. Section 3: shows a theoretical background about the feature analysis using DWT, fuzzy system, relief algorithm of feature selection, LR and MLP for recognition. Section 4: states the proposed system which contains the proposed architecture in addition to the proposed Fuzzy-relief algorithm. Experimental results and discussion are documented in Section 5. In Section 6, the most interesting conclusions are summarized.

2. RELATED WORK

Many studies have considered the speaker recognition problem.

In [8], the authors proposed a modification of spectral features of speech analysis called normalized dynamic spectral features resulted for treating with the noisy signal in the text independent system. In addition, K-NN is employed for identification, A comparative study with many levels of known feature analysis methods, the authors proved that the proposed method is robust and it enhanced the recognition accuracy.

The K-NN needs time .additionally, the intera-features of speakers have a negative effect on the recognition process.

In [9], wavelet packet was introduced as an alternative way from the standard features method such as Mel Frequency Cepstral Coefficient features (MFCC). The authors applied wavelet packet without following Mel scale. Multi-resolutions at seven layers to get a new feature. Also, the focus was on the structure of the filter. The proposed system was applied to a closed set. The results are compared with traditional features.

The wavelet features are so good but, for level7, the accuracy will decrease and the main features will be lost.

The authors in [10] worked on forensic speaker recognition on the Arabic language as a data set. The standard MFCC features are used in feature extraction while the Gaussian mixture model-universal is used for the classification the objective of the last model is to reduce error. The samples are tested in a noisy environment and in a short time.

The model needs to extract all eigenvectors values. Which requires the use of additional techniques to solve this problem and thus leads to increased complexity of the system.

The authors in [11] compared between two types of features MFCC and prosodic features for speaker verification. In order to evaluate the proposed system, three cases have been investigated. These cases resulted from taking every type separately and the third case combined between the two types. The results were better and more accuracy when the features are combined.

However, the MFCC features are better from prosodic, the latter features suffer from many disadvantages such as the difficulty of identifying the part of the signal that containing important information and determining an appropriate model of calculation as well as what is the amount of the robust and efficiency when combined with the other characteristics.

IN order to enhance automatic speaker recognition, the speech recognition is applied to identify the speaker identity in [12]. The MFCC

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

coefficients are extracted for speech analysis then hidden Markov model is employed to identify the word by computing the maximum likelihood values for the spoken word.

However, the hidden Markov model has a large number of unstructured coefficients. Additionally, the model cannot describe the correlations between hidden states.

3. THEORETICAL BACKGROUND

This section reviews the most important characteristic of methods and techniques that are dependent in this work for feature extraction, feature selection, the fuzzy inference system, which is employed to develop feature selection method, and the machine learning techniques for recognition problem.

3.1 Feature Extraction Using Discrete Wavelet Transform

The basic idea behind wavelet is to analyze a finite set of basic functions, each function called mother wavelet. The wavelets have generated using the dilations and translations of the Mother function. Thus, wavelet analysis has characteristic to process data in the time-frequency domain in various resolutions.

Furthermore, DWT has many families such as Daubechies (Db) family according to the numbers and values of filter coefficients. Db wavelet systems are very suite to reflect the polynomial behavior. Moreover, the filter values are convolved on raw data. The coefficients resulted are organized by two leveraged patterns, one that represents the smooth or approximation while the other represents the details [7].

Approximation and details parameters can be produced by one-level signal analysis. Where the coefficients of one level are calculated by means of the signal convolution with the low pass filter and then a down sampling process is performed. As for the details, they are calculated by means of the signal convolution with the high pass filter then applied the down sampling for each level.

In the other side, the DWT can be applied on the digital images in the image processing field. Here, the analysis will be different, let, I (rows, columns) represents the image, the following steps demonstrate the image analysis using DWT:

I. The Low pass filter coefficients have convolved with rows, the result is A signal which represents the approximation coefficients.

II. The High pass filter coefficients have convolved with rows, the result is a D signal which represents the details.

III. Columns of the A are convolved with the high pass filter coefficients to produce the LH band coefficients.

IV. Columns of the D are convolved with the high pass filter coefficients to produce the HH band coefficients.

V. Columns of the D are convolved with the low pass filter coefficients to produce the HL band coefficients.

VI. Columns of A are convolved with the low pass filter coefficients to produce the LL band which represents the most important band that contains the spectral form of each signal.

VII. Repeat all steps for generating a new coefficient to the next level but let LL band represents a new image.

3.2 Feature Selection Methods

Feature selection is a technique to choose a subset of variables from the multidimensional data. It can improve the classification accuracy in diversity datasets. In addition, the best subset selection can reduce the cost of size, time and complexity.

The feature selection methods can be classified into three types.

First: Filter; these methods use a proximity measure such as correlation, distance, consistency and statistical measures to evaluate the features of data. They are applied before the prediction model [13]. Such as: information gain [14], Correlation-based Feature Selection [15], and Relief [16].

Second: Wrapper Methods; are applied after prediction model. They are based on identifying sets of attributes and comparing them together to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

determine the best, however, they most often need search algorithms. The method may be repetitive to determine the best set of properties. The prediction model is employed to decide the accuracy of the attributes. Examples include wrapper: Naive Bayes [13] and most any modeling algorithm combined with a feature subset generation approach.

Third: Embedded Methods; the selection operation is executed through applying the prediction model. Examples include: Lasso, Elastic Net, and various decision tree based algorithms [17].

In this paper, the relief algorithm is developed using fuzzy. The original algorithm was first introduced by Kira and Rendell [18] for dealing with binary classification, then extended by Kononenko [19] to deal with incomplete, missing data, multi-classes and regression. Later many different versions to enhance the classification accuracy are proposed by many authors such as [16],[20],[21].

The basic approach behind Relief is to weight features according to the proximity among objects and matching decision of these objects with classes.

3.3 Fuzzy System

In real applications, three processes are required to implement a fuzzy inference system.

First: Fuzzification which contains applying membership function to convert data from crisp to fuzzy.

Second: Fuzzy inference; the control rules are projected with combining the membership functions to get fuzzy output.

Third: Defuzzification, in this process the fuzzy data is converted to a crisp by using many methods such as the center of gravity, mean of memberships and fuzzy operations [22].

3.4 The Classification Using Machine-Learning Techniques

Logistic Regression (LR) and Multi-Layer Perceptron are applied for identifying the speaker identity.

3.4.1 Logistic regression

In origin, the logistic regression computes the probability of a class that can only have two values (i.e. binary classification). So, for classifying multiclass many classifiers (numerical and categorical) are needed. It forms curve using natural logarithm of the objective variables. The values of this curve are limited between 0 and 1. Maximum likelihood is employed to get the coefficients.

The log likelihood continues in updating until no important change [23],[24].

3.4.2 Multi-layer perceptron

It is one of the most famous of ANN. It has feedforward topology, which consists of an input layer, an output layer, and one or more hidden neurons layers as well as weights that connect between them. MLP uses Backpropagation algorithm for training. There are two important coefficients learning ratio as well as momentum required to be defined by the user. The theoretical results have proved that the one hidden layer is sufficient network to approximate any continuous mapping of input patterns to output patterns. The root mean square error is calculated between desired and actual. The training successes when the error reaches to predefined small value [25].

Many approaches are employed for optimizing of error reduction [26], and training convergence [27].

4. THE PROPOSED SYSTEM

The methodologies of the proposed system include six different stages after the recording of the acoustic file. In the beginning, converting this file to two scopes, normalizing, feature extraction, feature selection, machine learning classifiers, and finally the evaluation as shown in Figure 1.

After recording the acoustic file, the normalization process is applied on the all the signals by 75% to reduce the effect of high and low speaker sound as well as the proximity and distance from the microphone at recording.

Additionally, the silence is removed from the beginning and ending the signal by the next step of processing.

After that, a new acoustic file has been stored as two files in two different formats. First format; the file is stored in “wav” extension. Second format: a copy of the image of the waveform has been stored with “bmp” extension.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

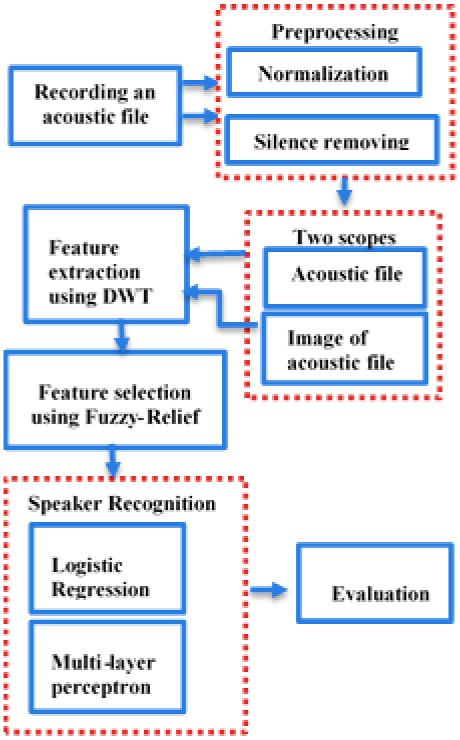

Figure 1: The Proposed System Architecture

At this stage, the db1 of DWT is applied to analyze acoustic file and image file separately at three levels.

The acoustic file is read as a vector, then this vector is divided into frames (the length of each frame = 256 value or sample). Thus, the first level produces 128 High (H) level frequency and 128 Low(L)level frequency coefficients. In the second level, only L coefficients are taken and analyzed to get 64L and 64H coefficients. Again, in the third level, only 64L coefficients are taken and analyzed to get 32H and 32L coefficients. The last coefficients are adopted for the next stage after computing the average of all frames.

The image of the acoustic file is read as a two-dimensional matrix. The image is divided into blocks with size 128 × 128 for each. The analysis is implemented on the rows and columns. The result of the first level is four bands (LL, LH, HL, HH) with 64 * 64 size for each. LL coefficients are taken only for analyzing in the second level to get the same number of bands but with half size for each of them. Finally, the 16*16 LL coefficients resulted which are adopted after computing the average of all blocks.

The fuzzy-relief algorithm is proposed to select the important wavelet features.

Let D represents a two dimensional array (N×M), where, rows are set of N objects $obj_i = (x_i, c_i)$, and columns are a set of features F_j =(F1, . . . , FM). The c_i is the class number which represents the class label of the object. Relief computes the Features weights in the case where c is the class number. Relief computes the Features weights in the case where c is a multiclass variable. The algorithm penalizes the features that indicate different values to neighbors of the same class, and rewards features that indicate different values to neighbors of different classes. The problem, it may give a different rank to the feature, which has the same weight merit. Therefore, the development here is to reassign weight depending on the fuzzy system using a new membership function (step 7 of the algorithm).

====================================

Algorithm Name: Fuzzy-Relief

====================================

Input: let x_{ij} represents the data where rows are the objects and columns are the features, except the last column is the class number (discrete number).

Output: Assign significance degree for each features F_j by forming three groups of features selected :Most- Significant (Mostsig) , Significant(Sig), and Neglected (Neg).

====================================

1. Split data into f folds; f-cross-validation and for each fold do
2. Sets initial values of weights w_j to 0, m=250 and k=10
3. For each i=1 to m do
4. Selects a random object x_i
5. Finds the k-nearest objects x_n to for each class, and updates, for each nearest neighbor x_b, all the weights for the feature selected F_i as:
6. If x_n and x_b ϵ the same class,

$$w_j^{i-1} = w_j^{i-1} - \frac{B_j(x_n, x_b)}{m} . dist_{np}$$

, where:

$$B_j(x_n, x_b) = \begin{cases} 0, & x_{nj} = x_{bj} \\ 1, & x_{nj} \neq x_{nj} \end{cases}$$

$$dist_{np} = \frac{dist1_{np}}{\sum_{l=1}^{k} dist1_{nl}}$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$dist1_{np} = e^{-(\frac{z(n,p)}{2})^2}$$

z(p,q) is the index of the q^{th} object among the nearest neighbors of the n^{th} object sorted by the distance.

End if

Else

$$w_j^{i-1} = w_j^{i-1} + \frac{p_b}{1-p_n} . \frac{B_j(x_n,x_b)}{m} . dist_{np}$$

, where: w_j^{i-1} j is the weight of the feature F_i at the i^{th} iteration step, p_b is the prior probability of the class to which x_b belongs, and p_n is the prior probability of the class to which x_n belongs.

The end for //k

7. Implement the fuzzification process for weighing for each feature using the following membership function:

$$\mu(w,\theta1,\theta2) = \begin{cases} 0, & w \leq 0 \\ 2\left(\frac{w}{\theta2}\right)^2, & 0 < w \leq \theta1 \\ 1-2\left(\frac{w-\theta2}{\theta2}\right)^2, & \theta1 < w \leq \theta2 \\ 1, & w > \theta2 \end{cases}$$

8. Compute fuzzy of weight (FuzzyW) for each feature using the center of gravity for all features weights as :

$$fuzzyW = \frac{\sum \mu(w) * w}{\sum \mu(w)}$$

End for// m

9. Assign the final weights (FinW) after computing the average of fuzzyW of all folds.

The end for //No. folds

10. Make the decision to form three subsets (Mostsig, Sig, and Neg) of features using the following rules:

If FinW < a then

F ϵ Neg

Elseif a <FinW ≤ θ1 then

F ϵ Sig

Else

F ϵ Mostsig

Endif

, where a is predetermined by user

11. .End

=================================

The Fuzzy-Relief algorithm shows the details. Initially, determining the number of folds (f) in cross-validation. Then, all features weights are set to 0. Another splitting of data is implemented based on user-defined (m). Later, iteratively the object is randomly selected. For this object find the k nearest neighbor for each class, the update of the weights is executed depending on the match or non-match of the class (step 6).

The fuzzy system is applied to divide the features into three groups (Mostsig, Sig, and Neg) according to the significant degree of merit weight and using the membership function. Thus, the third group will be neglected because of its weights less than the threshold value (a) (step 10).

LR and MLP are applied for speaker identification. The LR depends on maximizing the likelihood of logistic function with ridge value=1.0e^{-8}. While MLP is applied with two layers of input neurons= number of features, hidden neurons= number of features+2 (trial and error) and output neurons=20. As well as the learning value=0.1, and the momentum=0.2.

Last stage, the evaluation is implemented by computing many measures with 3-cross-validation such as: Accuracy, True Positive (TP) rate, False Positive (FP), rate, Precision, Recall, and F-Measure.

5. EXPERIMENTAL RESULTS AND DISCUSSION

The system has been applied to twenty speakers, ten females and ten males, ten recordings and two different sentences for each speaker, five recordings for each sentence and in normal circumstances of the room. The registration is done at sampling rate 44.1KH-16bit.

After silence removal and the applying normalization, the signals are treated with two scopes as image and acoustic files.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5.1 Feature Selection Using the Proposed-Fuzzy Relief Algorithm

The Feature extraction is implemented by db1 of DWT at three levels. The results are 32 coefficients for each acoustic file and 265 features of the image. Fuzzy-Relief divides the features into three groups Mostsig, Sig and Neg resulted from feature selection process.
Table .1 shows the number of features selected in each scope of groups, where a=0.

Table 1: Number of features selected

	all features	Mostsig	Sig	Neg
acoustic	32	12	7	13
image	256	54	9	193

These features are handled in three independent cases of the recognition process.
In the first case, all features are considered.
In the second case, the Mostsig and the Sig groups are merged together.
In the third case, only the Mostsig group is introduced. In the case2 and case3, the Neg. group is neglected.

The reduction percentage= [(number of all features - the number of selected features)/number of all features]*100.

Accordingly, the reduction percentage of the acoustic features reaches more than 62% while the reduction ratio of the image features reaches more than 78%.
Later, two models (LR and MLP) with 3-cross-validation are applied for speaker recognition in the three cases.

5.2 Performance of Acoustic Files

The accuracy rate of the three cases is shown in figures 2,3,4.

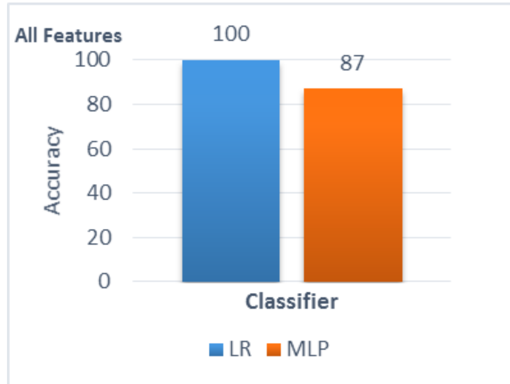

Figure 2. The accuracy of acoustic files using all features

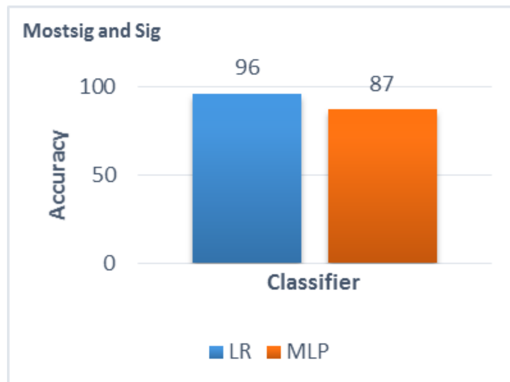

Figure 3: Accuracy of acoustic files using Mostsig and Sig features.

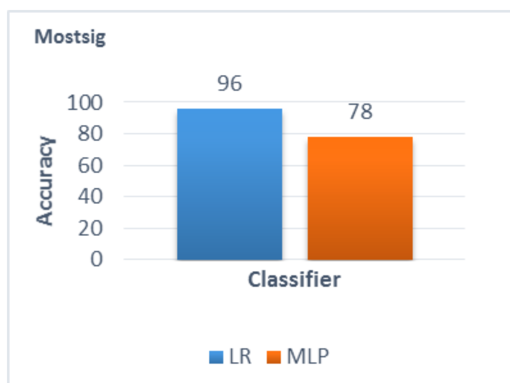

Figure 4: Accuracy of acoustic files using Mostsig features.

From the above figures 2,3,4, note that the LR has recorded better results. So that the accuracy rate is 100 for subset1(all features), 96 % for subset2(Mostsig +Sig features) at reduction rate equals to 40.6% and same accuracy value for subset3(Mostsig) features at reduction rate equals to 62.5%.While the results for MLP are 87%,87,and 78% for subset1,subset2 and subset3 respectively.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The 78% is the least accuracy achieved by the system.

Table 2. shows the others measures of speaker recognition, in general, the results are good but according to all measures, the results of the LR method are better. So that all measures values are perfect of all features and excellent for (Mostsig+Sig) and Sig.While for MLP the worst values occurred in the subset3.

Table 2: Confusion matrix measures of acoustic files

Features types	prediction model	TP Rate	FP Rate	Precision	Recall	F-measure
All	**LR**	**1**	**1**	**1**	**1**	**1**
	MLP	0.87	0.014	0.881	0.87	0.87
Mostsig and Sig	**LR**	**0.96**	**0.004**	**0.962**	**0.96**	**0.959**
	MLP	0.87	0.014	0.881	0.87	0.87
Mostsig	**LR**	**0.96**	**0.004**	**0.962**	**0.96**	**0.959**
	MLP	0.78	0.024	0.798	0.78	0.78

5.3 Performance of Visual Files

On the other side, taking the image for the audio signal is a new approach and the following figures state the accuracy results.

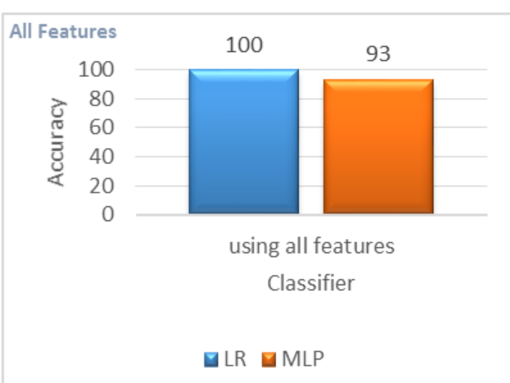

Figure 5: Accuracy of image files using all features

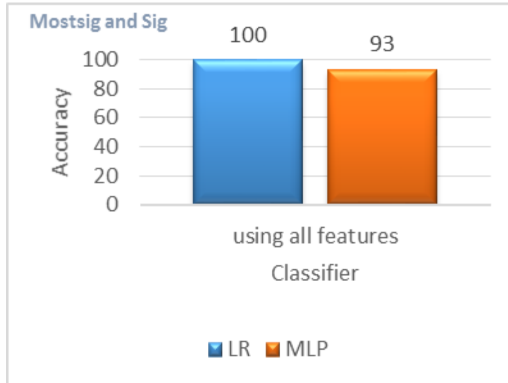

Figure 6: Accuracy of acoustic files using Mostsig and Sig features.

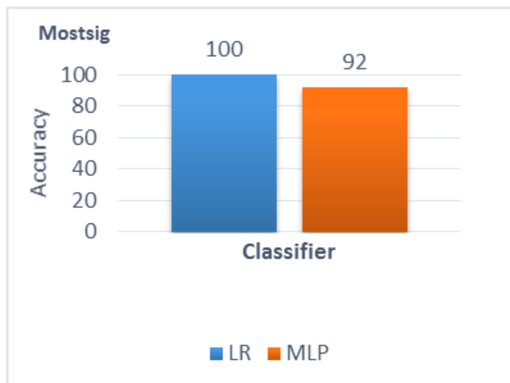

Figure 7: Accuracy of image files using Mostsig features.

From the above Figures 5,6,7, note that the LR has recorded better results, Also, the accuracy rate is 100% for LR, although the reduction percentage of the coefficients is greater than 78%. Also, the MLP has high accuracy which ranged from 92% to 93%.

In a similar way, the confusion matrix measures are calculated for image files as shown in the table.3, the results of LR are perfect in all cases. Thus all values are 1 which represents optimal value. While the results of MLP for all features are identical to the subset2 which contains Mostsig and Sig features, with reduction equals to 75.3%. Analysis the Mostsig features only satisfies close rates to the other subsets with reduction rate equals to 78.9.

Table 3: Confusion matrix measures of image files

Features types	prediction model	TP Rate	FP Rate	Precision	Recall	F-measure
All	**LR**	**1**	**1**	**1**	**1**	**1**
	MLP	0.93	0.008	0.942	0.93	0.93
Mostsig and Sig	**LR**	**1**	**1**	**1**	**1**	**1**
	MLP	0.93	0.008	0.942	0.93	0.93
Mostsig	**LR**	**1**	**1**	**1**	**1**	**1**
	MLP	0.92	0.009	0.929	0.92	0.917

Subsequently, the results are excellent but, according to all measures, the LR method results are perfect values with all subsets of features.

5.4 A Comparative Study

According to previous research studies, in the analysis of the signal or image using DWT, the decomposition level is predetermined only by the required resolution. It is not possible to balance a limited number of levels while maintaining the resolution. In this research, the appropriate level is determined and the number of coefficients is reduced at the same level using the proposed Fuzzy-Relief algorithm.

Table 4: Reduction percentage of Fuzzy-Relief

	Mostsig and Sig	Mostsig
image	**75.3%**	**78.9%**
acoustic	40.6%	62.5%

As shown in Table 4, the reduction value of the image file is higher than the sound reduction rate. In the worst case of the image, it is possible to reduce more than three quarters of the coefficients while about 40% of the audio files.

Another new aspect to compare with previous literature is the results of the visual file for this search, where all the results are better than the research method, which represents the analysis of the voice signal.

So that, the results have recorded ideal accuracy reached 100% of LR in all cases and the worst result is 92% of MLP.

6. CONCLUSION

The first contribution of this research is to capture an image of the voice signal of the speaker and thus treat it as a bmp image file in addition to the audio signal file which represents the traditional method in all previous research for automatic speaker recognition. After normalization for each file before saving.

The DWT is applied to three levels for the analysis of the audio and visual signal. This tool is suitable for dealing with sound and image as well as it works with the time domain and frequency domain at the same time. In the DWT decomposition, increasing the number of levels, the characteristics of signal or image are lost. While the low number of the levels increases the number of the coefficients generated. Three levels have been taken to ensure that's are not lost.

Hence, in order to achieve a balance for selecting an appropriate level of analysis and reducing the number of coefficients by selecting only the important features that positively influence the process of recognition, a Fuzzy-relief is proposed as a new method of feature selection. It is based on applying the fuzzy inference using a new membership function for determining the importance degree for coefficients which are distributed within three groups; the third group is neglected. The proposed method proved its success by reducing the number of coefficients reached 75% as well as maintaining an excellent recognition accuracy.

Two machine learning techniques LR and MLP neural network are applied to the recognition stage. In general for all experiments with all evaluation measures (accuracy, TP rate, FP rate, precision, recall, and F-measure) the performance of LR is better than MLP.

Finally, it is worth noting that the performance of the proposed system on the visual signal is better than the voice signal.

The research has proved that the speaker's voices have a visual print that can enhance the accuracy of identity identification.

REFERENCES

[1] Singh N, Agrawal A, Khan RA, ”Automatic Speaker Recognition: Current Approaches and Progress in Last Six Decades”, Global J Enterprise Inf System, 9, pp.38-45, 2017

[2] Singh N, Agrawal A, Chandra S, and Khan RA, “A Framework for Speaker Recognition System”, Journal of Biostatistics and Biometric Applications,3(1),2018.

[3] Speaker Biometrics Committee, ”Speaker Identification and Verification Applications”, VoiceXML Forum, February 2006.

[4] Sandouk U, ”Speaker Recognition Speaker Diarization and Identification”.,University Manchester School Compute Science, England, 2012.

[5] Shaver CD, Acken JM, ”The development of text-independent speaker recognition”, technology. J Electr Eng, pp.1-8, 2017.

[6] Singh N, Khan RA,” Extraction of Prosodic Features for Speaker Recognition”, Technology and Voice Spectrum Analysis. Int J Sci Eng Res, 5(600-5), 2014.

[7] Huda Naji, ”Visible Dual Watermarking using Wavelet Transform”, Journal of the University of Babylon, 24(9),pp.2374-2382,2016.

[8] Sharada V Chougule and Mahesh S. Chavan,” Speaker Recognition in Mismatch Conditions: A Feature Level Approach”, I.J. Image, Graphics, and Signal Processing,4, 37-43, 2017.

[9] Mangesh S. Deshpande and Raghunath S. Holambe,” Speaker Identification Using Admissible WaveletPacket Based Decomposition, World Academy of Science, Engineering and Technology, 37, 2010.

[10] Mohammed Algabri, Hassan Mathkour, Mohamed A. Bencherif, Mansour Alsulaiman, and Mohamed A. Mekhtiche,” Automatic Speaker Recognition for Mobile Forensic Applications”, Mobile Information Systems, Hindawi, March 2017.

[11] Utpal Bhattacharjee And Kshirod Sarmah, "SPEAKER VERIFICATION USING ACOUSTIC AND PROSODIC FEATURES " Advanced Computing: An International Journal (ACIJ), Vol.4, No.1, January 2013.

[12] Suma Swamy And K.V Ramakrishnan," An Efficient Speech Recognition System", Computer Science & Engineering: An International Journal (CSEIJ), Vol. 3, No. 4, August 2013.

[13] Liu, H., and Yu, L.,” Toward integrating feature selection algorithms for classification and clustering”. IEEE TKDE, 17,pp.491–502,2005.

[14] Bangsheng Sui, ”Information Gain Feature Selection Based On Feature Interactions”, A Thesis Presented to the Faculty of the Department of Computer Science, University of Houston,2013.

[15] E. Chandra Blessie and E. Karthikeyan, ”Sigmis: A Feature Selection Algorithm Using Correlation Based Method”, Journal of Algorithms & Computational Technology, (3),2012.

[16] Ryan J. Urbanowicza,_, Melissa Meekerb, William LaCavaa, Randal S. Olsona, Jason H.Mooreaa, ”Relief-Based Feature Selection: Introduction and Review “, Institute for Biomedical Informatics USA, April 4, 2018.

[17] Eman Al-Shamery and Ali Rahoomi Al-Obaidi,” Disease Prediction Improvement Based on Modified Rough Set and Most Common Decision Tree”, Medwell journals, Journal of Engineering and Applied Sciences,13(5), pp. 4609-4615, 2018.

[18] Kira, K., Rendell, L.,” A practical approach to feature selection”, In 9th Int. Conf. on Machine Learning, Aberdeen, Scotland, Morgan Kaufmann ,pp.249–256, 1992.

[19] Kononenko, I.,” Estimating attributes: Analysis and extensions of relief”. In: European Conf. on Machine Learning, Vienna, Springer Verlag ,pp.171–182, 1994.

[20] Zhou, Q., Ding, J., Ning, Y., Luo, L., Li, T., “Stable feature selection with ensembles of multirelieff”, In Natural Computation (ICNC), 10th International Conference on. IEEE, pp. 742–747, 2014.

[21] Zeng, X., Wang, Q., Zhang, C., Cai, H.. Feature selection based on relief and PCA for underwater sound classification. In: Computer Science and Network Technology (ICCSNT), 3rd International Conference on. IEEE, pp. 442–445, 2014.

[22] Ke-Lin Du and M. N. S. Swamy, ”Neural Networks and Statistical Learning”, Springer-Verlag London 2014.

[23] Faisal Maqbool Zahid & Gerhard Tutz,” Ridge Estimation for Multinomial Logit Models with Symmetric Side Constraints”, Elsevier, 23, 2009.

[24] Nabeel Al - A'araji, Eman Al-Shamery, and Alyaa Abdul- Hussein,” An Optimal Stream Prediction using Adaptive Regression Neural Network “, Medwell journals, Journal of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Engineering and Applied Sciences,12(10), pp. 8844-8850,2017.

[25] R.H.Bhesdadiya, Indrajit N. Trivedi and Pradeep Jangir, ”Training Multilayer Perceptrons in Neural Network using Interior Search Algorithm”, Advances in Intelligent Systems and Computing”, Springer, CCCCS,2016.

[26] Eman Al-Shamery and Ameer Al-Haq Al-Shamery, “An Optimized Feed Forward Neural Network for Reducing Error Based Stock Market Prediction”, Medwell journals, Journal of Engineering and Applied Sciences,13(5), pp. 4616-4621,2018.

[27] Eman Al-Shamery and Ameer Al-Haq Al-Shamery, ”A New Deep Neural Network Regression Predictor Based Stock Market “, Medwell journals, Journal of Engineering and Applied Sciences,13(5),pp.4794-4801,2018.