Open Access
Full Length Article

ORIGINAL ARTICLE

Sound Recognition Aimed towards Hearing Impaired Individuals in Urban Environment using Ensemble Methods

Syed Adnan Shah[1], Ahsan Malik[2], Sumair Aziz[3], Wakeel Ahmad[4]

[1,4]Department of Computer Science, University of Engineering and Technology Taxila, Pakistan.
[2,3]Department of Electronic Engineering, University of Engineering and Technology Taxila, Pakistan.

ABSTRACT

Automated content-based sound classification is an emerging research avenue with applications in sound analysis, surveillance, noise source identification, multimedia retrieval, smart homes/cities, and urban informatics. Traditionally, hearing aid users have been manually changing the instrument settings according to prevailing acoustic conditions. This change activates the appropriate frequency response, compression parameters, noise-cancellation parameters etc. that best fit the situation. Automatic sensing and classification of current acoustic conditions and subsequent automatic switching can relieve the deaf aid user from the annoying task of recognizing the acoustic environment and manual switching. This paper deals with the urban sound classification using an ensemble method. The dataset used for this research is an urban sound dataset containing 27 hours of audio recording with 10 sound classes. For sound classification, individual classifiers verses ensemble methods are applied to the urban sound dataset. According to results, ensemble methods are proved more efficient and robust due to higher recognition rates. The results obtained can be beneficial for designing hearing aids with automatic switching based on automatic sound classification in urban conditions.

Keywords: Acoustic Feature Extraction, Machine learning, classification, ensemble methods, urban sound.

Author`s Contribution
[1] Manuscript writing, Data analysis, interpretation and Active participation in data collection
[2] Conception, synthesis, planning of research, Interpretation, and discussion

Address of Correspondence
Syed Adnan Shah
Email: syed.adnan@uettaxila.edu.pk

Article info.
Received: Nov 07, 2017
Accepted: June 11, 2018
Published: June 30, 2018

Cite this article: *Shah SA, Malik A, Aziz S, Ahmad w. Sound Recognition Aimed towards Hearing Impaired Individuals in Urban Environment using Ensemble Methods. J. Inf. commun. technol. robot. appl.2018; 9(1):30-39.*

Funding Source: *Nil*
Conflict of Interest: *Nil*

INTRODUCTION

The Human ear can hear a sound which has frequency components ranging from 20 Hz to 20 KHz. A person having 25db hearing threshold cannot hear the sound of normal range frequencies might facing the issue of hearing disorder. Hearing loss can affect one and sometimes both ears which may cause difficulty in hearing conversations, speech, and loud sounds. Hearing disorder can be of different levels: i) mild, ii) moderate, iii) severe, or iv) profound. According to the World Health Organization (WHO) Factsheet 2015, 360 million people are suffering from disabling hearing disorder worldwide [1]. Disabling hearing loss with more than 40 (dB) in adults with better hearing and more than 30 (dB) in children with better hearing ear refers to hearing loss. A

big ratio of disabling hearing loss facing people lives in low-medium level income countries [2].

Individuals suffering from disabling hearing loss are mostly over 65 years of age group. The ratio of this age is greater in South Asia, Asia Pacific, and sub-Saharan Africa. Individuals facing hearing disorder can get benefit from hearing aids, cochlear implants, and other assistive devices. Microphone, amplifier, processor and a loudspeaker complete hearing aid. Loudspeakers are used to transmit the enhanced sound into the ear. Modern devices are designed according to the shape of the human ear along with the degree of hearing disorder. Modern aids are efficient enough to do more than just enhance the sound as the small processor is used to optimizing the sound. Modern devices are comfortable, invisible and small in size with the ability to speech recognition and hearing ease.

In the literature, general approaches have been reported for urban sound recognition systems that normally follow the following three steps: i) acquisition of sound, ii) extraction of features, and iii) sound classification. Researchers mostly use datasets of daily life sounds with the utilization of various sensors used in sound classification. With the objective to separate noise from sounds, a binary masking based algorithm has been developed [3]. The mask developed to achieve this target was evaluated by training the proposed algorithm on speech which is then not used in the testing phase. Speech shaped noise (SSN) and chatters were mixed with the original sentences at various signal-to-noise ratios (SNRs). Testing phase has been done by utilizing the individuals with normal hearing and hearing impaired (HI) listeners which results that intelligibility enhanced processing in all situations.

The researchers have been applied ensemble techniques for the classification of sounds. Ensemble methods are thought as learning algorithms that create a set of classifiers and by taking a vote of their predictions they used to classify new data points [4]. The goal of this research is to classify sounds using ensemble methods [5] and hence distinguishing important and unwanted urban sounds to help make a better understanding of the surrounding environment.

LITERATURE REVIEW

Several recent studies have been reported in the literature on sound impairment using urban sound datasets. S. Ali et al [6] presented their research work pertaining to human heart sounds classification using ensemble techniques. Authors applied the proposed framework for publicly available standard heart sound dataset and identified set of audio features which were used for human heart sounds classification. This study claims that the proposed technique proves to be more effective and robust as it increases the overall classification accuracy, and the results are higher than the existing solutions. Tuomas Virtanen et al [7] examine the engineering challenges in the computational environmental sound analysis. In this study, the authors compare environmental sound analysis to other major audio content analysis fields and identify the challenges in the field.

Søren Laugesen et al [8] discussed the significances of adding speech-like properties to ASSR stimulus. The authors were concerned to ensure that the hearing aid processes the stimulus as if it were real speech. The study concludes that reduced response magnitude and increased detection time seem acceptable, given the potential for allowing aided ASSR recordings to be carried out with hearing aids in their daily-life mode of operation. Klaylat, Samira, et al [9] proposed a two-phase novel model to enhance an emotion recognition system. This system claims recognition of three emotions, happy, angry, and surprised by a realistic Arabic speech corpus. This study compares thirty-five classification models with a newly developed model which gave the best result with 98% accuracy.

Lesica, Nicholas A. [10] and Irvine, Dexter RF [11] presented a brief review titled as “Why Hearing Aids Fail to Restore Normal Auditory Perception and “Auditory perceptual learning and changes in the conceptualization of auditory cortex” respectively. In these articles, the authors summarize the questions and the proposed solution presented till date on the said topic. Pierre Roy et al [4] addressed the issue of identification of acoustic features and recognition models. Purpose of this research was to solve problems raised in designing and implementing FDAI. The main two issues in FDAI are completeness as the difficulty level is high to describe all sounds in an urban environment and consistency as all

the sounds are not consistent. GMM and nearest neighbor classifiers were used for experiments. The datasets used were related to vehicles, birds and music sounds. They were succeeded to overcome the problem of completeness, but the issue of
Consistency remained questionable.

A.J. Torija et al [12] examined the utilization of sound spectra that is recorded as input data for the improvement of concurrent and immediate road traffic assessment models. They implemented a sequence of models related to the utilization of neural networks, multilayer perceptron, the Fisher linear discriminant and multiple linear regression for the guesstimate of road traffic, in addition, to categorize it according to the structure of motorcycles/mopeds and heavy vehicles. As a result, researchers found the highest correlation value with 50-400 Hz and 1-2.5 kHz frequency range and reported that the use of recorded sound spectra attains positive results. The relationship between measured frequency spectra and road traffic flow had been studied. Initially, highest correlated bands the third-octave bands with road traffic intensity are acknowledged and after that greatest correlation frequency bands with road traffic were used as latent input for a model for the estimation of general road traffic.

Arslan Shaukat et al [13] demonstrated the applicability of ensemble methods for automatic daily sound recognition. For this purpose, two datasets were used: i) RWCP database and ii) daily sound dataset. Individual base classifiers were applied on both datasets. The accuracy rate of applied classifiers was lower than stated in the literature. Subsequently, ensemble methods were applied on both datasets for classification purpose. Kernel discrimination analysis and Hierarchal Hidden Markov Models (HHMMs) were applied on the RWCP dataset and a combination of GMM and SVM were applied on the sound data set. They also performed 10-fold cross-validation on both datasets. In total, a combination of nine ensemble methods was applied to both datasets. Performance of ensemble methods was much higher than base classifiers.

David Opitz et al [14] applied Bagging and boosting (two commonly used ensemble methods) on 23 datasets to come to a decision that which ensemble method is more efficient when used for decision trees and neural networks algorithms. Research suggested that bagging is accurate than single classifier but less accurate than boosting but sometimes boosting seems to be less accurate than single classifiers especially when applied on neural networks. Experiments showed some

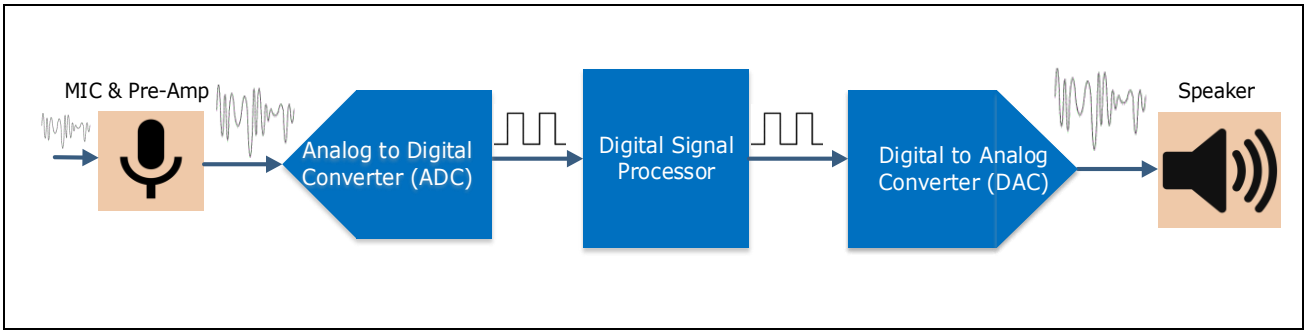


Figure 1: Overview of the Digital Hearing Aid System

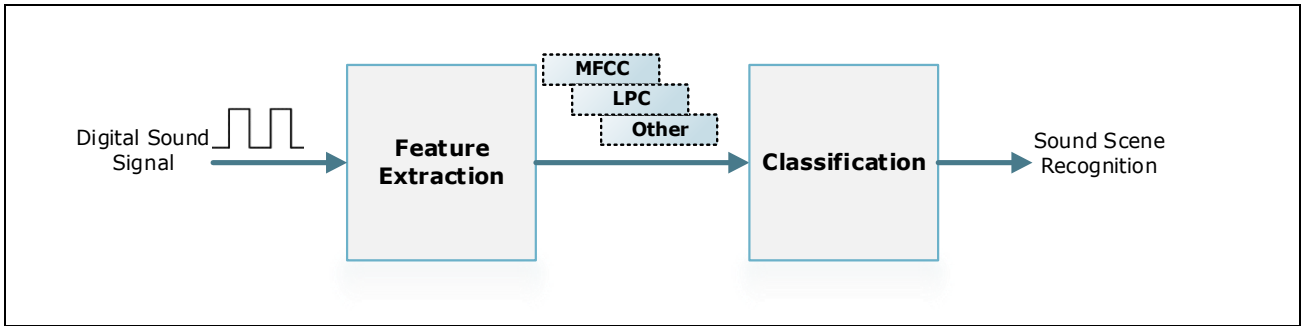


Figure 2: Block Diagram of Sound Scene Recognition System

relationship exists between boosting and dataset being used therefore Boosting over fits for noisy datasets. Thomas G. et al [5] reviewed the original ensemble methods and explained the reason why these methods are more efficient than single classifiers. This paper also compares the performance of different ensemble methods and concludes that AdaBoost mostly gives efficient results than others. Bagging and randomization have comparable performance. While using randomization gives better results than bagging when applied on large datasets.

To identify urban sounds Ensemble methods are successfully proved to be operative while compared with the individual classifiers. This study follows the aim to make a comparison between the proposed setup and literature stated results. The organization of the paper is as follows. Section III describes the experimental setup and methodology in details. Section IV reports implementation results, analysis, and brief comparison. The conclusion has been represented in the last section.

System Overview and Our Contribution

Audio received by the microphone is amplified and then Analog to digital (ADC) converter converts input audio signal to digital. A microprocessor or Digital signal processor analyzes the received signal and produces an amplified speech signal. Speech amplification is performed according to the environment in which the user is present for instance quiet room, loud streets etc. Finally, this digital signal is converted into an analog signal using a digital to analog (DAC) module. This output is attached to the miniature loudspeaker after necessary amplification for producing sound.

The core factor which contributes towards the efficient design of the DHA system is the environmental sound recognition ability in which user is present. This requires an efficient framework for sound recognition. Based on this recognition of the environment, an algorithm running in microprocessor changes the amplification level of the output sound produced by the load speaker. The main contribution of our work is aimed towards enhancing the classification accuracy for sound scene recognition.

EXPERIMENTAL METHODOLOGY

The details of the dataset, acoustic features, classification methods and ensemble techniques which are used in our experimental setup will be discussed in this section. Along with the proposed setup, previous methods from literature also enlightened in this section.

A. Urban Sound Recognition Dataset.

We have employed a standard dataset which has been created particularly with the purpose of urban sound recognition and has also been used by other researchers. The details of the dataset are mentioned here.

i. Urban Sound Database

The Urban Sound 8K dataset is created by the researcher. This dataset contains 8732 labeled sound excerpts (<=4s) of urban sounds from 10 low-level classes: air conditioner, car horn, children playing, dog bark, drilling, engine idling, gunshot, jackhammer, siren, and street music. Classes are drawn from the urban sound taxonomy. This is a standard dataset targeting 4 top-level groups: human, nature, mechanical and music in terms of sounds and recording conditions to meet the requirements for training scalable algorithms with the ability to analyze actual data through sensor networks or multimedia repositories [13]

Each class contains 1000 sound files in WAV format each class sounds have been transformed to 16bit for symmetric file characteristics before feature extraction phase. Table I. presents a number of samples and categories of each class. Samples used from the database have a maximum duration of 0.5 sec.

TABLE I. Urban Sound Dataset

Sound class	No. of files
Air Conditioner	1000
Car Horn	429
Children Playing	1000
Dog Bark	999
Drilling	1000
Engine Idling	1000
Gun Shot	374
Jackhammer	1000
Siren	929
Street Music	1001

B. Acoustic Features

For achieving the higher accuracy in sound classification phase effective feature extraction is very

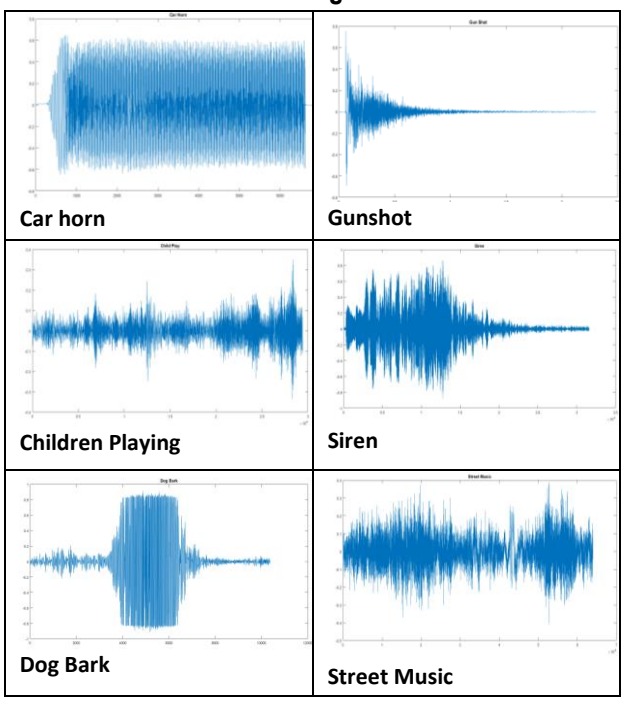
important. Features are represented in numerical values and decrease the information that a sound wave contains. Extraction of different acoustic features for sound classification has done in this literature. For sound classification, acoustic features have been explored in detail and a list of features useful for recognizing urban sounds have been identified and used in our work.

Table II Presents the acoustic features used in the proposed system. First 13 coefficients of Mel-frequency Cepstral Coefficients (MFCC's) [15], the first 10 coefficients of Linear Predictive Coding (LPC's) [16] and four Non-Spectral features are extracted. Root mean square of frames, a fraction of low energy window, relative difference function and zero crossing completes non-spectral features. Statistical measures like Standard deviation and mean are used to extract the features.54 total features are utilized in this study to achieve high accuracy.

TABLE II. Acoustic Features	
Non-Spectral	Root Mean Square of Frames
	The fraction of Low Energy Windows
	Relative Difference Function
	Zero Crossings
MFCC	First 13 coefficients
LPC	First 10 coefficients

Figure 3: Waveforms of the urban sound dataset

C. Classification using Ensemble Methods



In literature, individual classifiers have not performed well as compared to the ensemble methods because before final decision they consider the results of individual sub-classifier and also able to recover if any classifier fails to perform well. Bagging, Random Subspace [17] and AdaBoostM1 [18]ensemble methods are utilized in this study as they perform well with weak classifiers. For daily sound recognition, no such combinations of ensemble methods are used till today [19] Base classifiers have also been used individually to evaluate the effectiveness of ensemble methods [12].

1) Bagging

The name bagging by Breiman in 1996 comes from 'bootstrap aggregation'. It is an effective and most simple technique of ensemble learning method used for arching. The special case of model averaging is meta-algorithm which was designed for the classification often applied to the models like decision tree [13] but may also use for all type of model working for regression or classification. Using bootstrap, the method works for different versions of the training set like sampling with replacement. Each data set from them is used for the training of the different model. In case of regression the output of the models are merging by the averaging and in classification, case voting is performed creating a single output. The effectiveness of voting is high only when a small change of data set may cause a significant change of model means using as unstable for nonlinear models. For a training set of N, a learning models' class like neural networks or decision trees etc.

2) Methodology

By training, multiple models (m) on different splits of data means different samples. Now averaging the predictions; by averaging, the output of m models do predict the results. The objective is to improve the accuracy and efficiency of the model using its different copies another goal is that on different data pieces the average misclassification errors give a better estimation of predictive ability for the learning technique.

Training: for each iteration i=1 to me, with random sample S from the training data set, training the 'base model' like decision tree or neural network etc. on the given sample. Test: starting all of the base models (trained) for every test example predict by merging results of all M models. The output of bagging can be expressed

as

$$c^*(x) = \arg_{\max y \varepsilon Y} \sum_{i:c_i(x)=y} 1$$

3) Random Subspace

The Random subspace [17] also called attribute bagging a well-known ensemble method that contains multiple classifiers. Each individual classifier operates in a different subspace of base feature space and the output is based on the individual classifier's output. Decision trees, linear classifiers, Support vector machine, nearest neighbor and one class classifiers are some basic classifiers for which Random Subspace is used. Random subspace is the best choice for Classification problems in which numbers of features are greater than training objects, for example, fMRI (Functional Magnetic Resonance Imaging) data.

The Random subspace [17] also called attribute bagging a well-known ensemble method that contains multiple classifiers. Each individual classifier operates in a different subspace of base feature space and the output is based on the individual classifier's output. Decision trees, linear classifiers, Support vector machine, nearest neighbor and one class classifiers are some basic classifiers for which Random Subspace is used. Random subspace is the best choice for Classification problems in which numbers of features are greater than training objects, for example, fMRI (Functional Magnetic Resonance Imaging) data.

$$\beta(x) \arg_{\max y \varepsilon [-1,1]} \sum_{h} \delta_{\text{sgn}}(c^b(x)), y$$

Let:

T: number of training objects, F: number of features, C: number of individual classifiers. For each individual classifier c, choose fc (fc < F) to be the number of input variables for c.

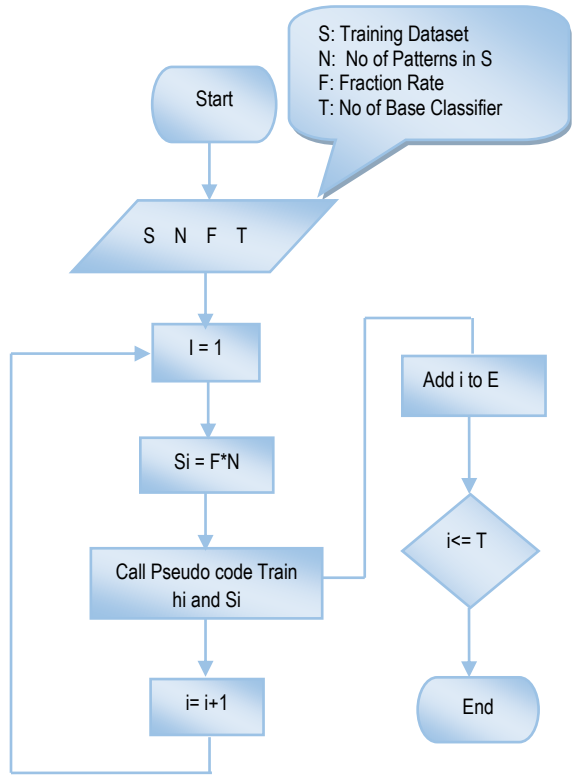


Figure. 4: Bagging

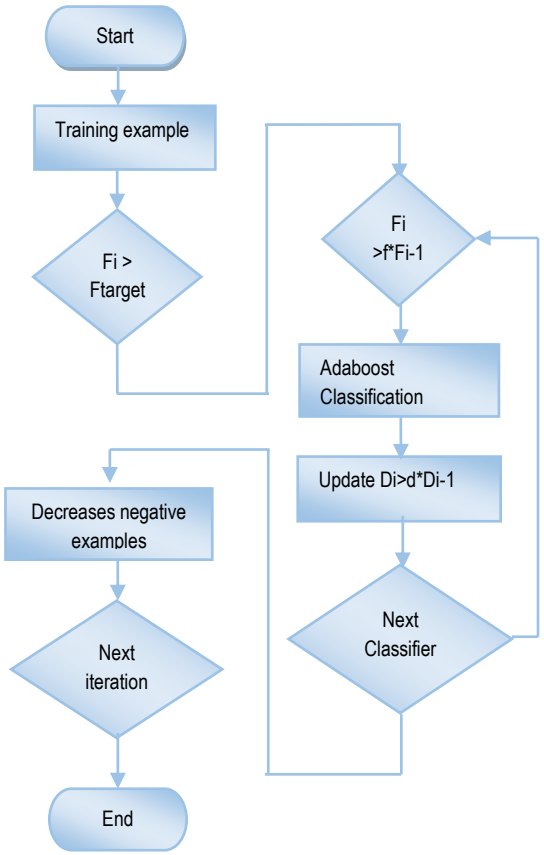


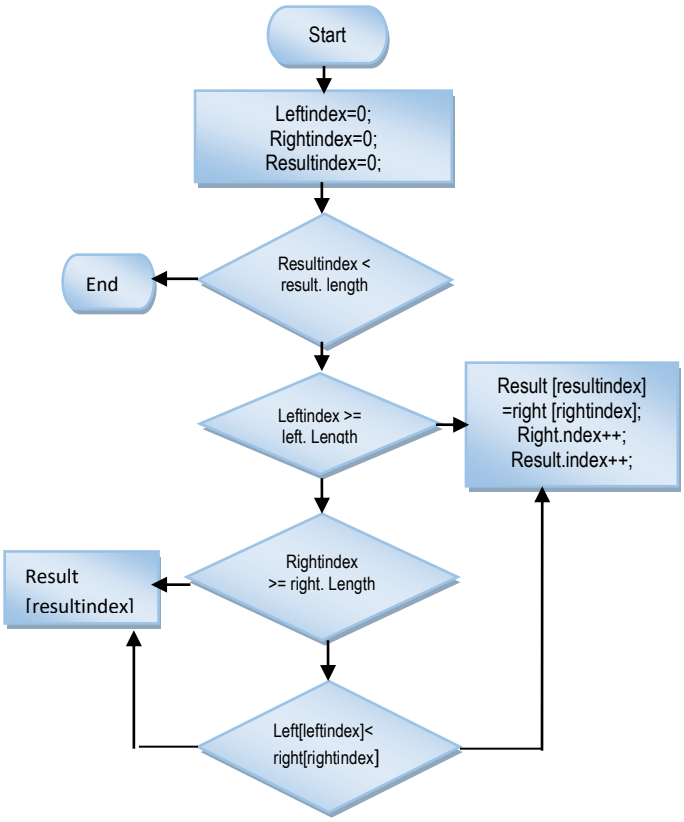
Figure. 5: AdaboostM1

Choose fc features from F, create a training set for each classifier c and train classifier. For classification of the new object, combine the results of L through majority voting. With ensemble methods, Random Forest, C4.5 Decision Tree [20] and Bayesian Network are used as base classifiers the reason been taken these base classifiers is less time taken by them to build a model with high accuracy with said ensemble methods [5].

4) Random forest

The tree predictors combination is known as Random Forests [21]. In this combination, the independent sampled value of the random vector is being dependent by each tree and the same distribution is applied to all the trees of the forest. For each node having error rates to be split by using a random selection of features that can display the comparison with AdaBoost [18] but with respect to noise, these are more robust. Strength, correlation and monitor error for internal estimates Are used to show a response to increasing the number of features used in splitting. These internal estimates are also important for the measurement of variable importance.

$$c = 1/B \sum_{b=1}^{B} Cb(x')$$



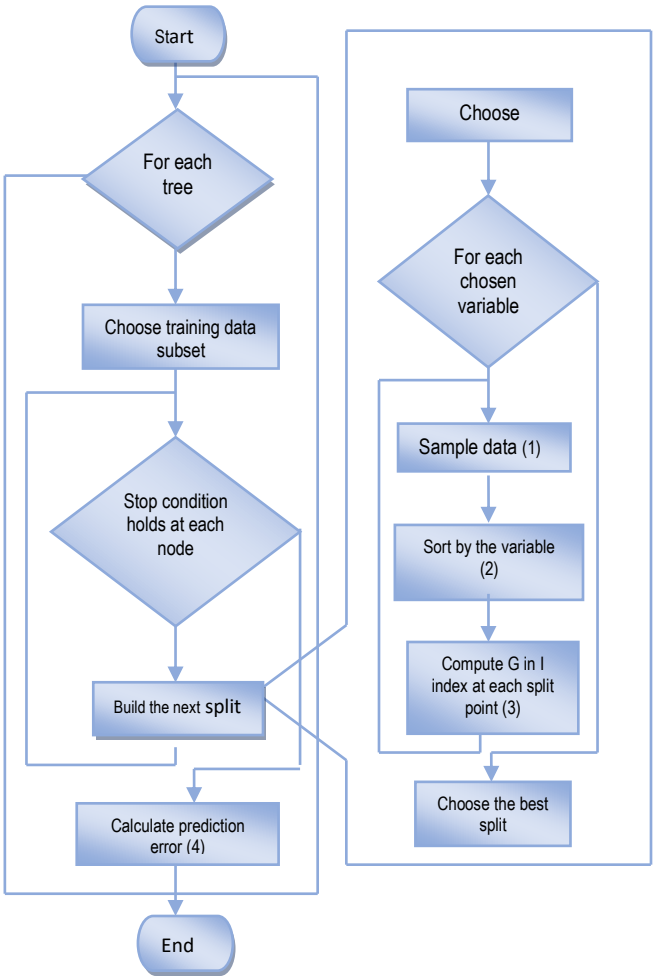


Figure 7: Random Forest

5) C4.5 Decision Tree

The classifier C4.5 is an extension of the ID3 algorithm. Using training data it develops a decision tree. The main improvement in C4.5 in comparison of ID3 is supporting the continuous and discrete attributes both beside this supporting the training data with a missing value of the attribute. It can also handle the attributes with differing prune trees and cost as well after creation.

$$1(P) = -(P_1 * \log(P_1) + P_2 * \log(P_2) + \ldots\ldots + P_n * \log(P_n))$$

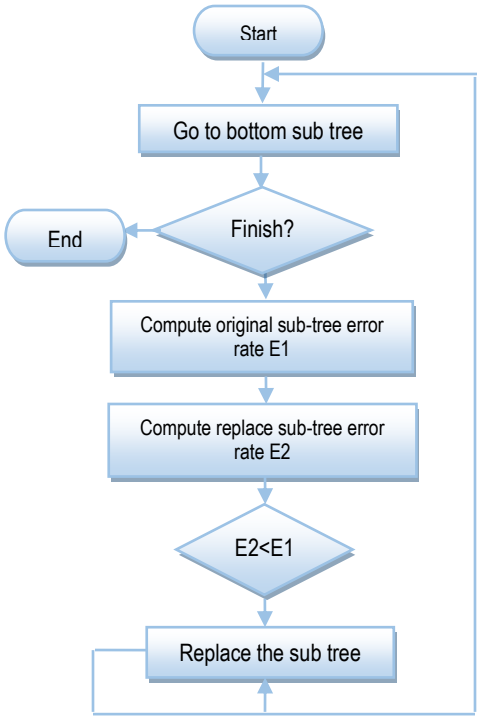


Figure 8: C4.5

6) Bayesian Network

$$P_B(x_1 \ldots\ldots x_n) = \prod_{i=1}^{n} P_b(x_i \mid \prod x_i) = \prod_{i=1}^{n} \theta x_i \mid \prod x_i$$

$$class(x) = \arg_{\max c \varepsilon Y} \sum_k P(M_k \mid S).P'M_k(Y = c_i \mid x)$$

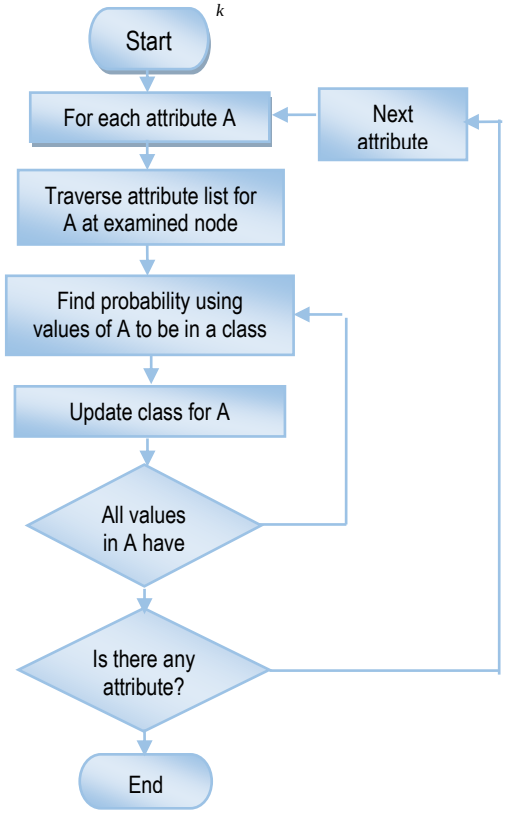


Figure 9: Bayesian Network

Experimental settings are the same as used in past studies to compare the results. Selina Et al [22] used the combination of GMM (Gaussian Mixture Model) and SVM (Support Vector Machine) with the objective of recognition of distress situation. 16 MFCC (Mel-Frequency Cepstral Coefficients) features extracted were extracted over a total of 18 sound classes of the sound data set. 75% accuracy rate was achieved using this method.

Ye et al [23] study present environmental sound recognition by using Kernel Discriminant Analysis and Gabor Transform was employed for future extraction. 105 sound classes of RWCP databases were evaluated with 10-fold cross validation which results in 94.41% accuracy. Peng et al [24] research the classification of sounds on the RWCP database using HHMMs (Hierarchical Hidden Markov Models). 12 MFCCs [25] and log frame energy along with 1st and 2nd derivatives were extracted. Three iterations of 50%-50% training-testing were applied over 105 sound classes that achieve 87.57% accuracy.

For the RWCP dataset, 10 fold cross-validation is used for classification in this study. 50%-50% training-testing approach is also used for the validation of classification which partitions the dataset into 50% training and 50% testing randomly. 10ms window size with 50% overlap is designed for the extraction of 54 features for sound samples. The window size is adopted from literature for a comparison point of view.

For sound dataset, classes' data have been split into three equal sections. section i) denotes for training section ii) has been used for testing and section iii) is reserved for development objective. Total of 704 instances is used in which 351 have been used in testing while remaining 353 are used in the testing phase. 54 features have been extracted utilizing 16ms window with 50% overlap for each sample. 10-fold cross validation is also performed for classification hence no such results were evaluated in literature.

At classification level three base classifiers with three ensemble methods were used which makes 9 combinations and their performance is evaluated on both datasets.

RESULTS & DISCUSSION

In this section, results of our experiments on the Urban sound dataset using acoustic features are presented. Performance of three individual classifiers is evaluated, afterward, the ensemble method based classification technique is applied to evaluate the performance of algorithms as compared to the individual classifiers.

The Sound recognition performance of individual classifiers i.e. Random Forest, Bayesian Network and C4.5 decision tree has been evaluated. Using all acoustic features discussed above and applying 10-fold validation schemes the result is shown in classification accuracies along with Standard Deviation in table III.

TABLE III. RECOGNITION RATES (%) USING 10 FOLD CROSS VALIDATION ON THE URBAN DATASET

Classifiers	Recognition Rates (Mean & Standard Deviation) %
Random Forest	81.87 (1.23)
Bayesian Network	65.87 (1.39)
C4.5 Decision Tree	71.38 (1.32)
Salamon et al.'s Method	70

This result shows that individual classifiers have the same ratio of classification accuracy and not improved result from literature. Therefore, an enhanced classification technique is required in order to improve sound recognition accuracy.

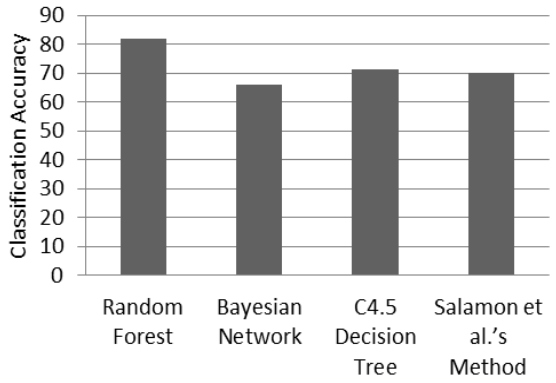


Table v shows the performance of nine combinations of ensemble methods using 10-fold cross-validation scheme.

TABLE IV. Recognition Rates (%) Using 10 Fold Cross Validation on the URBAN dataset

Ensemble Methods with Base Learners	Recognition Rates (Mean & Standard Deviation) %
Random Subspace-Random Forest	87.85 (0.32)
Random Subspace-Bayes Net	65.32 (0.70)
Random Subspace-C4.5	84.88 (0.43)
AdaboostM1-Random Forest	**88.25 (0.31)**
AdaboostM1-Bayes Net	70.04 (0.77)
AdaboostM1-C4.5	86.31 (0.39)
Bagging-Random Forest	86.26 (0.45)
Bagging-Bayes Net	67.14 (0.55)
Bagging-C4.5	82.15 (0.81)
Salamon et al.'s Method	70

The sound recognition rate of the proposed framework is higher as compared to the previous results. Our experiments demonstrate that the use of Ensemble methods classifiers for sound recognition of extracted acoustic features yield better output in terms of accuracy, but the process takes a bit longer time for training and testing which is acceptable in our scenario due to the availability of high performance embedded processors at a very low price that is required for implementing a complete system.

Confusion matrix of applying Random Subspace-Random Forest using 50%-50% training testing approach is shown in Table VI. Some classes such as hand clapping, hair dryer, and keys have 100% recognition rate other classes also show higher accuracy rate like 97.36% for door clapping and for an electric shaver, it shows 95.23% accuracy.

Table V-VI elaborate the results of applying Random Subspace-Random forest combination methods in 50%-50% and 10-fold cross validation mechanisms.

Results show the consistency and reliability of the proposed framework with a higher accuracy rate of recognition of daily sound as compared to literature. (Higher accuracy rates are highlighted in bold in each table). Figure 5 describes the results of the proposed method in terms of recognition rates.

CONCLUSION

This paper dealt with the issue of detecting daily domestic sounds using various ensemble methods. The findings can be used for hearing aids, surveillance, and smart homes etc. To achieve the goal, initially individual classifiers along with various acoustic features were used. Subsequently, ensemble methods were used for the classification task. The accuracy rates of ensemble methods in recognition of daily sounds were higher than individual classifiers as well as from individual three base algorithms in the literature. Hence, the classification accuracy attained for given two datasets using the proposed method of ensemble methods is much higher than previously reported results in the literature.

Future Recommendations

To achieve higher accuracy, it is suggested that more features should be evaluated and added in the future study as well as feature selection will also an interesting part. For future studies, our results on two datasets of sounds using ensemble methods will serve as baseline performance.

REFERENCES

1. Organization, W.H., Childhood hearing loss: Strategies for prevention and care. 2016.
2. Rochon, P.A., and J.H. Gurwitz Optimising drug treatment for elderly people: the prescribing cascade. BMJ: British Medical Journal, 1997. 315(7115): p. 1096.
3. Roman, N., D. Wang, and G.J. Brown, Speech segregation based on sound localization. The Journal of the Acoustical Society of America, 2003. 114(4): p. 2236-2252.
4. Defréville, B., et al. Automatic recognition of urban sound sources. in Audio Engineering Society Convention 120. 2006. Audio Engineering Society.
5. Dietterich, T.G. Ensemble methods in machine learning. in International workshop on multiple classifier systems. 2000. Springer.
6. Ali, S., et al., Human Heart Sounds Classification using Ensemble Methods. University of Engineering and Technology Taxila. Technical Journal, 2017. 22(1): p. 113.
7. Homsi, M.N., and P. Warrick, Ensemble methods with outliers for phonocardiogram classification. Physiological measurement, 2017. 38(8): p. 1631.
8. Laugesen, S., et al. On the cost of introducing speech-like properties to a stimulus for auditory steady-state response measurements. in Proceedings of the International Symposium on Auditory and Audiological Research. 2018.
9. Klaylat, S., et al., Enhancement of an Arabic Speech Emotion Recognition System. International Journal of Applied Engineering Research, 2018. 13(5): p. 2380-2389.
10. Lesica, N.A., Why Do Hearing Aids Fail to Restore Normal Auditory Perception? Trends in neurosciences, 2018.
11. Irvine, D.R., Auditory perceptual learning and changes in the conceptualization of auditory cortex. Hearing research, 2018.
12. Torija, A.J. and D.P. Ruiz, Using recorded sound spectra profile as input data for real-time short-term urban road-traffic-flow estimation. The science of the Total Environment, 2012. 435: p. 270-279.
13. Shaukat, A., et al. Daily sound recognition for elderly people using ensemble methods. in Fuzzy Systems and Knowledge Discovery (FSKD), 2014 11th International Conference on. 2014. IEEE.
14. Opitz, D. and R. Maclin, Popular ensemble methods: An empirical study. Journal of Artificial Intelligence Research, 1999. 11: p. 169-198.
15. Beritelli, F. and R. Grasso. A pattern recognition system for environmental sound classification based on MFCCs and neural networks. in Signal Processing and Communication Systems, 2008. ICSPCS 2008. 2nd International Conference on. 2008. IEEE.
16. Bradbury, J., Linear predictive coding. Mc G. Hill, 2000.
17. Ho, T.K., The random subspace method for constructing decision forests. IEEE transactions on pattern analysis and machine intelligence, 1998. 20(8): p. 832-844.
18. Collins, M., R.E. Schapire, and Y. Singer, Logistic regression, AdaBoost and Bregman distances. Machine Learning, 2002. 48(1): p. 253-285.
19. Lu, H., et al. SoundSense: scalable sound sensing for people-centric applications on mobile phones. in Proceedings of the 7th international conference on Mobile systems, applications, and services. 2009. ACM.
20. Blockeel, H., Top - down induction of first-order logical decision trees. Ai Communications, 1999. 12(1 - 2): p. 119-120.
21. Liaw, A. and M. Wiener, Classification and regression by randomForest. R news, 2002. 2(3): p. 18-22.
22. Chu, S., S. Narayanan, and C.-C.J. Kuo, Environmental sound recognition with time-frequency audio features. IEEE Transactions on Audio, Speech, and Language Processing, 2009. 17(6): p. 1142-1158.
23. Ye, J., et al. Kernel discriminant analysis for environmental sound recognition based on acoustic subspace. in Acoustics, Speech, and Signal Processing (ICASSP), 2013 IEEE International Conference on. 2013. IEEE.
24. Peng, Y.-T., et al. Healthcare audio event classification using hidden markov models and hierarchical hidden Markov models. in Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on. 2009. IEEE.
25. Ganchev, T., N. Fakotakis, and G. Kokkinakis. Comparative evaluation of various MFCC implementations on the speaker verification task. in Proceedings of the SPECOM. 2005.