iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** XI **Month of publication:** November 2017

DOI: http://doi.org/10.22214/ijraset.2017.11112

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue X1, November 2017- Available at www.ijraset.com

Vehicle Acoustic Signal Classification Using Machine Learning Algorithms

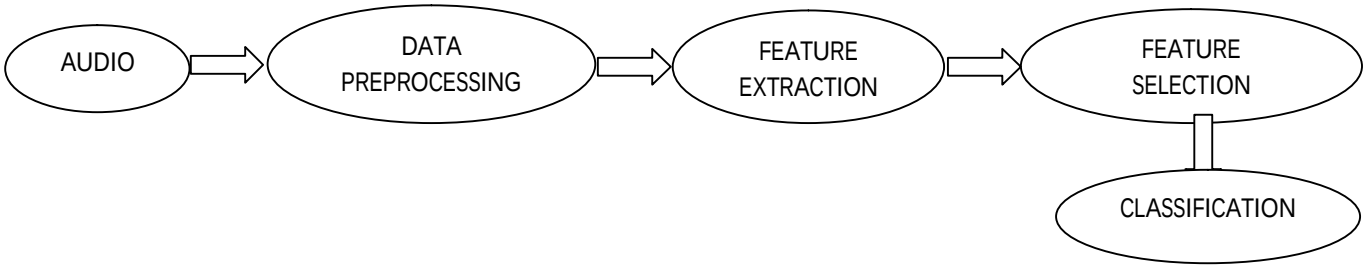
Alekhya Kodali[1]
[1]*Department of Computer Science, Maulana Azad National Institute of Technology Bhopal*

Abstract: ***Acoustic signals are vibrations in air produced by the movement of non-living objects in our daily life. These can be due to vehicles, turbines, machines and so many different objects. Classification can be done using Machine Learning techniques where we train the machine with some training data whose labels are known and get the class or label of test data whose class is not known. There are many pre-defined classification algorithms like KNN, SVM, RANDOM FORESTS, etc. In this project we are using KNN, Decision tree, NAÏVE BAYES for classification purpose. As Acoustic signal is a vast topic, here our application of acoustic signals is in vehicle classification. By using microphone we collect large amount of train data and classify vehicles depending on that data. We discuss further in the introduction part.***
Keywords: vehicle classification, feature extraction, KNN, decision tree, Naïve bayes.

I. INTRODUCTION

According to the new researches, Noise pollution problems have been increased to a large extent. Reducing Noise pollution completely is impossible, but it is possible to spread the noise evenly in the surroundings so that concentration at a given area can be reduced. To achieve this, in this report we are developing an algorithm to identify the vehicle making much noise in a given road.
Our project mainly comprises of the following three steps:

A. Collection of required testing data.
B. Preprocessing the collected data i.e. Feature extraction.
C. This extracted data is stored in a csv (excel) file.
D. Then classification algorithms are applied on the train data to train the machine.
E. As Machine is now trained we can test our data to see which class our data belongs to [2]



All classification systems employ the extraction of a set of features from the input signal. Each of these features represents an element of the feature vector in the feature space. The dimension of the feature space is equal to the number of extracted features. These features are given to a classifier that employs certain rules to assign a class to the incoming vector.

II. DATA COLLECTION AND ANALYSIS

A vehicle can be detected using microphones. Recordings from microphone will be taken and are run in the classification algorithms for vehicle detection.
The recordings in microphone will be used as training data. This data collection is done with the help of a microphone which is fixed on road side and records the vehicle sound passing through that road. Vehicle acoustic signal were collected from different locations by the side of the road. Recordings were done at the straight stretches of far away from the intersections /bus stops with minimum background noise. Data can be collected through many ways like by Magnetic Sensing or Inductive loops etc. But all these require high maintenance cost and installations cost. When compared to these methods Acoustic signal collection through microphone takes less installation and maintenance cost [6]. Collection of data can vary depending on their necessity i.e. data might be in the form of .wav file, .mp3 file or any other audio files but here we take .wav files

Based on their type vehicles are classified into heavy, medium and light vehicles. The heavy vehicles comprise of truck, bus and the medium vehicles are car, van whereas the light vehicles consist of bikes, autos and other two wheeler vehicles. In our project we are going to take heavy vehicles as truck, medium vehicle as car and light vehicle as bike and going to distinguish these with the help of KNN, decision tree and Naïve Bayes machine learning algorithms.

III. DATA PREPROCESSING

Collected audio files in the above method are divided into small windows of time interval 10ms and used for feature extraction. These windows are called as Analysis Windows.

ANALYSIS WINDOWS: Analysis window is a short window obtained by breaking signal into small segments of time and each segment can be processed for feature extractions[5]. In this project, we have taken frame time as 0.01 and sampling rate 40,000. From this we get frame length 400(sampling rate*frame length). As shown in figure below

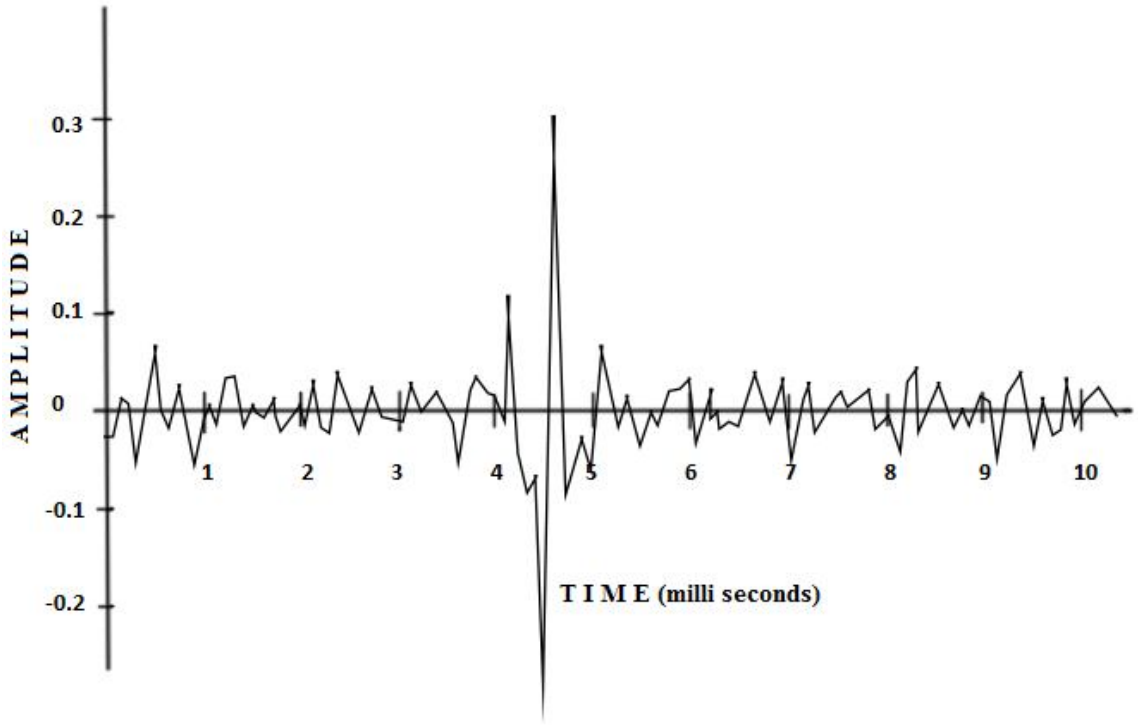


Fig. 1 Plot of analysis window

IV. FEATURE EXTRACTION

Features are the attributes that are used to describe a living or non-living object. As our aim is vehicle classification, features used to classify vehicles are of four types. They are temporal features, spectral features, energy and perceptual features. In our project we used mainly spectral features as they helped to distinguish the data in an accurate manner.

In order to find features suitable for classification, first of all consider feature sets containing single feature and check the distances between different classes in datasets. If a feature set is able to distinguish distinctly then we assign a priority one, then consider another feature sets of length two where one of the feature should be the one whose priority is given highest. Repeat the above process by increasing the length of feature sets gradually till we get the desired length. We have to take the features that are suitable for acoustic signals because number of features in the feature set should be reliable, keep the features that are able to classify the audio signals accurately and remove the features that are not able to classify because using unnecessary features makes the algorithm complex and increases time of execution. So it is important to choose the relevant number of features for classification to reduce the cost and complexity of programs [3]. Features are selected based on their discriminative power and uncorrelated to other features. Discriminative power of a feature is the ability to discriminate different classes thoroughly and uncorrelated to other features is that we see two different features give different values for the same audio signal so that they can distinguish signals properly without redundancies.

The following feature set is used in our project- spectral centroid [1], spectral roll-off, spectral zero crossing, spectral flux, low energy, spectral flatness and RMS

V. METHODS USED FOR CLASSIFICATION

Classification is a supervised learning technique which consists of two steps- learning step and classification step. In learning step a model is constructed based on the given training data. Training data is a data for whose class labels are known. Using this training data we train the machine so that it can classify the testing elements. In the classification step, based on the model constructed above we predict the categorical class labels of the testing data. There are classification algorithms based on ANN (artificial neural networks) such as back propagation algorithm, maximum likelihood method and multi-layer feed forward neural network and various other pre-defined machine learning algorithms[8]. In our project we used KNN (K- nearest neighbour), Decision tree, Naïve Bayes classification algorithms.

VI. EXPERIMENTAL SETUP AND RESULT ANALYSIS

With the K-NN classification we mainly want to classify a vehicle among bike, car and truck. The instances of training data taken are around 50-60 for each vehicle in order to train the machine. There are two methods to extract the features. One method is we can directly use the mathematical formulas for these features and get the values that could distinguish the data but here we are not sure whether our formula application is 100% correct or not. The second method is that in MATLAB we have a toolbox named as MIR toolbox which was designed to process and analyze the signals. Using MIR toolbox we can easily extract the features of the audio file [4].

After training the machine with bike, truck and car data it forms classes depending on the labels given

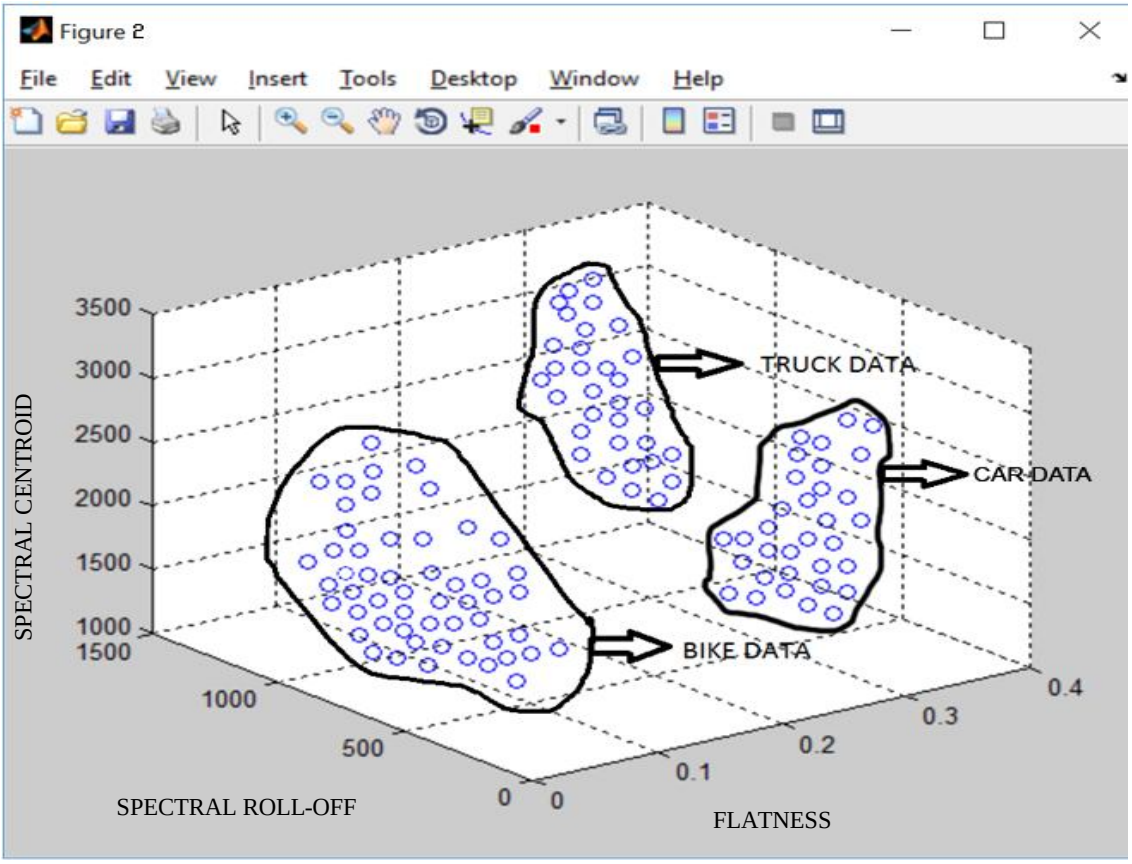


Fig. 1 Classification of Training Data

Results of accuracy of the given classifiers

DATA	KNN	DECISION TREE	NAÏVE BAYES
Data 1	78.57	66.42	64.1
Data 2	79.34	65.7	63.8

VII. CONCLUSION

In this paper a model for classification of vehicle acoustic signal has been proposed. This system has been able to detect the type of vehicles that causes the maximum noise pollution in a particular road. In this the preprocessing of data is done in order to eliminate the horn sound of vehicles and various other noises. Feature extraction is performed so as to differentiate various vehicles labeled as classes and suitable features are selected among them. Later, the classification algorithms are applied on the train data and then the test data is taken and results are evaluated. Through analysis and survey among the classification algorithms (KNN, Decision tree, Naïve Bayes) used it turned out that KNN is best among the other algorithms as it has got better accuracy and less error rate. So for vehicle classification one can employ KNN method and this is what we found through this project. With this project we are able to classify vehicles and use this classification for different purposes. The above proposed work is Data Classification. This work can be extended by making a model for Data Acquisition i.e. sensors like microphone or any other audio observing sensors and connecting them to our Data Classification can complete this model.

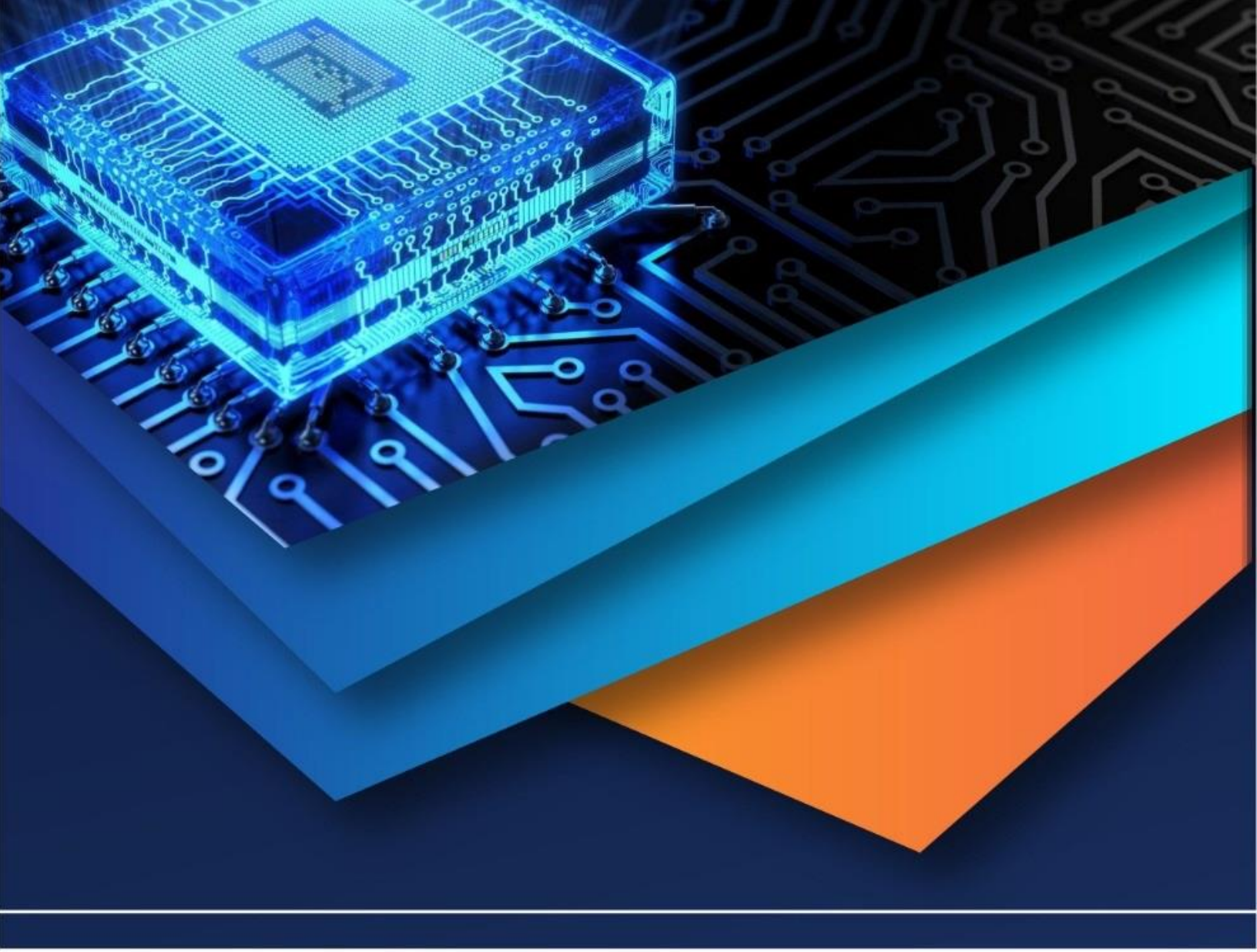
This model real time usage is very helpful to reduce Environment damage problems.

VIII. ACKNOWLEDGEMENT

I would like to thank Prof. Praveen Kaushik, Department of Computer Science Engineering, Maulana Azad National Institute of Technology, Bhopal, Madhya Pradesh, India for his valuable suggestions and clarifying doubts regarding the topic Vehicle Acoustic Signal Classification.

REFERENCES

[1] George Tzanetakis, Perry Cook, JULY 2002, Musical Genre Classification of Audio Signals
[2] Naoko Evans Department of Electronics, University of York, September, 2010, Automated Vehicle Detection and Classification using Acoustic Signals
[3] Jeroen Breebaart and Martin McKinney, Features for Audio Classification
[4] A. Y. Nooralahiyan' AND H. R. Kirby 1998, Vehicle Classification by Acoustic Signature
[5] Hariharan Subramanian, Dept, IIT Bombay, Submitted November2004, Audio Signal Classification
[6] Jobin George, Leena Mary, Riyas K S, 2013 International Conference on Control Communication and Computing (ICCC) Vehicle Detection and Classification
[7] From Acoustic Signal Using ANN and KNN
[8] Shuilin Guo, Rui Wang*, Bin Liu, Qiyue Wei, Yanping Li, 2016 IEEE, Vehicle Classification in Acoustic Sensor Networks based on Hybrid Dictionary Learning
[9] Yao He, Yuchuan Du, Lijun Sun, 8th International Conference on Traffic and Transportation Studies, 2012 Elsevier, Vehicle Classification Method Based on Single-Point Sensor
[10] Amol Ambardekar, Mircea Nicolescu, George Bebis, Monica Nicolescu, 2014 Springer, Vehicle Classification Framework: A comparative Study
[11] S.Gupte, O.Masoud, R.F.K Martin, N.P. Papanikolopoulos , 2002 IEEE Explore, Detection and Classification of Vehicles



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY