Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2022) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

Machine learning applied to acoustic-based road traffic monitoring

Karolina Marciniuk[a,*], Bożena Kostek[b]

[a] *Multimedia Systems Department, Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdańsk, Poland*

[b] *Audio Acoustics Laboratory, Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdańsk, Poland*

Abstract

The motivation behind this study lies in adapting acoustic noise monitoring systems for road traffic monitoring for driver's safety. Such a system should recognize a vehicle type and weather-related pavement conditions based on the audio level measurement. The study presents the effectiveness of the selected machine learning algorithms in acoustic-based road traffic monitoring. Bases of the operation of the acoustic road traffic detector are briefly described. Principles of several machine learning algorithms, data acquisition process, and information about the dataset built are explained. The study is conducted using the audio recordings prepared by the authors, registered in several locations and under different meteorological conditions of the road surface. For each recording containing a single-vehicle passage, a vector of 67 parameters extracted from the audio signal is calculated. Fisher Linear Discriminant Analysis and Regression Analysis, the fastest among algorithms employed, return the following values of accuracy: 0.968 and 0.978, precision: 0.919 and 0.853, recall: 0.882 and 0.974, and F1-score: 0.898 and 0.868 for vehicle type classification. In the case of the road pavement conditions, the obtained metrics are as follows: accuracy of 0.933, precision of 0.898, recall of 0.9, and F1-score of 0.884.

Keywords: vehicle acoustics; traffic road analysis; machine learning;

1. Introduction

An essential part of the environment in which human lives is sound. Sound provides a lot of important information about the environment. It plays a role in communication, expression, control, etc. It allows detecting a potentially critical threat before it is visible. It is therefore understandable to try to transfer the ability to listen and analyze to computing devices. The terms: "machine hearing" (and also Computer audition (CA); Computer/machine listening or Machine Hearing (MH)); "auditory scene analysis" (Auditory Scene Analysis (ACA); Computational Auditory Scene

* Corresponding author. Tel.: +48-58-347-19-67
E-mail address: karmarci@pg.edu.pl

Analysis (CASA)) emerged in the 1990s [24], but only recently - due to the development of methods, algorithms, and information technology - practical solutions and systems have become possible and feasible [30]. Within the "Machine Hearing" group, several subgroups of acoustic environmental analysis systems can be distinguished [29, 34]. Sound analysis is developed for, among others, autonomous robots, mobile devices [25, 26], assisting people with disabilities (e.g., sound scene analysis for people who are blind or visually impaired, hard-of-hearing people, etc.), and public safety considerations (e.g., systems for detecting gunshots, explosions, and other sounds in public spaces and environments) [27]. Acoustic CASA systems also find application in pipeline monitoring [28], the protection of which has both economic and safety dimensions.

Even though road noise is considered environmental pollution, it is possible to collect valuable information from the noise signal caused by vehicle traffic. The first document regulating the process of developing the road network in the context of road noise is the document from 1970, referred to as NEPA (The National Environmental Policy Act, Pub.L. 91–190). In contrast, the first project on acoustic road monitoring was developed in the 20s of the last century by Charles Adler, Jr. Acoustic road traffic monitoring is a passive technique, and unlike video cameras, the phenomenon of occlusion has a more negligible impact on the operation of the system. The most significant limitation of acoustic detection is the low durability of the sound sensor (microphone). However, the current development of mems microphones has resulted in a return to research on sound in traffic surveillance systems. During the last two decades, a few sound-based traffic detection systems were under the research program. They took place in various parts of the world [1, 22, 23]. Some of the projects considered using acoustic-only in the detection of possible dangerous events like the sounds of "crashers," "horns," and "lost cargo." Acoustic monitoring in tunnels [3, 16] or strategic places in cities may be given as an example. Besides, a few research solutions on acoustic vehicle detection and classification were also tested [11].

Detection and classification of traffic vehicles can be performed on the basis of several properties and physical phenomena that accompany either the movement or the presence of a vehicle. There are several groups of detectors that can be distinguished based on the method of operation (microwave, acoustic, optical methods) or on the method of installation (over the road, on the pavement). It is the location and the nature of traffic in a given location that dictates the target technology for vehicle detection and classification. Different detectors will be more effective for lower-speed built-up areas and others for high-speed roads.

Acoustic traffic analysis bases its operation on sounds accompanying a moving vehicle. The primary sound source that can be hearable for speeds higher than 30 km/h is caused by the contact of the tire with the road surface [13]. For lower speeds, the primary noise component is the engine and exhaust system [2]. The acoustic fingerprint of an individual vehicle pass-by is composed of several factors, such as the type of vehicle, its speed, the condition of the surface, and the technical characteristics of the vehicle (engine type, vehicle age).

Depending on the complexity of the traffic detector or the location of the measurement, it is advisable to obtain information about the number of vehicles in the time intervals, mean or individual speed, and traffic structure. For the purpose of this paper, the authors focused on investigating acoustic vehicle classification methods. Therefore, the overview of acoustic-based road traffic monitoring enabled to formulate the assumptions of this work. Since the applicability aspect of this research is important, i.e., road event detection in near real-time, that is why we propose to employ baseline machine learning algorithms such as regression analysis (RA), Linear Discriminant Analysis (LDA), Fisher Linear Discriminant Analysis (FLDA), and Naive Bayes Classification (NBC). As these algorithms are broadly used, thus only a brief description will be given. This is followed by depicting the experiment flow, dataset preparation, training, validation, and testing in the context of the vehicle body and road condition classification. Finally, conclusions are depicted, and future research is outlined.

2. Machine learning algorithms used in the evaluation of audio road events

Automatic Sound Recognition is one of the top interests in the Machine Learning (ML) trend. The growth of such an area is connected to massive technological progress in the field of computer hardware that allows for the development of ML methods, including deep learning [33]. For the purpose of this study, several supervised learning algorithms were tested in a road sound-based recognition system. Some of these algorithms are also used as prediction methods. The most important advantage of these particular ML algorithms is a short time of training. Also, they can operate on a limited segment of the population to be recognized - in this case, a limited number of representations occurring in the signal acquired from traffic in combination with a large variety of vehicles on the market.

2.1. Perceptron learning algorithm

The classical perceptron model was developed in 1958 by Frank Rosenblatt [17]. In 1969 the concept of the smallest human brain operational unit was proposed by Minsky and Papert [12]. Each of the inputs is the information that is weighed by the importance value. The typical use of this simplest ML algorithm is binary classification. For the purpose of the multiclass classification problem, the perceptron was used as a one-layer neural network.

2.2. Regression Analysis

Regression analysis (RA) is typically used for time series data (such as, e.g., stock quotes). During the learning process, it takes into account its forecasts on archival data and predicts on that basis. In road engineering, regression analysis can be found in determining the relationship between the accident rate and driver age. RA is used to recognize object classes; the task of the algorithm is to find the relationship between the characteristics of the object that determine its belonging to a particular class. Thus, parameters that have a significant impact on the correct classification of the object are discerned [6]. Multiple regression can be calculated using the following Equation:

$$Y = b_0 + b_1 x_1 + ... + b_n x_n \tag{1}$$

It is crucial to establish a proper number of objects from the population needed for analysis. According to Green [5], the minimum sample should be at least 50. The number should be increased by 8 for each additional class. As already mentioned, RA is a widely used algorithm in many fields of science, primarily because of its simple implementation and fast training.

2.3. Linear Discriminant Analysis and Fisher Linear Discriminant Analysis

The group of linear methods for classification is based on the assumption that there exists a proper way of space division of data classes using the calculated decision boundaries. In that case, we always have in mind the assumptions of accepted accuracy or precision. In two-classes LDA, there is an assumption that the classes share a common covariance matrix where $\sum_c = \sum \forall c$. The main idea of LDA is the computation of directions of the parameters that established the maximal separation between classes. FLDA is a specific type of LDA. In this case, it is to find a projection to a line with the best possible separation.

The algorithm focuses on the mean and variance of the values of the parameters. The efficiency depends on the differences between classes and convergence within a class [10]. LDA and FLDA are often used for feature reduction.

2.4. Naïve Bayes Classification (NBC)

Naïve Bayes Classification (NBC), like the classical Bayesian classification method, is based on the relationship of conditional probabilities of statistical quantities and can be described by Equation 2. Using a prepared model ready to compute the likelihood for each class, we obtain the generative model. It is a possible random process that describes the process of data formation.

$$P(c|\, x) = \frac{P(x|\, c)P(c)}{P(x)} \tag{2}$$

where:

c – class (label, a group of objects).

$P(c|\, x)$ – the posterior probability of c given predictor x (attribute, feature).

$P(c)$ – the prior probability of class c.

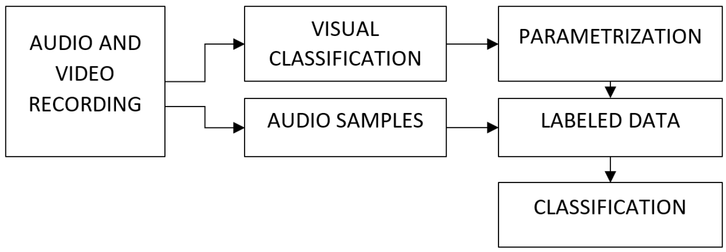

Fig. 1. The workflow of the conducted experiment

$P(x|c)$ – likelihood, the probability of predictor given class c.

$P(x)$ – the prior probability of the predictor.

The most significant advantage of NBC is low computational complexity – for a training stage is $O(np)$ and $O(p)$ for prediction. The second one is the ability to work with more than two classes. The outcome of NBC is strongly connected to the number of objects in classes, but the algorithm does not need too much for training to perform well. The main disadvantage lies in the assumption that predictors are entirely independent, which is nearly impossible in most of the classification data. That is one of the reasons we call it naive [6].

3. Research workflow

The experiment was performed in a few steps (Figure 1). For the purpose of the experiment carried out, an audio-visual dataset was prepared. Recording sessions took place in strategic (i.e., traffic noise context) agglomeration areas. The basis for the operation of the acoustic method is the registration of the signal emitted by the vehicle in motion. Therefore, the measuring stations were located on straight sections of roads, not blocked by traffic light signals. Also, there was no significant gradient in the longitudinal road section, which could increase the contribution of engine components to the vehicle acoustic footprint. Audio recordings were manually tagged based on the video. Thus, for the purpose of correct classification, recordings were made during daylight hours. High-speed open space routes were selected. The pavement condition at the measurement location was good, or very good, to avoid additional sounds in the recording such as storm drain overruns or expansion joints on the bridges.

For safety issues, the size of the recording equipment has been minimized, so the conducted measurements did not distract drivers. The recording devices were placed at the height of 1.5 meters at some distance from the road (1-2 meters). At the same time, the recording was carried out with the use of two microphones with different directional characteristics. Additionally, video documentation (also with sound) was carried out. Recordings were split into individual events (understood as sound samples with separate pass-by). Thanks to the simultaneously recorded video image, it was possible to tag files into a few categories – vehicle type, approximate speed, and pavement state (dry, wet, snow). This constitutes the ground truth of the study performed. Because of uneven vehicle class distribution, the dataset was unbalanced, which can be problematic for some ML algorithms. Another unequal condition was the presence of snow on the road (only one short session occurred).

3.1. Dataset preparation for classification

In this study, only a manually tagged audio dataset was used. The dataset for vehicle recognition was built from more than 400 audio files. At the time of the experiments, no other databases containing accurate tagging were found. Additionally, in most Western countries, the average vehicle age is up to 50% lower than for the area in which the study was conducted. In the process of parametrization, a feature vector (FV) was determined for each file. The full FV contains 67 parameters related to the spectrum distribution and mid-level features associated with the perception of the sound, categorized as bright or dark. The parametrization process was carried out with MIRtoolbox [9].

Recordings were split into individual events (understood as sound samples with individual pass-by of specific car types with no background noise). Video recording was used for tagging: vehicle types (according to the Department of

Table 1. Audio dataset - vehicle type at dry pavement.

Group	Description	No. of raw files	No. of used data
Light vehicles	typical small cars with gasoline engine	213	242
Medium vehicles	medium cars, mostly delivery, suv, bus (diesel engine - four wheelers, single unit)	45	92
Heavy vehicles	truck, buses, trailers, three axle or more	44	83
Motorcycles	all kinds (also with scooters)	78	110

Transportation (DOT), National Highway Traffic Safety Administration (NHTSA) 2011[35]), vehicle speed, pavement meteorological conditions. The dataset was additionally described with supplementary tagging for future experiments.

Before training, the database was cleared of incomplete data. Due to different numeric ranges of descriptors, the data had to be scaled so that each parameter is represented by a close order of values. Followed by that, the data were split into different categories (depending on the analyzed problem). As mentioned before – the dataset was unbalanced, so for better performance, a simple mechanism similar to bootstrapping was performed. The number of newly created objects depended on the class share in the population, but the number of objects in the classes was not equalized. The degree of noise (values in vector) took into account the distribution of individual density values in the population (see Table (1)).

The data prepared were then introduced to the training and test processes using individual algorithms described in the previous Section. The data were split into three groups – training, validation, and test sets. Each of the algorithms was subjected 300 times to the test, each time changing the set of training and validation data in the way of random mixing. The proportions assumed at the beginning of the experiment (45% training, 20% validation, 35% test) were maintained. Measurement was performed as iterated steps.

The process of data classification was performed in GNU Octave using the previously prepared script and toolboxes from the Octave Forge community [31]. The calculations were conducted utilizing i78550U processor with all eight cores (12395 points in Geekbench 3 – multithreaded performance test). These characteristics are important in the context of the practical feasibility of the study.

Information about time consumption is presented in Table 2. Two road traffic parameters were tested - vehicle type and pavement conditions. Each problem was analyzed as a one-label problem, which means a separate test cycle was carried out for each aspect.

Table 2. Time consumption for the whole process: 300 iterations in training and test (528 observations of different vehicle body types with 67 features per each)

Classification method	FLDA	LDA	NBC	PLA	REG
Time [s]	41.52	39.62	6.47	4.28	2.50

At first, we tested the data in the case of vehicle body type recognition. Taking into consideration the number of a class used in dynamic traffic management and the acoustic characteristics group of vehicles, we prepared a dataset with different labels – motorcycle, light vehicle, medium - van, and heavy. The second experiment tested the ability to recognize the pavement condition from the changes in the sound of a vehicle pass-by. The change in the type of surface and grip has an impact on safety.

Using the vectors of the expected classes and predicted classes, a set of classification efficiency metrics was calculated using a predefined script. For this classification problem, five efficiency metrics were chosen: accuracy (ACC), recall (TPR), precision (PPV), specificity (TNR), and F1-score [4].

3.2. Vehicle body recognition

Due to the different operating principles of the algorithms, several data transformation methods were tested at the pre-processing stage. Table 3 shows the best results achieved for the given algorithms (overall effectiveness in four class recognition problems). The summary of results is presented for the test set. The highest absolute difference was

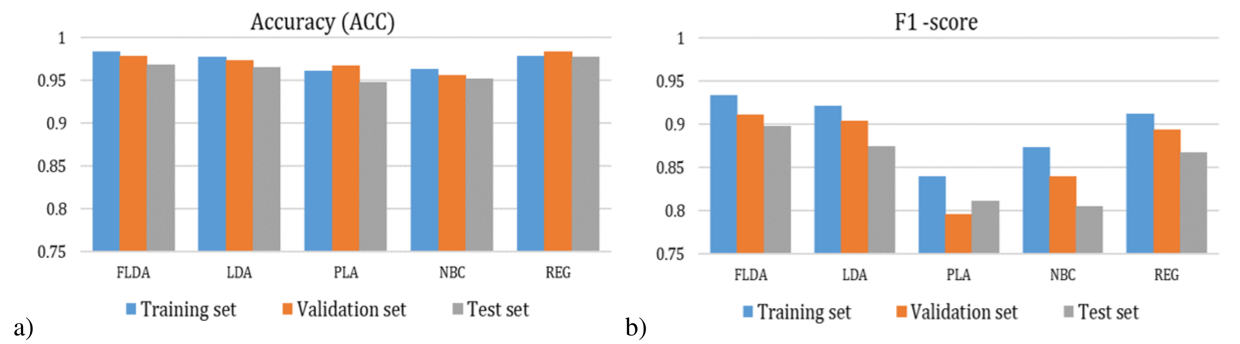

Fig. 2. a) Differences between training, validation, and test performance according to ACC b) Differences between training, validation, and test performance according to F1-score (vehicle type classification – four classes, 45:20:35 data ratio)

found in the PLA method (0.12 for TPR and 0.13 F1-score). In each column, the highest scores are highlighted in bold font to better illustrate the differences between the metric and the assessment algorithm. The FLDA turned out to be the most sensitive algorithm in the case of vehicle sound.

Table 3. The results of the effectiveness of individual algorithms in the test phase

	ACC	TPR	TNR	PPV	F1-score
FLDA	0.968	**0.919**	0.978	0.882	**0.898**
LDA	0.965	0.911	0.979	0.863	0.875
PLA	0.948	0.846	0.966	0.833	0.812
NBC	0.952	0.845	0.971	0.835	0.806
RA	**0.978**	0.853	**0.984**	**0.974**	0.868

For FLDA, PLA, and the RA model, the best results were obtained while using z-score normalization. For the other two classification methods, we use results with min-max normalization [15].
Graphs in Figure 2 show the average effectiveness for classifiers taking into account the best possible pre-processing method. Perceptron was the least effective algorithm in the tested group.

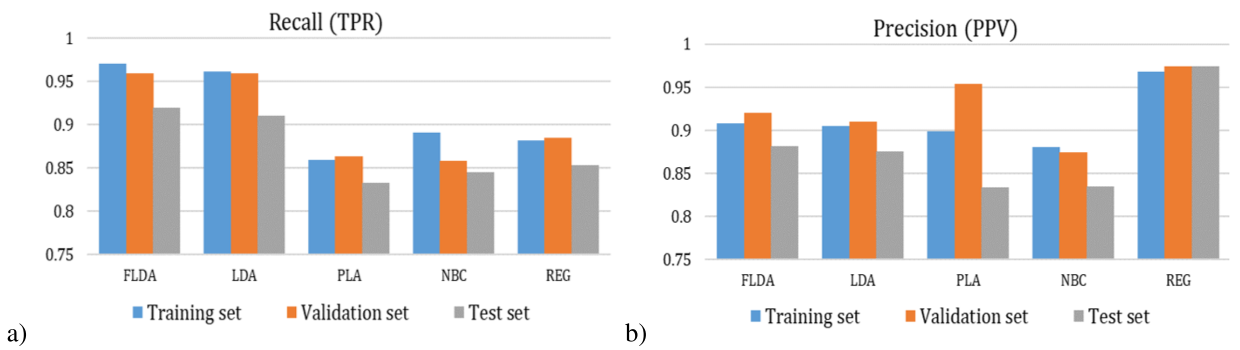

Fig. 3. a) Differences between training, validation, and test performance according to TPR b) Differences between training, validation, and test performance according to PPV (vehicle type classification – four classes, 45:20:35 data ratio)

Table 4 presents the results of the classification of particular classes taking into account the measures of effectiveness assessment described earlier. A motorcycle is the easiest type of vehicle to be recognized by an acoustic-based classifier. The distinction between the class of medium vehicles (van) depends on the visual division of vehicles by an adapted classification system that is based on a body vehicle type. The most considerable distortions of classification are caused by diesel vehicles as well as older vehicles.

Due to the comparison of more than two algorithms, instead of ANOVA or the Friedman test [18], the method chosen for the analysis was the paired parametrical test. Using F1-score and paired Wilcox test, only two pairs of algorithms show no significant difference ($p < 0.05$): LDA and FLDA, and NBC and PLA, rejecting the hypothesis of equal effectiveness of tested algorithms.

Table 4. The efficiency in different body types using FLDA and Regression Analysis, the fastest of the methods tested

FLDA	ACC	TPR	PPV	F1-score	TNR
Light	0.937	0.891	0.95	0.919	0.968
Van	0.955	0.786	0.733	0.759	0.972
Heavy	0.987	1	0.857	0.923	0.986
Motorcycle	0.993	1	0.986	0.993	0.988
Mean	0.968	0.919	0.882	0.898	0.978
REG	ACC	TPR	PPV	F1-score	TNR
Light	0.956	0.968	0.987	1	0.978
Van	0.942	0.545	0.923	1	0.853
Heavy	0.895	1	1.000	1	0.974
Motorcycle	0.946	0.615	0.909	1	0.868
Mean	0.948	0.993	0.993	1	0.984

3.3. Pavement condition recognition

The presence of fluids like water or snow may impact road safety. Even if most drivers tend to adjust vehicle speed to the weather conditions [19], for some inexperienced drivers, additional information about road condtition can be beneficial in terms of safety. This information is directly extracted from the sound from the interaction between the tires and the road captured by the microphones. In classic weather supervision systems, rainfall information is taken from rain sensors mounted at certain altitudes. There are optical-based meters for the measurement of water presence on the surface [14]; however, due to the high cost of a single device, this solution is used, preferably, at airports. For weather detection, the FLDA efficiency metrics are shown in Tables 5 and 6.

The most confusing condition for the algorithms was snow (recording when there was 1-2 cm of solid snow cover). Due to the weather conditions prevailing in the area when the recordings took place, it was the least numerous group of audio samples. To improve the proportion between the objects, only half of the sound base was included in the pavement condition classification. As resulted of the pavement condition classification, the best performance was obtained in the case of the Fisher Linear discriminant analysis with accuracy (ACC) of 93%.

Table 5. Confusion matrix for different pavement conditions (FLDA, FV=67)

	Dry	Wet	Snow
Dry	108	0	10
Wet	0	42	12
Snow	0	0	51

4. Conclusions

Due to the limited possibilities of introducing new road facilities in urban and intercity areas, the maximization of the efficiency of existing transportation networks is crucial for administration. Intelligent transportation systems (ITS)

Table 6. Efficiency metrics for different pavement conditions (FLDA, FV=67)

	Dry	Wet	Snow	Mean
ACC	0.953	0.944	0.901	0.933
Precision	0.915	0.778	1	0.898
Recall	1	1	0.699	0.9
Specificity	0.903	0.93	1	0.944
F1-score	0.956	0.875	0.823	0.884

enable to increase road capabilities and safety thanks to traffic analysis and control, along with communication with the road users. The experiments conducted based on the audio-based vehicle classification and their outcomes show that it is worth pursuing the development of acoustic technology for traffic supervision.

It is possible to acquire information from the road noise that can be useful for highway safety management. Acoustic monitoring is a passive method, and it does not require additional signals, as in the case of microwave and optical detectors. Moreover, the acoustic sensor can cover larger areas of operation and is not susceptible to lighting conditions and obscuring vehicles as other methods tend to do. Besides, time for assigning the output labels for the current sound events is short, so this is important in the context of real-time applications.

Commercially available traffic detectors are quite expensive, and cheaper solutions like inductive loops are very invasive and may damage the pavement. The proposed simple methods as well as the possibility of their implementation on relatively cheap solutions allows for a sighificant densification of the measurement network. In addition, the small size of the device allows for finstallation of existing road infrastructure like light poles.

In the future, several threads will be pursued. First of all, the system will be tested on data that are available as open access to improve the robustness of the developed system. Also, estimating traffic intensity by employing noise road analysis is envisioned. Another goal is to detect vehicles approaching from the side. This will be done by employing a sound intensity approach instead of microphones. This way, a sensor may measure sound intensity in two directions: parallel and perpendicular to the road. Moreover, based on sound intensity probes, incorrect movement direction can easily be detected. This may be especially important in applications such as smart-city traffic light control.

Acknowledgements

This research was funded by the Polish National Centre for Research and Development (NCBR) through the European Regional Development Fund entitled: INFOLIGHT Cloud-Based Lighting System for Smart Cities under Grant POIR.04.01.04-00-075/19.

References

[1] Barbagli, B., Magrini, I., Manes, G., Manes, A., Langer, G. & Bacchi, M. (2010) "A distributed sensor network for real-time acoustic traffic monitoring and early queue detection." *2010 Fourth International Conference On Sensor Technologies And Applications*. pp. 173-178.
[2] Cerrato, G. (2009) "Automotive sound quality–powertrain, road and wind noise." *Sound And Vibration*. **43**, 16-24.
[3] Foggia, P., Petkov, N., Saggese, A., Strisciuglio, N. and Vento, M. (2015) "Audio surveillance of roads: A system for detecting anomalous sounds." *IEEE Transactions On Intelligent Transportation Systems*. **17**, 279-288.
[4] Ghoneim, S. (2019) "Accuracy, Recall, Precision, F1-Score and Specificity, which to optimize on?" *Towards Data Science* (Accessed on April 2022)
[5] Green, S. (1991) "How many subjects does it take to do a regression analysis." *Multivariate Behavioral Research*. **26**, 499-510.
[6] Hastie, T., Tibshirani, R. and Friedman, J. (2003) "The elements of statistical learning, corrected ed." *Springer*.
[7] Kilpeläinen, M. and Summala, H. (2007) "Effects of weather and weather forecasts on driver behaviour." *Transportation Research Part F: Traffic Psychology And Behaviour*. **10**, 288-299.
[8] Klein, L., Mills, M., Gibson, D. & Others Traffic detector handbook: Volume I. (Turner-Fairbank Highway Research Center,2006)
[9] Lartillot, O., Toiviainen, P. and Eerola, T. (2008) "A matlab toolbox for music information retrieval." *Data Analysis, Machine Learning And Applications*. pp. 261-268.
[10] Li, C. and Wang, B. (2014) "Fisher linear discriminant analysis." *CCIS Northeastern University*.
[11] Marciniuk, K., Kostek, B. and Czyżewski, A. (2017) "Traffic noise analysis applied to automatic vehicle counting and classification." *International Conference On Multimedia Communications, Services And Security*. pp. 110-123.

[12] Minsky, M. and Papert, S. (1969) ”An introduction to computational geometry.” *Cambridge Tiass., HIT*.
[13] Ohiduzzaman, M., Sirin, O., Kassem, E. and Rochat, J. (2016) ”State-of-the-art review on sustainable design and construction of quieter pavements—Part 1: traffic noise measurement and abatement techniques.” *Sustainability*. **8**, 742.
[14] Partanen, P., Haavisto, V., Haavasoja, T. and Oyj, V. 10.14 Laboratory and Field Test Results from a Pavement Sensor.
[15] Patro, S. and Sahu, K. (2015) ”Normalization: A preprocessing stage.” *ArXiv Preprint ArXiv:1503.06462*.
[16] Rettenbacher, Graf, Rainer and Fiser. (2010) ”A pilot system for acoustic tunnel monitoring.” *EAA Euroregio*.
[17] Rosenblatt, F. (1958) ”The perceptron: a probabilistic model for information storage and organization in the brain.” *Psychological Review*. **65**, 386.
[18] Wong, T. (2017) ”Parametric methods for comparing the performance of two classification algorithms evaluated by k-fold cross validation on multiple data sets.” *Pattern Recognition*. **65** pp. 97-107.
[19] 91–190, P. The National Environmental Policy Act of 1969. (United States,1970), Pub.L. 91–190, Approved January 1, 1970. 42 U.S.C. § 4321 et seq
[20] Ye, J. and Ji, S. (2010) ”Discriminant analysis for dimensionality reduction: An overview of recent developments.” *Biometrics: Theory, Methods, And Applications. Wiley-IEEE Press, New York*.
[21] Saeed, N., Nyberg, R., Alam, M., Dougherty, M., Jooma, D. and Rebreyend, P. (2021) ”Classification of the Acoustics of Loose Gravel.” *Sensors*. **21**, 4944.
[22] Ahmad, A. and Tsuji, T. (2021) ”Traffic Monitoring System Based on Deep Learning and Seismometer Data.” *Applied Sciences*. **11**: 4590.
[23] Szwoch, G. and Kotus, J. (2021) ”Acoustic detector of road vehicles based on sound intensity. ”*Sensors*. **21**, 7781.
[24] Cooke, M., Brown, G., Crawford, M. and Green, P. Computational auditory scene analysis: listening to several things at once. *Endeavour*. **17**, 186-190 (1993), DOI: 10.1016/0160-9327(93)90061-7.
[25] Chu, S., Narayanan, S., Kuo, C. and Mataric, M. Where am I? Scene Recognition for Mobile Robots using Audio Features. *2006 IEEE International Conference On Multimedia And Expo*. pp. 885-888 (2006,7)
[26] Janvier, M., Alameda-Pineda, X., Girin, L. and Horaud, R. (2014) ”Sound representation and classification benchmark for domestic robots.” *2014 IEEE International Conference On Robotics And Automation (ICRA)*. pp. 6285-6292.
[27] Defense Advanced Research Projects Agency DARPA homepage. (https://www.darpa.mil/our-research), (Accessed on April 2022)
[28] Tejedor, J., Macias-Guarasa, J., Martins, H., Pastor-Graells, J., Corredera, P. andMartin-Lopez, S. ”Machine Learning Methods for Pipeline Surveillance Systems Based on Distributed Acoustic Sensing: A Review.” *Applied Sciences*. **7** (2017), DOI:10.3390/app7080841.
[29] Wang, J., Ma, Y., Yang, X., Li, T. and Wei, H. (2021) ”Short-Term Traffic Prediction considering Spatial-Temporal Characteristics of Freeway Flow.” *Journal Of Advanced Transportation*: **2021**
[30] Alford, A. MIT Researchers Investigate Deep Learning’s Computational Burden. (https://www.infoq.com/news/2021/12/mit-deep-learning-burden/, Accessed on April 2022)
[31] Schloegl, A. The NaN-toolbox: A statistics and classification toolbox for Octave and Matlab®. (https://pub.ist.ac.at/ schloegl/matlab/NaN/, Accessed on April 2022)
[32] Gov, A. Quiet Pavement Program, Arizona DOTn. (https://azdot.gov/business/environmental-planning/quiet-pavement-and-noise-program, accessed on April 2022)
[33] Sarker, I. (2021) ”Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions.” *SN Computer Science*. **2**, 1-2.
[34] Sun, S. Digital Audio Scene Recognition Method Based on Machine Learning Technology. *Scientific Programming*. **2021** (2021).
[35] U.S. Department of Transportation. Federal Highway Administatnion. (2013) ”FHWA’s 13 Vehicle Category Classification.” (https://www.state.nj.us/transportation/refdata/roadway/pdf/FHWA13ClassSchemeGraphic.pdf, accessed on April 2022)