Metrics of Expression of Results in Environmental Noise Monitoring Processing: Machine Learning Approach

E. Moke Aman[1], A. Lay-Ekuakille[2], S. Kidiamboko Guwa Gwa Band[1], J. Djungha Okadiowo[3]

[1]Dept of Electronics, ISTA University, Kinshasa, DR Congo
[2]Dept of Innovation Engineering, University of Salento, 73100 Lecce, Italy
[3]UTBC, Bel Campus University of Technology, Kinshasa, DR Congo

***Abstract*— Although modern instruments for environmental noise are equipped with sophisticated devices, the accuracy and uncertainty for discriminating a source from another one, hence a group from another, depend upon the variability of the signal content. Conventional techniques exhibit limits for this kind of variability. The paper proposes an approach according to machine learning technique for processing different levels of environmental noise, classifying them.**

***Keywords*—Environmental noise measurement, Machine learning, Spectral analysis, Signal processing, Artificial intelligence**

1. INTRODUCTION

Environmental noise, defined as unwanted or harmful outdoor sound created by human activities, can be generated by traffic, industry, construction, and recreation activities. Airports, (wind) power plants, rock-crushing, shooting ranges, and motorsport tracks are examples of noise sources for which sound propagation over several kilometers is relevant. One challenge in environmental noise monitoring is how to make sufficiently comprehensive measurements both in time domain and spatially. The changes in weather conditions have a significant effect on monitored noise levels and in order to obtain most of the variations the noise has to be monitored for extended periods of time. Also, a single point noise measurement is rarely representative for a whole neighbourhood and several sensor locations are needed. Because of high costs of the equipment and the amount of human resources needed, the reliability, validity, and representativeness of environmental data is usually unsatisfactory. Only a few reported scientific experiments with uninterrupted noise data captured from each relevant location over long periods of time exist. The typical need for measurements is to monitor the noise caused by a noise source (e.g. an airport, a church, an industrial plant) in a residential area [1]. Typical effects of noise are shown, in terms of quality, in Fig. 1 and the number of exposed people is depicted in Fig.2. However, also other noise sources exist and the captured noise level is usually a result of a combination of the target and interfering sound sources: wind-generated, cars, and birds being examples. Sound level meters used for noise monitoring either capture sound levels or time domain noise data and store the data locally – or nowadays more often – on a remote server.

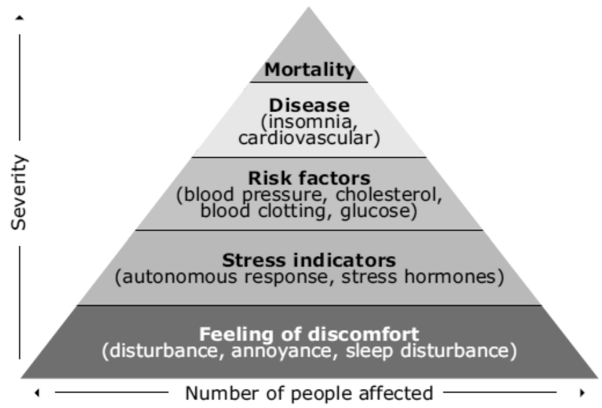

Fig.1. Pyramid of noise effects [1]

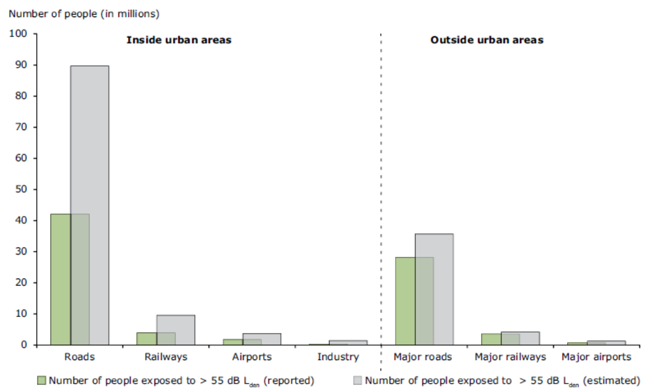

Fig.2. Number of people exposed to noise in Europe > 55 dB L_{den} in EEA member countries (2012): reported and estimated data [1]

The most common method to ensure the noise was caused by the original source is listening through all the samples afterwards. This requires a huge amount of resources because of a large amount of data due to often necessary long-term measurements. Also, if only noise levels are recorded, validation by listening is not possible. In mathematics, the concept of a generalised metric is a generalisation of that of a metric, in which the distance is not a real number but taken from an arbitrary ordered field. In general, when we define metric space the distance function is taken to be a real-valued function. The real numbers form an ordered field which is Archimedean and order complete. These metric spaces have some nice properties like: in a metric space compactness, sequential compactness and countable compactness are equivalent etc. These properties may not, however, hold so easily if the distance function is taken in an arbitrary ordered field, instead of in *R*. The metrics for noise detection are also connected to the band of exposure as illustrated in Fig.3.

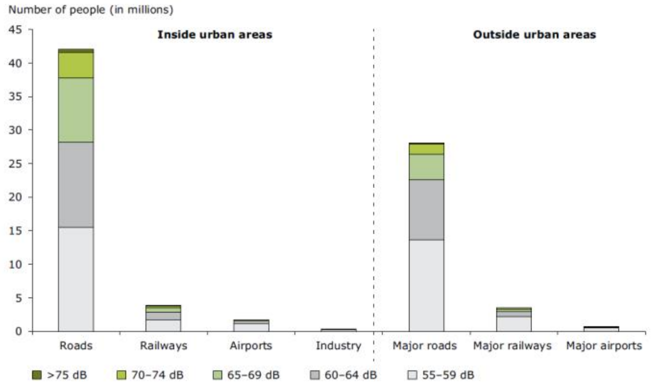

Fig.3. Number of people exposed to noise per decibel band in Europe L_{den} (2012) [1]

2. MACHINE LEARNING APPROACH

Contextually our work is based on Machine Learning and Deep Learning which are all branches of artificial intelligence. Machine Learning is a concept that states that there are generic algorithms that can reveal interesting information about data, without the need to build or develop a specific code. Instead of writing code, you feed these algorithms with data that will allow them to build their own logic. Take for example the classification algorithm [2]. It allows to classify data in groups. This algorithm, used for handwritten number recognition, can also be used to classify mails between spam and non-spam, without changing a line of code. It's the same algorithm but feeds differently it develops a different classification logic.

The basic concepts in computational methods used for analysis of sound scenes and events. Even though the analysis tasks in many applications seem different, the underlying computational methods are typically based on the same principles. We explain the commonalities between analysis tasks such as sound event detection, sound scene classification, or audio tagging. We focus on the machine learning approach, where the sound categories (i.e., classes) to be analyzed are defined in advance.
We explain the typical components of an analysis system, including signal preprocessing [3], feature extraction, and pattern classification. Finally, we explain the whole processing chain that involves developing computational audio analysis systems.

Supervised classifiers
Two types of supervised classifiers are considered: Gaussian mixture model (GMM) as a representative of generative classifiers and artificial neural networks (ANN) as a representative of discriminative classifiers. A GMM represents a class by a distribution of its correspondent feature vectors¨ [4]. The probability density function of a GMM for an observation x is the weighted average of its multi-variate Gaussian distribution components as

$$p(x|\lambda) = \sum_{i=1}^{M} w_i (2\pi)^{-\frac{k}{2}} |\Sigma_i|^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i)} \quad (1)$$

where M is the number of Gaussian components. The parameters of the density model are collectively denoted as

$$\lambda = \{w_i, \mu_i, \Sigma_i; i = 1...M\} \quad (2)$$

The weight, mean and covariance matrix of *i:th* Gaussian component are denoted as w_i, μ_i, Σ_i respectively, satisfying

$$\sum_{i=1}^{M} w_i = 1 \quad (3)$$

The GMM parameters of a class are iteratively estimated using the training data with the expectation maximization (EM) algorithm. Classification can be made using GMMs by outputting the class whose GMM gives the highest likelihood on an input vector *x*.
Like regression, classification also consists of finding the link between a variable *X* and a discrete random variable following a multinomial distribution *Y*.

Further classifications
Given a random variable *X* and a discrete random variable Y, the objective is to approximate the function

$$E(Y|X) = f(X) \quad (4)$$

The problem data is a sample of points: $\{(X_i X_i) \mid 1 \le i \le N\}$ with $\forall i \in \{1,...,N\}, Y_i \in \{1,...,C\}$ and a model parameterized with θ:

$$\forall i \in \{1,\dots,N\}, \forall c \in \{1,...,C\}, \mathrm{P}(Y_i = c \mid X_i \theta = h(\theta, X_i, c) \quad (5)$$

with $n \in \mathbb{N}, h$, h is a function of parameter θ that is included in [0,1] and verifying the constraint:

$$\sum_{c=1}^{c} h(\theta, X, c) = 1 \quad (6)$$

The first example is a classification in two classes, it consists of discovering the link which unites a real random variable *X* and a discrete random variable and $Y \in \{o, 1\}$,for that we have a list:

$$\{(X_i, Y_i) \in \Re \times \{0,1\} \mid 1 \le i \le N\} \quad (7)$$

3. PROPOSED ALGORITHM

In the proposed automatic target source detection system, noises are defined into two classes. Sounds propagating [5] from the target sources belong to a target class, whereas interfering noises as well as silence belong to a background class. Examples of possible target (church) sounds are plant noise and aircraft noise. Possible back-ground noises may be caused by e.g. traffic, wind, rain, thunder, and birds. The activity of the target sources is detected by analysing continuous audio input and making binary classification between the background and the target. The audio input is the same as the signal used for SPL measurement, but without the A-weighting filter.

The detection system consists of two stages: the training stage and the monitoring stage (see Fig.4). Acoustic models are learned from training examples, captured audio with manual annotation, in the training stage. The learned acoustic models are used to classify audio captured on a sensor, to detect the activity of target, in the monitoring stage. The training algorithm needs only annotation of target sounds. Traffic sounds, regarded as background are annotated to help understanding the system output.

It is not easy to directly determine a function h which approximates $Y|X$ because h and Y are both discrete. Therefore, rather than directly solving this problem, it is better to determine the marginal law so that $\mathbb{p}(Y = c|X) = f(X, \theta, c)$ f is then a function whose outputs are continuous and can be chosen to be differentiable. For example, f can be a neural network whose outputs verify: $f(X, 0)\ f(X, 1) = \mathbb{p}\ (1|X) = 1$

$$f(X,0)\ f(X,1) = p(1|X) = 1 \qquad (8)$$

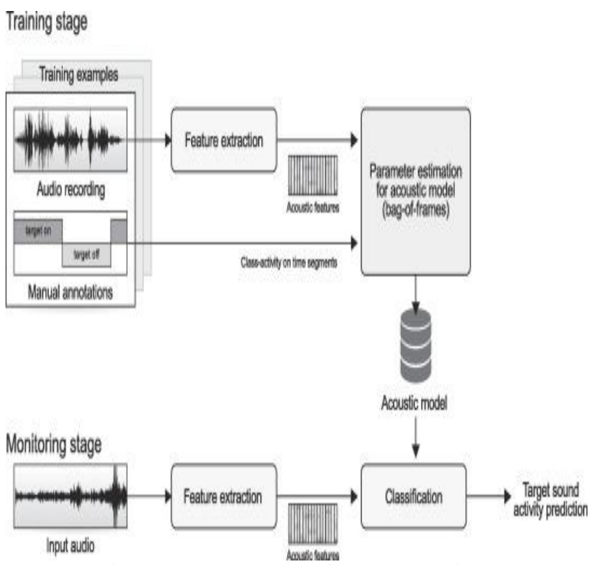

Fig.4. Flowchart of the proposed algorithm for noise classification

The neural network used for this task is slightly different from the previous one. A plane has been divided into two semi-planes by a straight-line delimiting two classes, the neural network whose hidden layer contains two linear neurons, has found this separation despite the few misclassified examples.

4. RESULTS AND CONCLUSIONS

The measurement instrument is a portable system with an electret transducer. An electret microphone is a variation of the condenser microphone. Instead of requiring an external voltage source to charge the diaphragm, an electret microphone uses a permanently charged plastic element (electret) placed in parallel with a conductive metal backplate. The portable instrument exhibits the following features: (i) measuring range: 30-60 dB, 50-80 dB, 70-100 dB, 90-120 dB; (ii) resolution: 0.1 dB; (iii) accuracy: 30-60 dB±3 dB, 60-120 dB±2 dB; (iv) frequency range: 31.5 Hz-8kHz.

Table 1 Noise intensity from 64 religious areas

Sites	Intensity in dB	Sites	Intensity in dB
	81,1		69
	70,2		63,2
	73,2		78,3
	90		64,7
	79,6		75
	79,9		82,8
	81,4		70,1
	78,4		68,2
	75,2		74,3
	78,1		71,4
	64,5		72,1
	77,8		69,3
	75,8		70,1
	84,5		68,1
	76		73
	66		75,1
	89,6		79
	86,2		68,2
	75,2		78
	74		76
	72,5		73,1
	79,2		69,8
	73		76,2
	82,2		84,2
	79,1		81,3
	77,9		63
	81,1		71,3
	68		

A campaign has been conducted in 2019 for measuring noise in 64 religious activities (Table 1), such as churches, and 51 recreational centers (Table 2) such as leisure. The proposed algorithm using a specific classifier brings to the results described below.

Table 2 Noise intensity from 51 recreational centers

Site	Intensité dB	Site	Intensité dB
	84,7		90
	87		85
	91		87,4
	92,1		88
	82		79
	90		83,1
	79		89,1
	81,7		83
	89,7		93
	91,3		86
	82		83
	90		89
	88		79
	87,1		71
	84,6		79
	89,8		89
	77,9		82
	80		75
	92		80
	85		82
	80		78
	90		79
	82,1		82
	87,9		83
	76,1		86
	78		85
	79		85,1
	80		

The algorithm based on classification delivers two maps by displaying natural and artificial noises. With natural noise (blue colour), we intend human voice and wind, whilst artificial noise (red colour) is characterized by vehicle traffic, musical instruments, and aircraft).

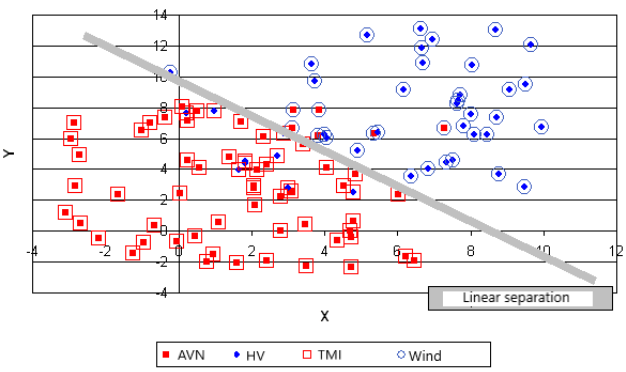

Fig.5. Results of the first classification

From Fig.5 we can see an almost clear separation between natural and artificial noises but in Fig.6 the separation/classification includes an area with a little unclear area. But we are able to understand some areas close to the airport of Kinshasa (DRC) where part of the campaign has been conducted. We also understand the importance of wind in displacing noise and sound from one area to another.

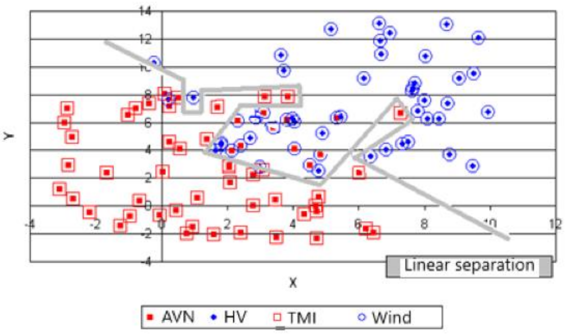

Fig.6. Results of the second classification

Finally, a quantification of the noise is performed, and the algorithm delivers a quantification per class, namely human voice, traffic, and wind. The maximum is 80 dB, mostly caused by traffic, and a little bit by human voice. Fig.7 reports these data.

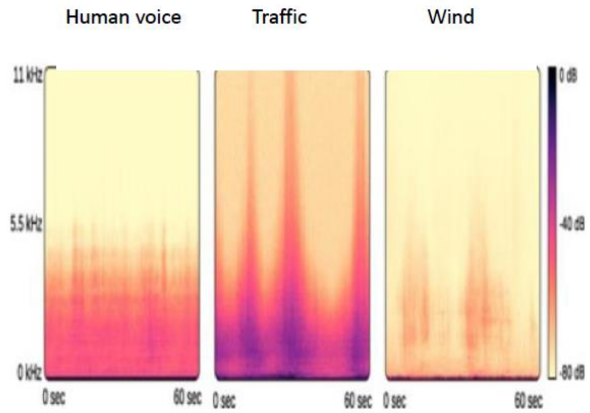

Fig.7. Intensity quantification and classification

In the concluding remarks, we notice the algorithm is able to classify the different sources of noise in detailed manner, and it is also able to grouped them by homogenous categories. Environmental noise detection still remain an issue related to spectral estimates [6] [7] [8]. The artificial intelligence approach is a strong support to classify the diverse sources of noise.

6. REFERENCES

[1] European Environmental Agency, “Noise in Europe 2014”, Publications Office of the European Union, 2014, doi:10.2800/763331

[2] D. Verma, A. Jana, K. Ramamritham, “Classification and mapping of sound sources in local urban streets through AudioSet data and Bayesian optimized Neural Networks”, Noise Mapp., Vol. 6, pp. 52–71, 2019

[3] H. Purwins, et al., “Deep Learning for Audio Signal Processing”, Journal of Selected Topics of Signal Processing, Vol. 13, n. 2, pp. 206–219, 2019.

[4] C. E. Rasmussen, “The Infinite Gaussian Mixture Model“, in Advances in Neural Information Processing Systems 12 S.A. Solla, T.K. Leen and K.-R. Muller (eds.), pp. 554–560, MIT Press, 2000

[5] M. Avoci Ugwiri, I. Mpia, A. Lay-Ekuakille, "Vibrations for Fault Detection in Electric Machines", IEEE Transactions on

Instrumentation & Measurement Magazine, Vol 23, n. 1, pp. 66-72, 2020

[6] A. Lay-Ekuakille, P. Vergallo, G. Griffo, F. Conversano, S. Casciaro, S. Urooj, V. Bhateja, A. Trabacca, "Multidimensional Analysis of EEG Features Using Advanced Spectral Estimates For Diagnosis Accuracy", 2013 IEEE Memea, 4-5 May, 2013, Gatineau – Ottawa, Canada.

[7] F. Conversano, R. Franchini, S. Casciaro, A. Lay-Ekuakille, "In Vitro Evaluation and Theoretical Modelling of the Dissolution Behaviour of a Microbubble Contrast Agent for Ultrasound Imaging", IEEE Sensors Journal, vol.12, n.3, pp. 496-503, 2012.

[8] A. Lay-Ekuakille, V. Telesca, M. Ragosta, G.A. Giorgio, P. Kapita, S. Kidiamboko, " Supervised and Characterized Smart Monitoring Network for Sensing Environmental Quantities: Characterization", IEEE Sensors Journal, Vol.17, n.23, pp. 7812-7819, 2017