

A ROBUST EVENT DETECTOR FOR ENVIRONMENTAL NOISE BASED ON L_{Aeq} TIME PROFILES

Daniel de la Prida¹, Miguel Molina-Moreno^{2,3}, Luis A. Azpicueta-Ruiz², Antonio Pedrero¹

¹ Grupo de Investigación en Acústica Arquitectónica, Universidad Politécnica de Madrid
{e-mail: daniel.prida@upm.es; antonio.pedrero@upm.es}

² Department of Signal Theory and Communications, Universidad Carlos III de Madrid
{e-mail: migmolin@ing.uc3m.es; lazpicue@ing.uc3m.es}

³ Department of Immunobiology, Yale University
{e-mail: miguel.molinamoreno@yale.edu}

Abstract

Environmental noise monitoring systems often require the detection and identification of the different types of sound events, in order to evaluate the acoustical impact generated by the various types of acoustic emitters that coexist in such environments, like cars, trains, aircraft, etc. These systems often require the joint analysis of several complex acoustic characteristics or the comprehensive analysis of recordings of the environments in which they are located. This fact hinders both the implementation of the detection task in the measuring devices themselves, as well as the immediate detection of sound events, requiring this task to be performed after the monitoring.

In this paper we present a fully interpretable event detection system, based on Machine Learning, which only requires the temporal profile of the L_{Aeq} for its operation. This fact makes the detector lightweight and easily implementable into sound level meters and other simple noise measurement devices. In addition, a data augmentation procedure is proposed, which allows the system to adapt to highly noisy acoustic environments, but being trained only in favorable environments. The performance of the system is demonstrated using a train detection case study.

Keywords: noise monitoring, environmental noise, sound events, machine learning.

Resumen

Los sistemas de monitorización de ruido ambiental requieren, en muchas ocasiones, la detección e identificación de los distintos tipos de eventos sonoros, para evaluar la afección sonora generada por los distintos tipos de emisores acústicos que conviven en estos ambientes, tales como coches, trenes, aviones, etc. Estos sistemas requieren, a menudo, el análisis conjunto de diversas características acústicas complejas o del análisis completo de grabaciones de los entornos en los que se sitúan. Este hecho dificulta tanto la implementación de la tarea de detección en los propios aparatos de medida, como la detección inmediata de los eventos sonoros, precisándose realizar esta tarea de manera posterior a la monitorización.

Se presenta en esta comunicación un sistema de detección de eventos basado en *Machine Learning*, totalmente interpretable y que sólo requiere, para su funcionamiento, del perfil temporal del LAeq. Este hecho hace que el detector sea ligero y fácilmente implementable en sonómetros y otros dispositivos sencillos de medición sonora. Además, se propone un procedimiento de *data augmentation* que permite al sistema ser entrenado en ambientes favorables, pero que funcione correctamente en ambientes sonoros altamente ruidosos. El desempeño del sistema se presenta en un caso de estudio destinado a la detección de trenes.

Palabras clave: monitorización de ruido, ruido ambiental, eventos sonoros, aprendizaje automático.

PACS nº. 43.50.Rq, 43.50.Yw

1 Introducción

Noise is currently one of the main pollutants in urban environments. The emitters contributing to noise in these environments are diverse and have multiple particularities. However, the noise generated by infrastructures such as roads, airports or railways constitutes one of the main components of noise in urban environments [1].

With the aim of assessing the impact caused by each type of infrastructure, there are rules and regulations aimed at evaluating and controlling the emitters of these types of infrastructures [2].

In order to effectively assess the impact of an infrastructure and its emitters on an environment, it is essential to identify the noise events that solely belong to the infrastructure under study. For this purpose, there are various approaches, ranging from the manual identification of noise events, when the measurements aimed at assessing the infrastructure are of short duration, to complex automatic identification. In fact, this variety of techniques is contemplated in standards, such as ISO 20906:2009 [3], related to the unattended monitoring of aircraft sound events. Sometimes these techniques are even used together, with the first stage involving an automatic detection and, in doubtful cases, a subsequent manual evaluation by an experienced operator. Human experts, while requiring in some cases the listening to the sound recordings made synchronously with the measurement of acoustic parameters, are usually able to identify whether a noise event is of interest by simply observing the acoustic profile in the measurement environment, i.e., by observing the time-dependent evolution of a global parameter such as LAeq. Automatic systems generally rely on the analysis of audio recordings from the environments under study [4,5] or at least of relatively complex acoustic descriptors derived from their spectral behavior [6–8], which usually entails that this identification must be carried out offline, on a different and more powerful device than the measurement one, and at a later point in time. This means not only that the identification of the events is not obtained in real time, but also that it is usually necessary to send heavy data from the measuring device to the one in charge of the identification.

This communication presents an automatic detector, useful for identifying sound events of interest in environmental noise, which tries to mimic the behavior of a human expert. For this purpose, it bases the detection of the noise events of interest only on the analysis of the acoustic profile of a measurement, i.e., the time-dependent evaluation of the LAeq. This fact makes feasible the implementation of this automatic detector into the noise monitoring systems and allows the identification of sound events in real time. The detection process consists of two elements: (1) the sound event identifier in the acoustic profile of the evaluation environment (called Acoustic Profile Event Detector - APED) and (2) the classifier of interest of the identified events. In this paper we present two variants of the classifier, one based on an extremely efficient and interpretable Machine Learning algorithm, Logistic Regression, and another somewhat more complex, but also efficient, implementable and interpretable algorithm, based

on a Long Short-Term Memory Autoencoder (LSTM AE). The fact that both approaches are interpretable makes it possible to infer the characteristics of the sound profile that are most decisive when determining whether an event belongs to the categories of event of interest or event of no-interest.

A data augmentation strategy is also presented, aimed at training these classifiers in a more robust way, for better performance in a wide variety of noise scenes different from those existing in the data collected for the training of the systems.

These detection and data augmentation approaches are tested in a case study aimed at the identification of railway pass-by (RPB) events in urban environments, which demonstrates the great applicability, low computational cost and high performance of the proposed approaches. Nevertheless, the presented system is specially designed to be transferable, with appropriate adaptations, to the detection of other types of noise events, by re-training the classifiers with proper databases comprising events of other types.

2 Methodology

2.1 Gathering of training and test data

The noise event detector is tested in our research for the identification of RPB events. For that purpose, measurements were conducted in urban environments with a complex variety of noises, coming not only from trains, but also from other types of sources, such as traffic noise, pedestrian noise, birdlife noise, etc.

For the training and testing of the detector, around 930 h of noise measurements were carried out in two different scenarios:

- clean scenario (scenario A): the sound analyzer was placed close to the track, 1.5 m above the ground and about 15 m from the track in a straight line (Figure 1:left). The environment was quiet, and the events were mostly due to RPB. About 20 hours were measured in it, considered of high SNR.
- noise polluted scenario (scenario B): the sound analyzer was placed at a distance of about 28 m from the track and at a height of about 5 m (Figure 1:right). The microphone was pulled out through a window of a building and placed 2 m from the façade. Between the railway and the microphone there was a street with heavy traffic and pedestrian noise. Approximately 900 h were measured in this scenario, which is considered to have a low SNR.

In these two scenarios, both the A-weighted equivalent continuous noise level for each second ($L_{Aeq,1s}$), for the total duration of the measurement, together with the audio signal, were captured. A Brüel & Kjaer Type 2250 sound analyzer was used for this purpose, with its calibration duly verified at the beginning and end of each measurement session by means of the Brüel and Kjaer Type 4231 acoustic calibrator.

As mentioned in the introduction, one of the most important contributions of our proposal is that the detector only requires the $L_{Aeq,1s}$ of a noise monitoring situation to classify the events. However, it was decided to record, for this experiment, the audio signal, for reasons that will be described in detail in the next section, and that are connected to the data augmentation policy. This being so, it is important

to emphasize that the use of audio recordings is never necessary, in our approach, for the execution of the detection tasks in the wild.



Figure 1: Left: clean scenario (scenario A). Right: noise polluted scenario (scenario B)

In addition to the recording of LAeq,1s profiles and sound recordings in the two scenarios, an automatic annotation of RPB was performed to provide labels of event of interest (RPB events) and event of non-interest (other events), with which to train and test the detector. In both scenarios, the annotation could be performed by automatically time-aligning a public database of pass-by times from the rail infrastructure manager (ADIF) with the temporal tags of the LAeq,1s profiles recorded by the sound analyzer.

2.2 Data augmentation policy

Data augmentation is the technique of performing realistic transformations on the data fed as input to Machine Learning algorithms. This process can help to make the training of the detector more robust to discrepancies between the acoustic characteristics of the training environment and those of the different measurement environments in which the potential measuring device embedding the detector is used.

The data augmentation technique, usually employed in computer vision [9], basically takes the available data and generates a new collection of data, with certain variations of some of its characteristics. In this way, if the training of the detector is carried out with data with greater variability (i.e., the one obtained after data augmentation), it can be more effective in scenarios different from the one where the detector was trained.

Data augmentation techniques are varied, and in the case of audio signal processing they usually involve processes such as shifting the pitch of the audio signals, stretching the spectrograms or adding white/pink noise [10].

In our case, the data augmentation technique is based on the following principle: the acoustic profiles measured in the clean scenario A mainly have their abrupt changes in the LAeq due to RPB. If these profiles are contaminated with acoustic profiles from other noisier urban environments in which the profile variability is due to a myriad of sound events of a nature other than railways, such as vehicles,

aircraft, etc., the trained detector will be more robust to changes in the background noise of the evaluation environments.

The data augmentation based on the described approach can be carried out in, at least, two ways:

- (1) the addition, directly to the acoustic profiles of scenario A, of the profiles of other noisy situations. This would have resulted in the incoherent second-by-second energetic addition of the LAeq profiles of scenario A with the profiles of other noisy environments, with different SNRs between scenario A and the other noisier environment. This approach, which is not the one followed in this research, would have entailed the assumption of perfect incoherence between the noise from scenario A and that of the other noisy environments.
- (2) the summation of the samples of the sound recordings of scenario A and those of other noisier environments, with different SNR, and the subsequent computation of the combined LAeq acoustic profile. In this case, when summing the samples of the temporal signals, it is not necessary to make any assumptions on the coherence or incoherence of the noises to be added. Of course, this approach requires not only the acoustic profiles, but also the recordings in the environments. However, as it has been done before, it is imperative to emphasize that the need for the recordings was only for data augmentation and is never necessary for the detection task, which is carried out solely and exclusively on the basis of the LAeq profiles captured by a monitoring system.

Therefore, for the data augmentation policy, additional sound recordings were made in various urban environments with a wide variety of noise sources. From these recordings, sound events of sources other than railway noise were extracted, which then were added to the acoustic profiles of scenario A by using the approach presented above in point (2). In this way, a new collection of profiles from a synthetically polluted version of scenario A was obtained, contaminated with the profiles of other noises. In the results section, detection performance metrics will be presented for both when the detector is trained with the clean scenario A data and with the synthetically polluted scenario A after data augmentation.

In addition to the data augmentation policy described above, other simple domain adaptation [11] techniques (i.e. adding samples of sound events that do not belong to RPBs to the training data) were also used during the research, which allow, without applying the data augmentation process described above, to improve the robustness of the detection in presence of events different from those existing in the training data environment.

2.3 Acoustic Profile Event Detector (APED)

The first stage of the detection process consists of the identification, in the long-term LAeq profiles, of noise events that are particularly noticeable with respect to the baseline background noise of the monitored acoustic environment. This first stage is agnostic, meaning that the identified events are not classified into events of interest (i.e. RPB in our case study) and events of non-interest (other events), but are simply marked as events with their start and end times returned.

The APED is very simple and fully parametric. To identify that an event that deserves assessment has occurred, the APED scans the acoustic profile and marks as events those time periods that meet the following two conditions:

1. all consecutive LAeq,1s values in a time interval exceed a configurable S_{th} threshold in dB,
2. the duration of this interval is above a certain time threshold L_{th} , in seconds.

These parameters can be chosen in several ways, such as on the basis of expertise, or through a preliminary evaluation in the monitoring environment, which was the case here. In our case study, the selection of the S_{th} and L_{th} values, that will be shown in the results section, allowed us not to miss any of the RPB events that occurred in scenarios A and B.

2.4 Interest event classifier

Once the event identification task has been performed by the APED, the events go through the classifier, which will determine the probability that identified events correspond to RPB events (events of interest) or other types of events (events of non-interest).

In this research, the classification task has been approached from two different Machine-Learning strategies.

On the one hand, a Logistic Regression based classification approach (LR classifier [12]) has been used. This strategy has the advantage of being highly interpretable. It computes a set of features from the evolution of the $L_{Aeq,1s}$ profiles, and allows to identify which ones of them have the highest relevance on the classifier to determine the interest or non-interest of an event. On the other hand, this approach has the disadvantage that the features to be extracted from the profile to make the decision must be handcrafted by human expertise. Accordingly, this approach is divided into two stages: (1) feature extraction from the evolution of the $L_{Aeq,1s}$ profiles during an event and, (2) the classification on the basis of the computed features.

The second strategy is based on an Autoencoder using a Long Short-Term Memory [13] network (LSTM AE). This strategy has the advantage of not requiring neither two individual stages nor the human selection of features, since the AE system is both in charge of extracting the relevant features from the profile evolution as well as of making the classification. However, its computational cost is somewhat higher and it is not as interpretable as the LR-based approach.

A summary of the pipeline followed by both approaches, which will be explained below, can be seen in Figure 2.

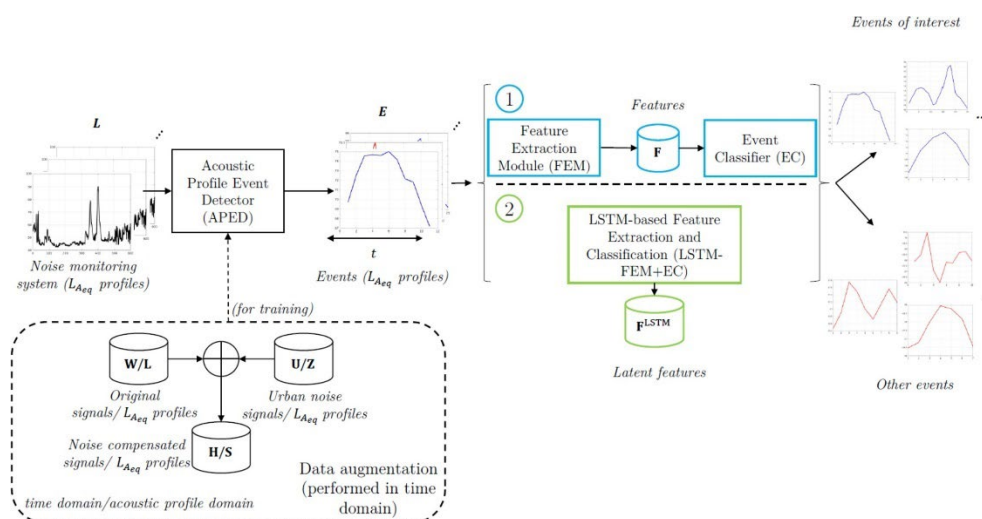


Figure 2: Pipeline of our robust detection system, considering both strategies, for the classification of the events [14].

2.4.1 LR classifier

The LR classifier [12], as already mentioned, requires the knowledge of certain features of the acoustic profiles, both for the training and test of the system. These features, which are not extracted by the classifier itself, as in the case of LSTM AE, were determined, in our research, on the basis of human expertise. Experts in sound event identification generally rely on the observation of several features of the evolution of sound pressure levels across an event to make their judgements. These characteristics are related to the magnitude of the event profile, its shape and the variations of levels throughout the event. In this sense, and with this in mind, eight features, which were considered especially relevant, were selected to be used by the classifier. These were: the mean and standard deviation of the LAeq,1s over the length of the event, the skewness and the kurtosis of the event's profile, the crossing rate, as well as the mean and width of the crossing rate. The crossing rate assesses the number of times the LAeq,1s crossed the mean value of the LAeq,1s across the duration of the event. Further details on the particulars of these features and the reasons for their selection can be found in [14].

This classifier provides an output with the probability that the event under study is of interest according to the following rule:

$$p_i = \frac{1}{1 + \exp(\beta_0 + \beta_1 * f_i^1 + \dots + \beta_8 * f_i^8 \dots)} \quad , \quad (1)$$

where β are the weights of each of the eight characteristics used and f_i is the value of each of the features.

Therefore, in a test environment, the LR classifier can give a probability that an event is of interest by simply applying the above formula, based on the described features.

For this formula to be applicable in a test environment, it is mandatory to establish, during training, the value of each of the β weights. This calculation is performed by minimizing a Binary Cross-Entropy (BCE) cost function with L2 regularization, as described below:

$$L = \frac{1}{N} \sum_{i=1}^N y_i \log(p_i) + (1 - y_i) \log(1 - p_i) + \lambda \sum_{j=1}^G \|\beta_j\| \quad , \quad (2)$$

with λ being the optimal weighting parameter between L2 regularization and Binary Cross-Entropy loss, obtained by cross-validation.

2.4.2 LSTM AE classifier

The proposal based on an LSTM Autoencoder is very novel and addresses the task of feature extraction and classification jointly [13,15], not requiring manual feature selection in a different step. For this purpose and because a human expert would observe the evolution of the dynamics of the LAeq,1s values of an event to determine its interest, the encoder follows a similar process. Based on an LSTM-layer, the encoder extracts certain unknown features which are then contained in its hidden state, and which are subsequently summarized into a reduced set of latent features, by applying a block of LayerNorm+FullyConnected layers. From these features, a FullyConnected layer acts as classifier and outputs the probability that an event is of interest.

On the other hand, the decoder, using the latent features, tries to reconstruct the time-dependent evolution of the LAeq,1s profiles of the events. The minimization of the error between the reconstructed

signal and the original signal therefore forces the latent space to contain as much information as necessary for the reconstruction from the original scenario and the latent features to be meaningful.

A complete representation of the process followed by the LSTM AE classifier can be seen in Figure 3:

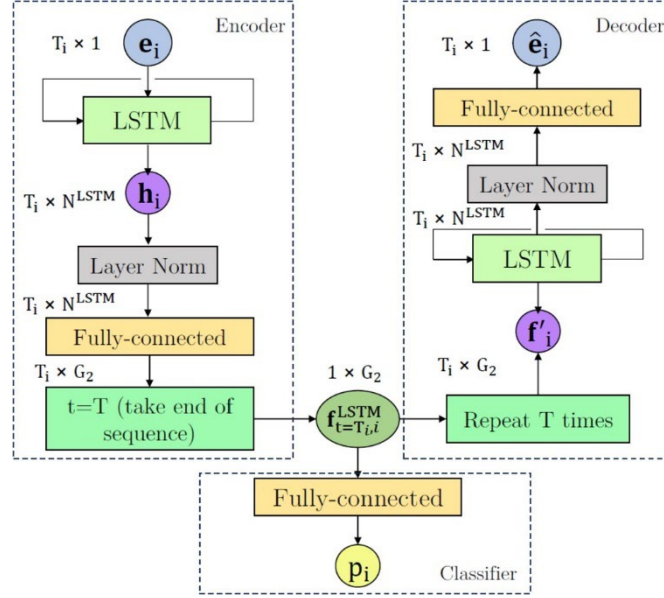


Figure 3: Diagram of the process followed by the LSTM AE classifier: the encoder takes the LAeq profile of an event (e_i) and summarizes its evolution in G_2 latent features. From these latent features, the classifier outputs the probability (p_i) of an event to be of interest. The decoder forces the latent space to be representative by reconstructing (\hat{e}_i) the dynamics of the original profile [14].

The system is trained by the combined, unweighted minimization of two cost functions:

- (1) BCE loss cost function (like the previous approach, but without L2 regularization), to minimize the classification error:

$$L_{BCE} = \frac{1}{N} \sum_{i=1}^N y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \quad , \quad (3)$$

- (2) MSE loss cost function, to minimize the error of reconstructing the LAeq profiles of the events by the decoder from the latent features:

$$L_{MSE} = \frac{1}{N} \sum_{i=1}^N \|e_i - \hat{e}_i\|^2 \quad , \quad (4)$$

Since the contribution of these two cost functions was considered to be equally important, the total loss function is simply an unweighted sum of both of them :

$$L = L_{BCE} + L_{MSE} \quad , \quad (5)$$

3 Results

In this section we present the fundamental results of our proposal in terms of event identification by APED, as well as correct classification of events of interest by both the LR classifier and the LSTM Autoencoder classifier. Additional results on the interpretability of the classifications, as well as a comparison of the performance of our system with respect to more complex systems in the state of the art, based on more resource-consuming features, can be found in our recent journal publication [14].

3.1 APED event identification

As mentioned in the methodology section, the APED relies on two parameters to consider a time interval of the LAeq,1s acoustic profiles as constituting an event. These parameters were the level threshold S_{th} and the temporal threshold L_{th} .

The experimental selection of $S_{th} = 66 \text{ dB}$ and $L_{th} = 4 \text{ s}$ allowed us not to miss any RPB event in either of the two scenarios measured (Figure 4). With these parameters, the APED identified 1434 time intervals, of which 273 belonged to scenario A and 1161 to scenario B. Of these 1434 intervals, 749 actually belonged to RPB events and 685 to other events, leaving the additional work of event classification for the second stage of the system.

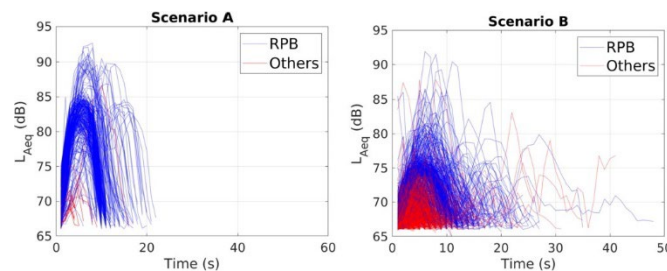


Figure 4: LAeq,1s event profiles for both the scenario A and scenario B.

3.2 Event classification by the LR classifier and the LSTM AE classifier

This section briefly presents performance results of the two classification approaches described in the methodology. The performance of both classifiers is presented in four situations:

1. (B/A): Training the classifiers with the data from noisy scenario B data and testing with clean scenario A data. This constitutes a rather unrealistic upper limit of detection, where the classifiers are trained on rather complex data and then used in much simpler environments.
2. (A/B): Training the classifiers with the clean scenario A data and testing them with the data from the noisy scenario B. This is a more realistic case study, where it is considered that the classifier may encounter more complex test situations than those used for training.
3. (A+UN/B): Training of the classifiers with the data from scenario A, clean, as well as with additional data from events not related to RPB (a simple strategy for domain adaptation). Testing with data from noisy scenario B. The fact of feeding the training both with data from scenario A but also with data from other usual types of events, but different from train events, should make the system more robust to the presence of other types of events.

4. (A+UN+AUG/B): Training of the classifiers with the clean scenario A data, with additional data from events other than RPB (simple domain adaptation, as in 3.), but also with scenario A data but contaminated with different urban background noises, with different SNRs (our proposed data augmentation policy). Testing on the data from scenario B. This case should be the most robust in our case, since, in addition to training the system with data from other types of events, the classifiers are trained with RPB data from scenario A but highly contaminated with other types of events, plausible in urban environments.

The results of performance for each of the classifiers, considering each of the training and test situations, are presented in Table 1. The Area Under the Receiver Operator Characteristic (AUC) metric, in percentage, is used to evaluate the performance:

Training dataset	B	A	A+UN	A+UN+AUG
Test dataset	A	B	B	B
LR classifier	94.19	81.82	86.21	88.46
LSTM AE classifier	95.53	87.55	87.71	90.58

Table 1: Performance of the detector in AUC (%)

As can be seen, both classifiers have, in general, good abilities to identify and classify the events of interest. However, as it can be observed, both classifiers have difficulty in correctly classifying some of the more complicated events, the upper limit being an AUC of 95.53%.

It can also be observed that the LSTM autoencoder-based classifier generally performs better than the LR-based classifier. As observed in the analysis of the LSTM AE classifier, this approach uses four latent features for the decision-making, compared to the eight handcrafted features used by the LR classifier. This fact highlights how the imbrication between feature extraction and classification performed by the LSTM AE classifier allows to reduce the dimensionality of the features while improving the performance in comparison to the LR classifier.

From the third case (A+UN/B) it can be derived that domain adaptation, by itself, was useful for the LR classifier but not for the LSTM AE classifier.

Finally, it can be observed that the strategy (A+UN+AUG) presents very high detection results. The fact that the system is trained not only with data from the clean scenario A, but also with this data, but polluted by challenging background noise from other urban environments, makes the training data of this strategy very similar in characteristics to the data from the very noisy and poorly controlled urban scenario (scenario B) of the test data.

Finally, the fact that both classifiers use such few features which, moreover, are all based on the LAeq,1s profile captured by a sound analyzer, highlights the high efficiency of these detectors, which could be easily implemented in monitoring devices, providing real-time detection results. Additional findings on the computational cost of the proposed solutions, as well as a comparison with more complex solutions described in the state of the art, can be found in [14].

4 Conclusions

This paper has presented two Machine Learning-based approaches for the efficient classification of acoustic events of a specific type (e.g., railway pass-by in our case). One of the main contributions of the research is that the presented classifiers only rely on the knowledge of the acoustic profile of an

environment (i.e., the evolution of LAeq,1s along the time) to perform their detection tasks, without the need to use audio signals or other complex spectral or spectro-temporal metrics. To evaluate its applicability to real situations and its efficiency, the proposal has been tested with a case study aimed at detecting rail pass-by events in noisy urban environments.

The research results highlight several facts:

- (1) the assessment of the evolution of LAeq,1s values in a monitoring environment should be sufficient for the detection of noise events for particular applications. This fact has been verified with the high classification accuracy results in our RPB detection case study, which have reached accuracy values of around 90%.
- (2) the high efficiency of the proposal presented in this research should allow embedding event detection tasks in noise monitoring systems, even for the most low-cost ones. The two classification approaches presented use between four and eight features, purely based on the acoustic profile, to carry out their classification tasks.
- (3) the system presented should be adaptable and robust to noise environments different from those used for training, considering the data augmentation policy proposed. This has been proven with increases of between 3% to 8% in detection accuracy, when going from a training based on a clean scenario to a scenario with a training based not only on clean data but also on noise-polluted data obtained by using data augmentation techniques.

Furthermore, the good interpretability of the two proposed Machine Learning solutions allows to infer which characteristics of the magnitude, shape or variability of the acoustic profiles are the most relevant when identifying an event of interest in the trained application.

In summary, our research has highlighted that, in continuous noise monitoring situations where it is necessary to carry out an identification of the types of acoustic emitters that contribute to noise, such as train or aircraft noise monitoring, these tasks could be efficiently carried out by the monitoring system itself, in real time, and unattended, through the efficient analysis of LAeq profiles with the proposed Machine Learning detectors.

Acknowledgments

The work of D. de la Prida and L. A. Azpicueta-Ruiz was partly supported by Grant PID2021-124280OB-C21, funded by MCIN/AEI/ 10.13039/501100011033 and by the “ERDF A way of making Europe”.

M. Molina-Moreno was supported by the Spanish Ministry of Education, Culture and Sports FPU Grant FPU18/02825 and the Regional Grant 2020/NMT-6660 of the Autonomous Region of Madrid.

References

- [1] Peris E. Environmental noise in Europe - 2020. vol. 1. Publications Office of the European Union; 2020. <https://doi.org/10.2800/686249>.
- [2] de España G. Real Decreto 1367/2007, de 19 de octubre, por el que se desarrolla la Ley 37/2003, de 17 de noviembre, del Ruido, en lo referente a zonificación acústica, objetivos de calidad y

- emisiones acústicas. BOE nº 254. de 23/10/2007. <http://www.boe.es/buscar/act.php>; 2007.
- [3] ISO 20906:2009 - Acoustics -- unattended monitoring of aircraft sound in the vicinity of airports. International Organization for Standardization; 2010.
- [4] Valero X, Nencini L, Alías F, Vinci B. Feasibility of automatic noise source recognition in collaborative wireless sensor networks. Proc AIA-DAGA 2013.
- [5] Haddad K, Song W, Valero X. Environmental sound classification in realistic situations. Proc. Forum Acusticum, 2014, p. 6.
- [6] Sobreira-Seoane MA, Rodríguez Molares A, Alba Castro JL. Automatic classification of traffic noise. J Acoust Soc Am 2008;123:3823.
- [7] Socoró JC, Ribera G, Sevillano X, Alías F. Development of an Anomalous Noise Event Detection Algorithm for dynamic road traffic noise mapping. Proc. 22nd Int. Congr. Sound Vib. (ICSV22), Florence, Italy, 2015, p. 12–6.
- [8] Ullo SL, Khare SK, Bajaj V, Sinha GR. Hybrid computerized method for environmental sound classification. IEEE Access 2020;8:124055–65.
- [9] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data 2019;6:1–48.
- [10] Mushtaq Z, Su S-F. Environmental sound classification using a regularized deep convolutional neural network with data augmentation. Appl Acoust 2020;167:107389.
- [11] Farahani A, Voghoei S, Rasheed K, Arabnia HR. A brief review of domain adaptation. Adv Data Sci Inf Eng Proc from ICDATA 2020 IKE 2020 2021:877–94.
- [12] McCullagh P. Generalized linear models. Routledge; 2019.
- [13] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9:1735–80.
- [14] Molina-Moreno M, de la Prida D, Azpicueta-Ruiz LA, Pedrero A. A noise monitoring system with domain adaptation based on standard parameters measured by sound analyzers. Appl Acoust 2024;218:109892.
- [15] Kramer MA. Nonlinear principal component analysis using autoassociative neural networks. AIChE J 1991;37:233–43.