



## PROJETO MIRA – DESENVOLVIMENTO DE UMA PLATAFORMA INTELIGENTE DE MONITORIZAÇÃO DE RUÍDO AMBIENTAL

*Luís Conde Santos\**, *Cristina Leão*, *João Figueiredo*<sup>1</sup>  
*Afonso Melo*, *Guilherme Covas*, *André Carvalho*<sup>2</sup>

<sup>1</sup> dBwave.i - Acoustic Engineering, SA, Oeiras, Portugal ([www.dbwave.pt](http://www.dbwave.pt))

<sup>2</sup> Instituto de Soldadura e Qualidade, Oeiras, Portugal ([www.isq.pt](http://www.isq.pt))

### RESUMO

A poluição sonora, identificada pela Organização Mundial de Saúde como um dos principais fatores de Risco Ambiental para a Saúde [1], é objeto de atenção crescente por parte da sociedade, gerando frequentemente reclamações e conflitos entre a população e as entidades responsáveis pela emissão de ruído, cada vez mais com recurso à internet e às redes sociais, muitas vezes com aspetos subjetivos associados à perceção de incomodidade acústica.

Neste contexto, a monitorização contínua de ruído ambiental assume uma importância crescente, de modo a reduzir a incerteza e subjetividade inerentes ao processo de avaliação de ruído ambiental e de incomodidade acústica.

Nesta comunicação será apresentado o projeto MIRA (Monitorização Inteligente de Ruído Ambiental), que consistiu no desenvolvimento e teste de uma solução tecnológica para a monitorização contínua de ruído de atividades ruidosas, constituída por estações de monitorização de ruído, plataforma digital integrada e algoritmos de inteligência artificial. A solução permite, por um lado, um registo contínuo de parâmetros objetivos do ruído provenientes de até 100 estações de monitorização e, por outro, a identificação de padrões de variação do ruído, de eventos sonoros ou de ruídos anómalos ou particularmente incómodos (tonais, impulsivos), com capacidade de deteção e classificação em tempo quase-real.

### ABSTRACT

Noise is one of the main environmental risk factors for health. Exposure to environmental noise can lead to stress reactions, sleep disturbance, mental health and well-being, cognitive impairment in children, as well as negative effects on the cardiovascular and metabolic system. The World Health Organisation has identified noise as the second most significant environmental cause of disease in Western Europe, after air pollution.

It is in this context that MIRA project arises, as described in this paper, which lays on the development of an intelligent and seamless continuous monitoring system for environmental noise, with multiple configurations, consisting

\* *Autor de contacto:* [luis.conde@dbwave.pt](mailto:luis.conde@dbwave.pt)

*Copyright:* ©2023 Luís Conde Santos et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

of sound level meters and sensors (hardware) that communicate with a management, processing and data analysis platform (software), capable of detecting and classifying different types of sounds events, through the incorporation of *machine-learning* (ML), as well as notification of alerts and information for decision-making.

This system provides a service to different entities from the most diverse sectors, enabling them with the knowledge to comply with noise legislation and regulations. This solution gathers an innovative and highly differentiating characteristics from the actual products in the market.

*Key Words* — *audio classification, machine-learning, noise, sound-level-meter, monitoring, automation.*

### 1. INTRODUÇÃO

Os métodos tradicionais para avaliar o ruído ambiental, quer sejam a realização de medições de ruído mediante amostragens pontuais de curta duração com um sonómetro, operado no local por um técnico (pouco representativas), ou mesmo a elaboração de mapas de ruído (estáticos), revelam-se, já hoje, insuficientes para dar resposta às aspirações da população e às necessidades das entidades com responsabilidades nesta matéria. O problema decorre da enorme dinâmica do ruído num ambiente exterior, conjugada com o impacto imediato que um ruído excessivo ou incómodo pode ter nas pessoas a ele expostas.

No que respeita às medições de ruído “tradicionais”, um dos problemas práticos que surge é o facto de, devido à necessidade de presença física do técnico para realizar a medição e de se terem de recolher amostras em dois dias distintos e em três períodos de referência (diurno, entardecer e noturno), os laboratórios acreditados para este tipo de ensaios tenderem, dentro dos requisitos normativos, a realizar o menor número de amostragens, com a mínima duração possível e evitando as horas noturnas mais tardias.

No que respeita aos mapas de ruído, muito úteis para efeitos de planeamento e ordenamento do território e para orientar a elaboração de planos de ação de redução de ruído, estes baseiam-se em cálculos de valores médios de longa duração (1 ano) dos indicadores de ruído, sendo por isso

eminentemente estáticos e alheios à dinâmica temporal inerente ao ruído ambiental. São também por isso insuficientes para traduzir o impacto imediato que um ruído excessivo ou incómodo pode ter na população exposta.

De facto, num contexto de rápidas mudanças na sociedade e nas tecnologias ao seu dispor, os métodos e tecnologias utilizados na avaliação e gestão do ruído ambiental têm evoluído lentamente, não conseguindo, acompanhar as verdadeiras mudanças de paradigma que vêm ocorrendo à sua volta.

A dBwave, empresa de engenharia acústica do Grupo ISQ, procurando atender a este tipo de questões, desenvolveu e tem vindo a aplicar metodologias de monitorização de ruído em contínuo, por períodos mais ou menos alargados (desde 48h a várias semanas), com análise manual dos registos em pós-processamento, metodologias essa que começam atualmente a revelar-se insuficientes para as necessidades de uma franja crescente dos seus clientes mais exigentes. Além disso, este tipo de análises, feitas manualmente, consomem muito tempo e não permitem qualquer acesso remoto aos dados durante a monitorização, nem qualquer controlo em tempo real de ruídos anómalos e que só são identificados à posteriori.

A adicionar, a quantidade de marcas e modelos de sonómetros existentes no mercado obriga a utilização de vários *softwares* de monitorização e tratamento de dados, cada um específico de cada marca e modelo, e a análise e classificação sonora após monitorização torna-se laboriosa, com a elaboração de relatórios demorada e outros obstáculos relevantes no âmbito atual de medições de ruído.

## 2. MODELOS DE MACHINE LEARNING PARA CLASSIFICAÇÃO AUDIO

O desenvolvimento de modelos de IA para classificação de eventos sonoros, e sua consequente integração nos sistemas atuais, ainda se encontram numa fase bastante embrionária, com poucos *use cases* relevantes para a tecnologia em questão.

Devido à complexidade de análise a ficheiros áudio, por vezes mais difícil do que a análise de imagens ou vídeo, graças à sua estrutura complexa, não estruturada, muitas vezes com ruído branco, ruído de fundo, e necessidade de um elevado pré-processamento, traz grandes dificuldades à extração de informação útil do conjunto de dados original e à criação de modelos matemáticos capazes de modular o problema a estudar [5]. Igualmente, a falta de um grande volume de dados para treinar modelos capazes de identificar inúmeras fontes de ruído, torna esta questão ainda mais complexa. No entanto surgiram novos algoritmos para o efeitos de classificação, é um exemplo o modelo [3] capaz de identificar somente quatro classes diferentes: (1) porta de madeira a abrir e (2) fechar, (3) metal a cair e (4) o som de água a correr).

Recentemente, começaram a surgir *datasets* em grande escala, bem como modelos de ML mais complexos capazes de processar grandes volumes de informação. Estes avanços proporcionaram a investigação e desenvolvimento de algoritmos mais eficientes no reconhecimento de padrões sonoros. Uma das abordagens é a classificação de ficheiros áudio, convertidos em espectrogramas e usados como *input* para *Convolutional Neural Network (CNN)* [4]. No entanto, as *CNNs* foram inicialmente desenvolvidas para análise de imagens, que contêm informação estática no tempo, sendo que o espectrograma tem uma dimensão adicional, o intervalo temporal, tornando os dados sequenciais.

Mais recentemente, foi desenvolvido outro algoritmo [6] com o propósito de preencher a lacuna dos muitos dos trabalhos desenvolvidos anteriormente, *i.e.*, estes concentravam-se em tarefas particulares com um número limitado de classes, não ambicionando reconhecer uma vasta gama de fontes sonoras. Em 2017 a *Google* disponibilizou publicamente um *dataset* - *AudioSet* [2], que incorpora um grande volume de dados, composto por 1.9 milhões de ficheiros áudio, com 527 classes diferentes.

Tirando partido deste *dataset*, foi desenvolvido um modelo designado por *Pretrained Audio Neural Networks (PANNs)* [6]. Este modelo *PANNs* inclui uma camada de processamento com a transformada de *Fourier*, permitindo assim ao algoritmo aprender informação no domínio tempo-frequência. A informação no domínio tempo-frequência alimenta uma rede neuronal que permite a identificação de 527 padrões áudio (presentes no *dataset AudioSet*). Existem também outros modelos desenvolvidos utilizando este *dataset* [7] [9] [10]. Para a nossa aplicação e os requisitos definidos no projeto considerou-se a utilização do modelo *PANNs*, devido ao seu código público e vasta documentação.

## 3. O PROJECTO MIRA

Para responder aos desafios anteriormente descritos, a solução inovadora do projeto MIRA, idealizado pela dBwave.i, foca-se na monitorização de ruído em contínuo, desacompanhada, pretendendo capacitar todo o processo com automatismos desenvolvidos para identificação de sons ou ruídos específicos, bem como a geração de alertas e relatórios diários de medições realizadas. Sendo assim, o MIRA agrega diferentes módulos, todos interligados por uma arquitetura robusta e utilizados através de uma Plataforma Digital Integrada, que:

- É agnóstica em relação aos diferentes equipamentos (sonómetros), com integração rápida de novas marcas e/ou modelos;
- Armazena e trata os dados de centenas de equipamentos em tempo real;
- Monitoriza em contínuo e controla múltiplas métricas sonoras, assim como permite a sua visualização em tempo real;
- Gera alertas e relatórios de forma automática;

- Classifica rapidamente ruídos distintos através de métodos de IA;
- Pode alocar os equipamentos a múltiplos projetos, independentes entre si e com acessos personalizados, para cada um dos quais é exibida num mapa a localização instantânea de cada equipamento (GPS).

O sistema desenvolvido proporciona também:

- Substituir uma boa parte do trabalho laborioso do técnico;
- Reduzir o tempo e esforço na análise, aumentando a produtividade dos estudos de ruído ambiental;
- Caracterizar as diferentes origens sonoras e identificar em tempo útil se o ruído advém de fontes do ruído particular em análise (indústria, via rodoviária, construção, etc) e/ou com características mais incomodativas e prejudiciais para a saúde, que deverão merecer foco de atenção, ou se são sons naturais ou outros integrantes do ruído residual (animais, vento, chuva, etc);
- Gerar alertas por email ao responsável da instalação ruidosa em estudo de modo a permitir atuar de imediato na resolução do problema;
- Elaborar relatórios periódicos (diários, semanais, mensais) de forma automática;
- Acompanhar a evolução do enquadramento normativo e regulamentar em matéria de ruído ambiental que, em traços gerais, aponta novos objetivos e metodologias que acentuam ainda mais a necessidade de novas técnicas de monitorização.

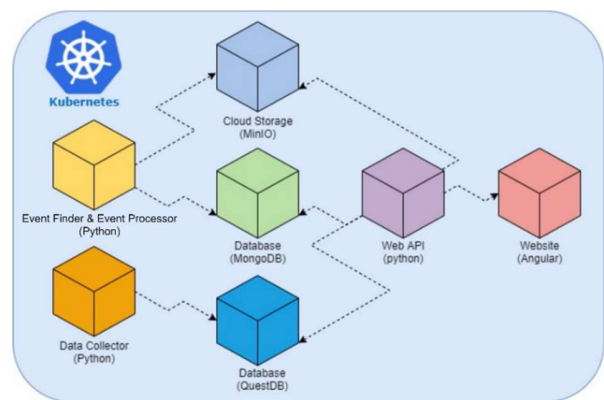
Assim, a solução apresentada permite, de modo muito significativo, melhorar tecnologicamente o método de avaliação e análise de ruídos, reduzindo os custos dessa análise e aumentando a sua abrangência, para além de permitir oferecer novos serviços e soluções, desde o controlo de atividades ruidosas em tempo real e registo continuado de evidências de *compliance* com os limites regulamentares. Contribui, assim, para a melhoria da qualidade de vida e da saúde da população, designadamente nas áreas urbanas e na vizinhança de áreas industriais ou de infraestruturas ruidosas, potencia o desenvolvimento de novos serviços e ajuda empresas e outras organizações a melhor gerir e minimizar os impactos acústicos da sua atividade, reduzindo o seu nível de risco, aumentando a previsibilidade e melhorando a sustentabilidade do seu negócio/atividade.

#### 4. PLATAFORMA DIGITAL INTEGRADA

A plataforma digital integrada permite integrar e centralizar diversas tecnologias já existentes, bem como novas desenvolvidas com características inovadoras e diferenciadoras da concorrência. Procura assim responder às necessidades identificadas no mercado, em setores como a indústria, construção, infraestruturas de transporte, municípios e espetáculos musicais e desportivos.

A plataforma MIRA é desenhada de forma a permitir flexibilidade, performance, escalabilidade e fácil utilização. O sistema está dividido em diferentes módulos com o principal foco na capacidade de escalar consoante o tráfego que estiver a ocorrer no momento. Para o efeito, é utilizado um serviço de *cloud* que permite flexibilidade das capacidades computacionais contratadas de acordo com a necessidade do momento. Este aspeto é crítico pois permite garantir uma grande segurança na qualidade e responsividade do serviço, mas também controlo dos custos associados, já que estes estão diretamente relacionados com o número de equipamentos/trabalhos a ser realizados em cada momento.

A Figura 1 ilustra os vários módulos existentes e como eles comunicam entre si.



**Figura 1:** Esquema da arquitetura do MIRA.

Destacam-se então os seguintes módulos:

- *Website*: proporciona um acesso seguro, interativo e acessível às funcionalidades da plataforma;
- *REST API*: responsável por servir informação aos vários componentes da plataforma, orquestrando respostas e processos resultantes de pedidos do utilizador;
- *Data Collector*: consiste num módulo isolado que se conecta aos equipamentos (sonómetros/estações de monitorização de ruído) recolhendo as diferentes métricas sonoras e armazenando-as numa base de dados de séries temporais;
- *Event Finder & Event Processor*: consiste num módulo isolado, sem interação com o exterior, que verifica, de forma cíclica em intervalos de tempo regulares pré-definidos, se existem ficheiros áudio no sonómetro, criados pela ativação de um *trigger* pré-definido (evento); Este módulo é também responsável pela identificação dos ruídos em cada ficheiro áudio (evento), através de um modelo de ML, sendo o seu *output* disponibilizado na plataforma;
- Bases de dados: A escrita e/ou leitura de dados é um processo contínuo e decorre em paralelo com todos os outros componentes da plataforma. Recorre-se então a três géneros de bases de dados, sendo a mais crítica apropriada para entradas indexadas no tempo.

- Outros módulos: A plataforma está ainda munida de outras funcionalidades que trabalham em paralelo permitindo dar uma melhor resposta às necessidades do utilizador e que se despoletam por pedidos à *REST API*;
  - Alertas: um alerta é gerado após a ativação de um *trigger* pré-definido na plataforma, onde o utilizador receberá a informação através de sms e/ou correio eletrónico;
  - Relatórios: Os utilizadores podem gerar relatórios referentes a a um dado período de análise;
  - Exportar dados: Os utilizadores podem descarregar os dados recolhidos pela plataforma referentes a um período de análise.

#### 4.1 Aquisição de dados em tempo real

Sendo um dos objetivos do projeto a recolha, processamento e visualização de dados em tempo real, a cada segundo (capaz de detetar sons impulsivos de curta duração), o módulo *data collector* anteriormente referido, torna-se parte essencial da arquitetura do sistema.

Este módulo depende intrinsecamente da base de dados, que necessita de ser indexada ao tempo, permitindo a inserção de um grande volume de dados por segundo, bem como rápida velocidade de leitura.

É possível recolher múltiplas métricas e espectros a cada segundo, com possibilidade de configuração pelo utilizador, dentro dos parâmetros e indicadores disponíveis em cada equipamento, por exemplo:

- Métricas globais: *LAF*, *LAI*, *LAFeq*, *LAIeq*, *LAFMax*, *LAFMin*, *LAIMax*, *LAIMin*, *LCpeak*, *LAI3*;
- Espectros: *LAF*, *LCF*, *LZF*, *LAFeq*, *LCFeq*, *LCZeq*

Dada a quantidade de métricas globais e espectros recolhidos e a frequência de amostragem necessária, uma vez por segundo, é então realmente crítico a escolha da base de dados.

Considere-se um exemplo onde existe: uma taxa de amostragem de 1 segundo, 100 equipamentos, 10 métricas (por equipamento) e 8 espectros com 34 bandas de frequência (por equipamento). Isto totaliza ~1.69 milhões de valores escritos por minuto e 101 milhões por hora.

Existindo diversas soluções presentes no mercado, otimizadas para armazenamento e leitura de dados baseados no tempo a escolha recaiu sobre uma base de dados pública e bem documentada, com alto desempenho de escrita, capacidade analítica fácil recorrendo à linguagem SQL, bem como simplicidade de integração com outros módulos do sistema.

Dado que a inserção das múltiplas leituras na base de dados não deve ser apenas baseada na sua própria capacidade de escrita, foi utilizado adicionalmente um sistema de filas de espera extremamente rápido e escalável, que desacopla os fluxos de informação, resultando numa baixa latência, ou seja, num pequeno tempo de processamento de dados simultâneos através de uma conexão de internet, garantindo-

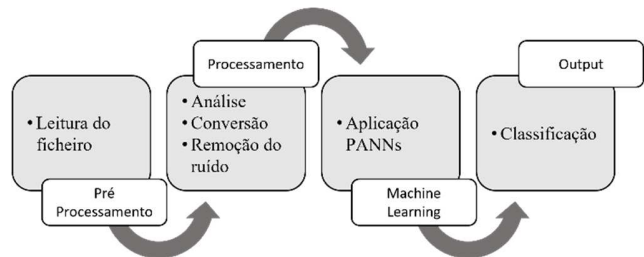
se assim a capacidade de receber dados segundo a segundo de cada equipamento.

As soluções acima descritas, aliadas a tecnologia pública que permite gestão automática, alta disponibilidade e escalabilidade de micro serviços conforme necessário, garante a recolha e armazenamento dos diferentes valores das métricas e espectros a cada segundo, com uma elevada confiança, ajustando e escalando o sistema ao número de equipamentos, trabalhos e tráfego em curso.

O objetivo é então garantir que as amostras são sempre recolhidas com a cadência de 1Hz e, caso a latência aumente para valores maiores ou iguais a um segundo, o número de recursos computacionais escala automaticamente, sendo que estes diminuem caso o contrário se verifique.

#### 4.2. Event processor - identificação de audios

Um dos aspetos mais críticos e inovadores para o sucesso da plataforma MIRA, é a identificação e classificação de forma automática da fonte sonora, de forma a permitir uma análise mais minuciosa e detalhada de um determinado ambiente. Para tal, foi desenvolvida e implementada uma metodologia, Figura 2, capaz de receber e processar um áudio e utilizar modelos de ML para detalhar/classificar com precisão a origem da maioria dos sons medidos.



**Figura 2:** Metodologia para recolha e processamento de áudios.

Para esta componente de desenvolvimento foram estabelecidos alguns requisitos e pressupostos gerais, tais como:

- Alta precisão nas classificações sonoras;
- Grande quantidade de classes que podem ser identificadas;
- Escalabilidade;
- Possibilidade de melhorias nas classificações;
- Capacidade de adicionar novas classes de sons a identificar;
- Rapidez de resposta.

Um dos requisitos definidos foi a necessidade de existir um treino personalizado do modelo de ML para novas classes que não estejam presentes no *dataset* original. Ruídos específicos da indústria, ou de um ambiente novo, para o qual o modelo não tenha referência exigem um estudo e classificação prévia do ruído local, de forma a treinar o modelo na classificação destas novas classes.

A fim de solucionar esta restrição foi aplicada a técnica designada por *transfer learning* [8] que permite adaptar um modelo já previamente treinado, *i.e.* com algum conhecimento do domínio do problema, para um novo conjunto de dados ou novas classes. Quando esta técnica é utilizada, o conhecimento adquirido anteriormente é reaproveitado, facilitando a adaptação ao novo desafio e, consequentemente, havendo assim uma transferência de conhecimento e uma aprendizagem mais rápida e requerendo muito menos dados de treino.

Deste modo, e como referido anteriormente na secção 2, o modelo de ML recaiu sobre a rede neuronal *PANNs* (2020) [6], treinado e otimizado para reconhecimento de múltiplos padrões em ficheiros áudio e que também permite a utilização de *transfer learning*.

## 5. RESULTADOS

Os resultados aqui apresentados foram obtidos pela plataforma num caso de estudo em que se instalaram 3 estações de monitorização, que mediram em contínuo durante vários meses junto aos limites de uma instalação fabril.

Os sonómetros foram colocados no terreno e conectados por 4G à plataforma digital integrada MIRA. Os valores de ativação dos *triggers* para iniciar a gravação de áudio foram definidos em cada sonómetro, e os valores para geração de alertas (email e/ou sms) foram definidos na plataforma digital integrada. Através desta deu-se início à tarefa de avaliação e monitorização de ruído ambiental da qual provieram os resultados apresentados em seguida.

Ao iniciar uma monitorização, pode-se visualizar o gráfico das diferentes métricas e espectros em *real time* assim como os eventos desencadeados através do *trigger* configurado no sonómetro, como ilustrado na Figura 3.

Pode ainda ser visualizado o histórico de um trabalho, onde é acessível informação como: Gráficos históricos das métricas e dos espectros, listagem de eventos e sua classificação, listagem de alertas emitidos, entre outros, conforme ilustrado na Figura 4.



Figura 3: Gráfico de métricas e eventos em tempo real.

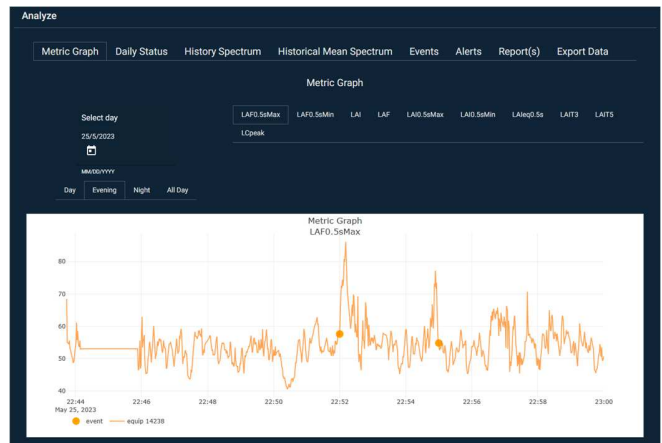


Figura 4: Gráfico de histórico das métricas e eventos.

Informação relativa aos eventos desencadeados pelos *triggers* é acessível no separador *Events* onde são disponibilizadas três versões do ficheiro áudio recolhido pelo sonómetro, (1) com e (2) sem redução de ruído de fundo e (3) ficheiro original, assim como as classes identificadas no processamento do ficheiro conforme ilustrado na Figura 5. É também possível pesquisar eventos por intervalo temporal assim como por tipo de ruído detetado, *e.g.* martelo pneumático ou vento.



Figura 5: Visualização do histórico de eventos desencadeados ao longo da execução de uma monitorização.

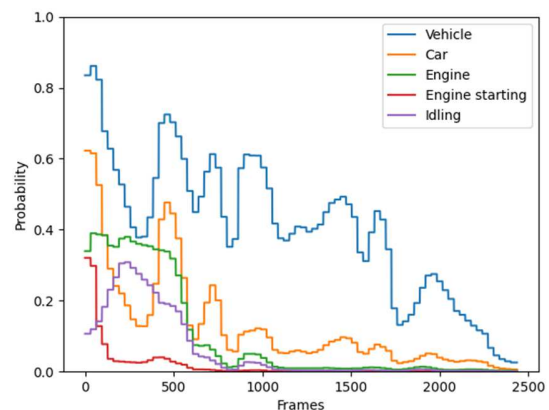
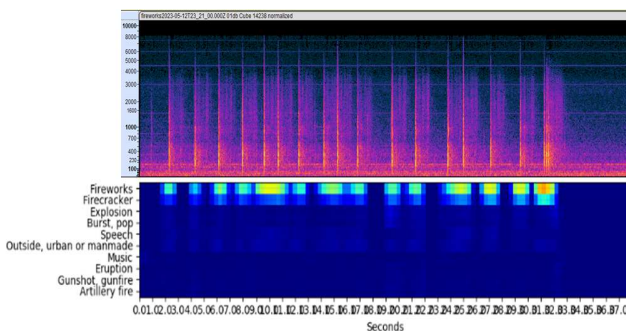


Figura 6: Representação temporal do output do modelo de ML.

Foram analisadas as deteções da origem de ruído dos excertos de áudio recolhidos pelos sonómetros aquando da ativação dos *triggers*. O resultado do processamento do modelo de ML é representado sobre a forma probabilística do ruído corresponder a uma determinada classe conhecida pelo

modelo. Para cada instante (*frame*) de ruído é executada esta avaliação resultando em um grau de confiança/probabilidade ao longo do tempo para cada uma das possíveis classes, como demonstrado na Figura 6, onde se observa um carro a ser ligado nos primeiros *frames* e de seguida a afastar-se do sonómetro.

A avaliação dos resultados do modelo de ML comparada com os dados recolhidos pelo sonómetro mostra uma grande correspondência entre as classes detetadas e o ruído medido. Um exemplo destes resultados que é visualmente observável, é o caso da deteção de Fogo de Artifício que ocorreu na execução do trabalho, que pode ser observado na Figura 7.



**Figura 7** - Comparação da representação do espectrograma dado como *input* do modelo (em cima) e da deteção de classes do modelo (em baixo).

## 6. CONCLUSÃO

O projeto da plataforma MIRA, descrita resumidamente neste artigo, teve como foco o desenvolvimento de uma solução digital para monitorização contínua de ruído de atividades ruidosas, envolvendo a capacidade de instalar uma rede de monitorização que permita evidenciar registos de monitorização em contínuo de ruído, com potencial para identificação e classificação de diferentes ruídos e eventos sonoros, bem como gerar alertas sobre a ocorrência de ruído excessivos, incomodativos ou fora do “normal”, para atuação corretiva rápida.

A plataforma inclui diversas funcionalidades inovadoras, destacando-se a capacidade de reconhecimento de diferentes tipos de ruído (mais de 527 tipos diferentes atualmente). É também capaz de ser adaptada e customizada para diferentes envolventes com fontes de som atípicas e mais específicas de cada ambiente. É desta forma capaz de se adaptar a um leque muito variado de âmbitos graças a técnicas de *transfer learning*.

O sistema foi também desenvolvido com vista a escalabilidade e capacidade continua de adaptação ao número de sonómetros em operação, permitindo que a monitorização em tempo real ocorra de segundo a segundo capturando métricas, espectros, bem como a sua associação aos eventos detetados (ficheiros áudio) e sua respetiva classificação via

técnicas de ML. A plataforma disponibilizada pela dBwave.i é agnóstica a marcas e modelos de sonómetros, aceitando dados de um largo conjunto de sonómetros que tenham capacidade de transmitir informação via web.

**Agradecimentos:** Este trabalho foi realizado com o apoio financeiro do Programa Operacional Temático Competitividade e Internacionalização (COMPETE 2020), cofinanciado pelo Fundo Europeu de Desenvolvimento Regional (FEDER), no âmbito do PT2020 - MIRA - Nº 070283.

## 7. REFERÊNCIAS

- [1] World Health Organization. *Burden of disease from environmental noise - Quantification of healthy life years lost in Europe*. WHO Regional Office for Europe, 2011.
- [2] Jort F. Gemmeke, Daniel P. W. Ellis, Dylan Freedman, Aren Jansen, Wade Lawrence, R. Channing Moore, Manoj Plakal, and Marvin Ritter. Audioset: An ontology and human-labeled dataset for audio events. In Proc. IEEE ICASSP 2017, New Orleans, LA, 2017.
- [3] J.P. Woodard. Modeling and classification of natural sounds by product code hidden markov models. *IEEE Transactions on Signal Processing*, 40(7):1833–1835, 1992.
- [4] Keunwoo Choi, György Fazekas, and Mark B. Sandler. Automatic tagging using deep convolutional neural networks. In International Society for Music Information Retrieval Conference, 2016.
- [5] Matt Benatan and Kia Ng. Cross-covariance-based features for speech classification in film audio. *Journal of Visual Languages Computing*, 31:215–221, 2015. Special Issue on DMS2015.
- [6] Qiuqiang Kong, Yin Cao, Turab Iqbal, Yuxuan Wang, Wenwu Wang, and Mark D. Plumbley. Panns: Large-scale pretrained audio neural networks for audio pattern recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28:2880–2894, 2020.
- [7] Ramírez, J., Flores, M.J. Machine learning for music genre: multifaceted review and experimentation with audioset. *J Intell Inf Syst* 55, 469–499, 2020.
- [8] Stevo. Bozinovski and Ante Fulgosi. "The influence of pattern similarity and transfer learning upon the training of a base perceptron B2." (original in Croatian) *Proceedings of Symposium Informatica* 3-121-5, 1976.
- [9] Victor Guedes, Felipe Teixeira, Alessa Oliveira, Joana Fernandes, Leticia Silva, Arnaldo Junior, João Paulo Teixeira, Transfer Learning with AudioSet to Voice Pathologies Identification in Continuous Speech, *Procedia Computer Science*, Volume 164, pp. 662-669, 2019.
- [10] Y. Yu, S. Beuret, D. Zeng and K. Oyama, "Deep Learning of Human Perception in Audio Event Classification," 2018 IEEE International Symposium on Multimedia (ISM), Taichung, Taiwan, pp. 188-189, 2018.