inter·noise 2023
CHIBA,GREATER TOKYO 20-23 AUGUST

AI-technology for efficient noise monitoring and analysis in complex urban soundscapes

Karl Henrik Ejdfors[1]
Norsonic AS
Gunnersbråtan 2, N-3409 Tranby, Norway

ABSTRACT

This paper presents a user-friendly, project-based noise monitoring and data analysis software called NorCloud. This software has been developed to exploit the advantages of noise monitoring terminals that allow localizing dominant sound sources in complex urban soundscapes. NorCloud can be used with multiple microphones simultaneously and has data analysis features that enable identification, classification and extraction of sound events. The efficiency of these processes is greatly increased by using AI-technology for automatic removal of non-relevant sound sources. In the article, the advantages and limitations of using automatic noise monitoring and analysis systems are explained. Current applications and possibilities for further development are also presented.

1. INTRODUCTION

Noise pollution has become a growing concern in modern societies and across Europe [1]. Urban soundscapes are complex, and can be shaped by factors such as construction work, industrial activity and traffic. These activities can have a significant negative impact on resident's and worker's health and quality of life. Excessive noise have proven effects on health, including sleep disruptions, hypertension, heart disease, and hearing loss [2]. In addition, the economic impact of noise is also significant. As of 2012, World Health Organization estimated that at least one million healthy life-years in Western Europe were being lost annually to environmental noise [3].

Sound monitoring for noise control can mitigate the negative effects of sound pollution. It can provide the possibility to identify sources of excessive noise, and a way to evaluate the effectiveness of noise control measures. Combined with automatic identification of sounds using advanced machine learning, it is possible to distinguish between different noise pollution sources. This enables more accurate identification of problematic sources of noise, and more targeted mitigation measures can be taken.

Norsonic has for over 50 years developed high precision and quality sound instrumentation equipment, and their products are among the most advanced within sound monitoring. Their cloud based sound and vibration monitoring system NorCloud is a powerful software for storing and analyzing sound data over time. Since 2019, Norsonic have developed a new machine learning system, called NoiseTag, together with one of Europe's largest independent research organisations, SINTEF, and Norwegian University of Science and Technology for automatic recognition of complex sounds in urban soundscapes. This paper presents a user-friendly and project based approach for efficient noise monitoring with NorCloud and explaines how artificial intellegence (AI) in NoiseTag is utilized to improve soundscape analysis.

[1]kejdfors@norsonic.com

2. DESCRIPTION OF THE SYSTEM

NorCloud is a cloud-based monitoring system developed by Norsonic that allows for remote monitoring and analysis of sound and vibration data both in real-time and non-real-time. The system is designed to work with Norsonic's sound level meters and vibration sensors and offers a user-friendly platform for data collection, storage, and analysis. NorCloud enables users to access sound and vibration data from any location with an internet connection, allowing for efficient monitoring and control of noise and vibration levels in a range of contexts.

One of the key features of NorCloud is its scalability. It can accommodate large numbers of sound level meters and vibration sensors, making it ideal for use in large-scale industrial or construction projects. The platform can be customized to meet the specific needs of each project, allowing for efficient and effective noise and vibration monitoring and control. In addition, NorCloud offers a variety of visualization tools, including interactive maps and graphs, which allow users to easily analyze and interpret the data collected by the sensors. This can help to identify specific sources of noise pollution.

Another benefit of NorCloud is its ability to provide advanced alarm functionalities that enable users to create custom triggering thresholds for specific noise and vibration functions and receive alerts when those triggers are activated. This feature is particularly useful in environments where noise and vibration levels must be strictly controlled, such as in construction sites, harbors, or motorsport arenas.

The triggering functionality in NorCloud allows users to combine different sound level profiles and functions for frequencies ranging from 6.3Hz to 40kHz, and sound directivity with the assistance of NoiseCompass. This feature is made possible through the connection of NoiseCompass to sound level meters, which facilitates sound directivity by collecting the most prominent sound in 3D from eight MEMS microphones. By leveraging this capability, the triggering functionality enables efficient event generation by selectively generating events from the sectors of choice. This process is described as a flow chart in Figure 1. This feature is particularly valuable for users seeking to efficiently monitor and analyze sound data in specific areas of interest. By limiting the scope of event generation to a designated sector, the triggering functionality reduces the volume of irrelevant data that users would need to sort through, enabling them to focus on more relevant data.

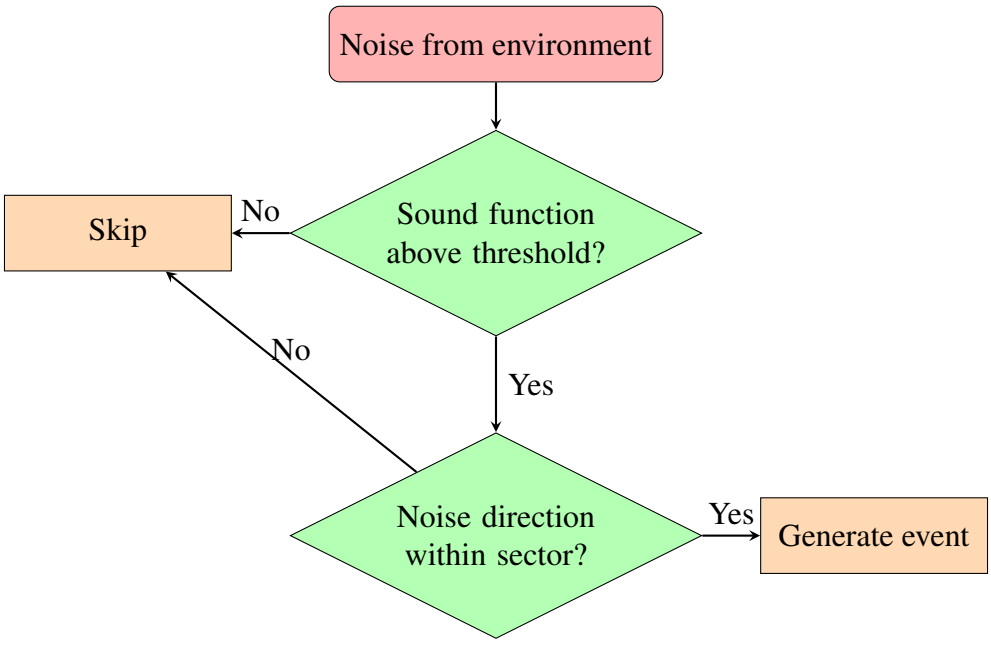

Figure 1: Event generation flow chart

Despite the ability of NorCloud's triggering functionality, there is a possibility of some sounds not being relevant to the project within that sector. The advanced sound level meters have the capability to record sound from the surrounding environment, which can be attached to the events generated by the triggers. This additional data can be valuable in providing context to the generated events and identifying the sources. However, the processing of this additional data can require a significant amount of manual labor and associated costs. In order to distinguish the sound sources of interest, it is essential to have a team of experts to listen to and analyze the recordings. This process can be time-consuming and labor-intensive, as the experts must carefully review each recording to identify any relevant information. Additionally, the experts must be highly skilled and trained in sound analysis to accurately interpret the data.

The AI-based system called NoiseTag can further enhance the analysis by automatically recognizing and tagging sounds based on the recorded data. The system utilizes machine learning algorithms to analyze the sound data and identify specific sounds, such as traffic noise, construction noise, or bird song. By leveraging this technology, users can gain more granular insights into the sources and patterns of sound pollution, and decrease costs and time spent on manually listening to the recordings. With NoiseTag, the users are able to exclude irrelevant sound data in the event list and from calculations such as e.g. Lden [4]. Lden, which stands for "day-evening-night" level, is a common noise metric used to measure sound exposure over a 24-hour period, taking into account the different levels of background noise during the day, evening, and night. By excluding non-relevant sounds, the Lden calculation can provide a more accurate representation of the overall sound exposure in a given area in space, and enhances the quality and accuracy of the analysis. This capability is essential in ensuring that the sound analysis is comprehensive, precise, and free from the potential biases that could arise from the inclusion of irrelevant sounds.

With NorCloud's built-in report generator, users have the ability to customize their reports using drag and drop functionality to include relevant sound and vibration data in the form of tables and graphs. This feature enables users to create reports that are tailored to their specific needs and requirements, providing them with a more comprehensive and detailed analysis of the collected data. Additionally, the ability to customize reports in this manner facilitates more effective communication of findings to stakeholders and decision-makers.

3. SOUND PROCESSING

Upon receipt of recordings from NorCloud, NoiseTag employs its advanced machine learning algorithms that leverage new techniques in the field of deep learning for data augmentation and sound classification. Specifically, NoiseTag utilizes convolutional neural networks (CNNs) to perform sound classification. By using this cutting-edge technology, NoiseTag is able to accurately and efficiently identify and tag specific sounds [5].

In recent years, CNN architecture has been increasingly employed for sound classification, even though it was originally developed for image classification. Compared to traditional machine learning classification techniques, the CNN approach has been shown to achieve better accuracy and robustness [6]. However, due to the limited amount of high-quality labeled data from NorCloud, additional techniques such as transfer learning and data augmentation have been employed. Transfer learning allows for the extraction of features learned from an existing neural network, typically a sophisticated deep network trained on a large dataset, and transferring these features into the current network. This approach allows for efficient training with limited amounts of data, without sacrificing accuracy [7].

Data augmentation is another technique used in the sound classification, whereby the label remains the same even when deformations are applied to data. This technique has been widely acknowledged in deep learning as an effective way to increase additional training data and to improve performance [8].

NoiseTag's cutting-edge sound recognition system is built on these machine learning techniques

and powers its sophisticated API. The NoiseTag API receives recordings from NorCloud, processes it, and returns the label, allowing for quick identification of specific sounds. This process is shown as a flow chart in Figure 2.

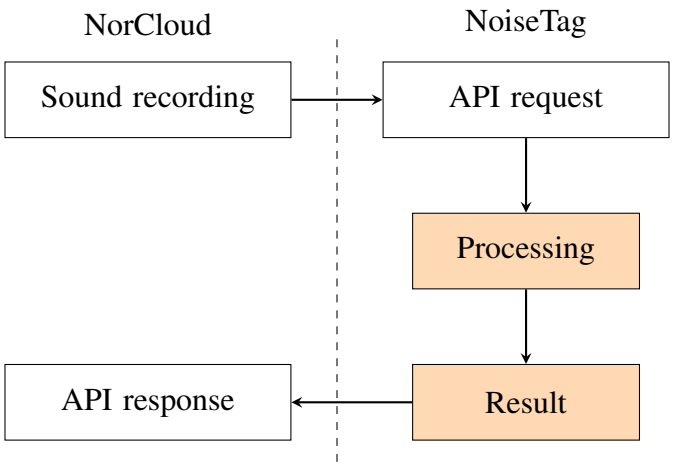

Figure 2: Sound processing flow chart

4. LIMITATIONS OF THE SYSTEM

NorCloud is capable of collecting data from various instruments and can be setup to generate events based on different function thresholds. Nonetheless, in complex soundscapes, there is a significant amount of non-relevant data that can potentially interfere with intended events and calculations. To address this issue, NoiseCompass can be introduced, which enables localization of sound sources and consequently, the exclusion of extraneous data from events and calculations.

It is noteworthy that NoiseCompass is limited to identifying the most prominent direction of sound. This yields high-quality results when there is a single dominant sound source present in the environment. However, in the presence of multiple dominant sources, the accuracy of the output may be reduced. Furthermore, the performance of NoiseCompass may be subject to limitations when it is positioned close to reflective surfaces. In NorCloud, users have the option to specify an acceptance threshold for the quality of directional output provided by NoiseCompass.

NoiseTag is an innovative AI service that is designed to extract meaningful information from sound recordings. Its primary function is to accurately label these recordings by identifying their unique acoustic characteristics. The successful operation of the model is dependent upon a robust dataset of labeled audio data. Through testing conducted in on-site locations, NoiseTag has demonstrated the potential to significantly mitigate uncertainties from irrelevant events and calculations. Nonetheless, it is important to acknowledge that NoiseTag, like other machine learning models, is not error-free and may produce erroneous outcomes in the form of false positives and false negatives. Such inaccuracies may lead to automatic exclusion of events and calculations that are erroneously tagged, requiring the execution of manual checks to rectify the errors.

As the machine learning model continues to evolve, it is important to provide it with additional labeled data. This will enable NoiseTag to enhance its accuracy and precision in identifying specific sounds, as well as expand its range of capabilities. By leveraging a large and diverse collection of labeled audio data, NoiseTag can be further optimized to classify sounds.

Therefore, it is important to access labeled audio datasets that can be used to train NoiseTag. Such datasets should encompass a broad spectrum of audio content from the broad spectrum of projects in NorCloud. To ensure the accuracy of the model, selected experts will be responsible for correctly labeling any reported mislabeled recordings. By doing so, the model can be enriched

with a wide range of labeled audio data, thereby enhancing its generalizability and adaptability to different contexts and scenarios.

5. CURRENT APPLICATIONS

NorCloud has established itself as a leading cloud-based platform for sound and vibration. Its applications span across diverse fields, including environmental noise monitoring, industrial noise control, and harbour noise control. Among these, environmental noise monitoring is perhaps one of the most prominent areas where NorCloud has proved to be a valuable asset.

The construction of the new hospital in Drammen, Norway, which began in 2019, has provided a compelling use case for NorCloud's environmental noise monitoring capabilities. Norsonic has stationed a monitoring system on site to capture and analyze the acoustic data generated by the construction work. The data collected by the monitoring system is a documentation of the changes in the soundscape throughout the construction process. An example of the data analyzation in NorCloud is shown in Figure 3.

Figure 3: Analyzation of sound data in NorCloud from the construction site in Drammen.

The soundscape of the construction site is inherently complex, as it is characterized by a multitude of sound sources that can make it challanging for the entrepreneur to determine their daily noise limit. Furthermore, the presence of a train track adjacent to the site, where trains periodically sound their horns, and a nearby scrap dealer further compound the acoustic environment. In such scenarios, the NoiseCompass serves as a critical tool for the entrepreneur to effectively discern between their intended noise and unwanted sounds from external sources such as the train and scrap dealer, enabling them to take appropriate measures to control noise levels and comply with regulatory requirements.

The construction work generates acoustic data from various events that have been utilized to train the NoiseTag machine learning model. Furthermore, recordings from the nearby train and scrap dealer have also been integrated into the training dataset to increase its diversity. This strategy has enabled NoiseTag to learn from a wide range of acoustic content. Moreover, the NoiseCompass has been employed to enhance the interpretation of the recorded soundscape, simplifying the manual annotation of sounds during the NoiseTag training phase.

Another use case for NorCloud and NoiseCompass involves monitoring the soundscape of the

streets in Paris. The primary objective of this project is to assist a construction site situated in a vibrant entertainment area, featuring numerous bars and restaurants, in distinguishing the external noise from the noise generated by the construction work. This application has demonstrated the efficacy of the system in identifying the dominant sound sources in an urban environment. The diversity of sounds in urban areas can make it challenging to identify the most significant sources of noise, which can hinder effective noise reduction and mitigation efforts.

6. FURTHER DEVELOPMENT

NorCloud is constantly evolving to satisfy the requirements of its customers. In order to further enhance its capabilities, NorCloud has been subject to ongoing development initiatives. One of the latest developments is the release of the NoiseTag service in beta version. This innovative feature will be made available to selected customers in the near future, thereby providing them with the opportunity to experience this cutting-edge technology firsthand.

The beta version of the NoiseTag service is expected to offer a range of benefits to early adopters. Selected customers will be able to test the service in real-world scenarios and provide valuable feedback to NorCloud's development team. This feedback will be used to refine and improve the service before its full release to the wider market.

In addition to the release of the beta version, NorCloud's development roadmap includes a focus on research related to unsupervised model training for NoiseTag. This area of research is expected to play a major role in the future development of NorCloud, as it has the potential to significantly improve the efficiency of the monitoring system.

Unsupervised model training involves the use of advanced algorithms to enable the system to learn and adapt to new data, without the need for supervision or intervention from human operators. This approach has the potential to dramatically reduce the time and resources required to train NoiseTags's machine learning models.

7. CONCLUSION

The issue of noise pollution is a growing concern in modern societies, and monitoring noise is a step towards detecting and mitigating its negative effects. Norsonic, a company with over 50 years of experience in developing high-quality sound instrumentation equipment, has created NorCloud, a powerful cloud-based sound and vibration monitoring system for storing and analyzing data. One of the key features of NorCloud is its advanced alarm functionalities, which enable users to set triggering thresholds for specific noise and vibration functions, and combine them with sectors of interest with NoiseCompass.

However, despite these advanced triggering functions, it is still possible for some sounds to be irrelevant to the project being monitored. This is where the newly developed AI-based system, NoiseTag, can be particularly useful. By automatically recognizing and tagging sounds based on recorded data, NoiseTag can significantly enhance monitoring and reporting efficiency.

Overall, NorCloud, NoiseCompass and NoiseTag are important tools to analyze and understand noise pollution. These technologies offer an efficient and effective means of monitoring and analyzing sound and identifying sources of noise. As the issue of noise pollution continues to grow, the development of new systems like NorCloud, NoiseCompass, and NoiseTag will be critical in better understanding and mitigating the negative impacts of excessive noise.

REFERENCES

1. European Environment Agency. Environmental noise in Europe, October 2020.
2. Hiral J. Jariwala, Huma S. Syed, Minarva J. Pandya, and Yogesh M. Gajera. Noise pollution & human health: a review. *Indoor Built Environ*, 1(1):1–4, 2017.
3. World Health Organization. Regional Office for Europe. *Burden of disease from environmental noise: quantification of healthy life years lost in Europe*. World Health Organization, 2011.
4. European Environment Agency. Lden, 2001.
5. Karen B. Evensen, Rolf T. Randeberg, and Herold Olsen. Noisetag report h1: classification. Unpublished internal document, 2021. SINTEF.
6. Shawn Hershey, Sourish Chaudhuri, Daniel P. W. Ellis, Jort F. Gemmeke, Aren Jansen, R. Channing Moore, Manoj Plakal, Devin Platt, Rif A. Saurous, Bryan Seybold, Malcolm Slaney, Ron J. Weiss, and Kevin W. Wilson. CNN architectures for large-scale audio classification. *CoRR*, abs/1609.09430, 2016.
7. Izzi Yi Liu. Noisetag technical report at m2. Unpublished internal document, 2020. SINTEF.
8. Olusola O. Abayomi-Alli, Robertas Damaševičius, Atika Qazi, Mariam Adedoyin-Olowe, and Sanjay Misra. Data augmentation and deep learning methods in sound classification: A systematic review. *Electronics*, 11(22), 2022.