

Pattern recognition and neural networks for acoustic monitoring and conservation of the historic port of Ancona

Samantha Di Loreto^{1,*}, Alessandro Ricciutelli¹, Valter Lori², Fabio Serpilli² and Sergio Montelpare¹

¹ Università degli studi G. D'Annunzio, Viale Pindaro 42, 60123, Pescara (Italy)

² Università Politecnica delle Marche, Via Brece Bianche 12, 60124, Ancona (Italy)

Abstract

The soundscape of Ancona's historic port is a vital aspect of its cultural heritage, capturing the essence of its maritime and urban environment. This research focuses on the application of advanced sound event recognition techniques to characterize and preserve the unique auditory landscape of the port. Utilizing convolutional neural networks (CNNs) we developed an automated system capable of identifying and classifying sound events within the port's diverse acoustic environment.

Our approach involves the extraction of significant "soundmarks" - unique sounds that define the auditory identity of the port, such as ship horns, dockyard activity, and traditional marketplace sounds. By capturing and analyzing these soundmarks, our system not only characterizes the existing soundscape but also aids in the preservation of these auditory elements, ensuring that they are protected in the event of environmental changes or disasters.

The findings of this study underscore the importance of soundscape preservation as a critical component of heritage conservation, particularly for historic sites where auditory elements play a significant role in cultural identity. Our methodology provides a framework for the automatic extraction and classification of soundmarks, offering a valuable tool for heritage conservationists and urban planners in safeguarding the acoustic heritage of historical buildings and sites globally.

Keywords

Soundscape, Cultural Heritage, Pattern recognition, Neural network model

1. Introduction

The soundscape strategy, which encompasses both desired and undesired sounds while considering the contextual perception of the acoustic environment, represents a cutting-edge approach in integrating soundscapes as a central element in architectural design practices[1]. In this context, the term "soundscape" refers to the array of sounds present in a specific environment, including both wanted and unwanted sounds. Soundscape-based design transcends the purely visual aspect, considering the auditory experience of the environment and actively seeking to shape the acoustics of designed spaces [2].

VIPERC2024: 3rd International Conference on Visual Pattern Extraction and Recognition for Cultural Heritage Understanding, 1 September 2024

* Corresponding author.

✉ samantha.diloreto@unich.it (S. Di Loreto);

🆔 0000-0002-1901-4280 (S. Di Loreto).



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

This perspective suggests that for a more holistic architectural practice, it is essential to consider the soundscape as a design element, integrating it into design decisions to create more harmonious environments that are mindful of the users' overall sensory experiences [3]. To study the influence of the interaction between the geographic landscape and acoustic stimuli on environmental perception, various authors have attempted to identify the informative, aesthetic, or affective qualities of sound, which substantially contribute to the quality of a given landscape [4].

Landscape factors, which are not negligible, have also been considered in many studies, especially in relation to sound perception and soundscape [5]. The soundscape is typically interpreted through the identification and description of different sound sources in a location. Kang et al. [6] demonstrated that the human evaluation of loudness and acoustic comfort depends on a series of intrinsic environmental factors. For instance, a large-scale subjective survey was conducted on the underground commercial streets of Harbin, China, to determine how individual sound sources affected loudness and the assessment of acoustic comfort [7].

To study the influence of the interaction between geographic landscape and acoustic stimuli on environmental perception, various authors have attempted to identify the informative, aesthetic, or affective qualities of sound, which substantially contribute to the quality of a given landscape [8], [9]. In an urban environment, there are different zones, and in each zone, a dominant sound persists. Based on a series of case studies in Europe and China and an extensive literature review, Xu and Wu [10] evaluated the basic elements of the soundscape: sound, space, people, and environment. The evaluation of the relationship between the acoustic/auditory environment and the responses of the people inhabiting it is a crucial correlation for characterizing the human perception of daily environmental noise.

In [11], a system that, by means of a tangible user interface, integrated by pattern recognition and computer vision techniques, supports cultural heritage experts in creating Smart Interactive Experiences by properly tailoring the behavior of the involved smart objects. An experimental evaluation of the used techniques has been performed and is presented and discussed.

The contribution of this work lies in the development of a pattern recognition algorithm that systematically identifies and extracts soundmarks from the historic port's acoustic data. These soundmarks are crucial for characterizing the sonic footprint of the location. In the unfortunate event of the historical port's loss, our methodology provides a framework for not only structural reconstruction but also for recreating its soundscape. This ensures that the cultural and auditory identity of the port can be preserved and restored, maintaining a holistic approach to heritage conservation.

Our methodology provides a framework for the automatic extraction and classification of soundmarks, offering a valuable tool for heritage conservationists and urban planners in safeguarding the acoustic heritage of historical buildings and sites globally.

This strategy reflects a holistic approach that recognizes the fundamental role of sound in the perception of space and the overall quality of the built environment.

2. Material and methods

2.1. The case study

The city of Ancona is characterized by a very varied territory on which numerous contexts arise, including the port.

The transformation of the waterfront is an important topical issue that concerns not only large urban centers, but also small and medium-sized waterfront towns.

The port of Ancona (*international abbreviation IT AOI*) is a port located in the innermost part of the Gulf of Ancona and is therefore in its oldest core a natural harbour; it is Italy's leading port for international vehicle and passenger traffic.

The survey was carried out mainly in the historic Port, the only area of the port of Ancona that can be reached without the aid of means of transport and the oldest area of considerable historical interest due to the architectural works present. The historic port of Ancona is located on the north-east of the entire port and it is one of the oldest and most significant in the Adriatic Sea, with origins tracing back to ancient times. Established by the Greeks from Syracuse in the 4th century BC, Ancona's port is strategically situated along the maritime routes that linked the eastern Mediterranean with Italy and central Europe.

During the Roman era, Ancona emerged as a crucial port due to its prime location and the shape of its natural harbor, which offered safe refuge for ships. The port experienced one of its most significant expansions under Emperor Trajan, who built a new pier, now known as the Arch of Trajan, a monument honoring the emperor and his contributions to the city. In the Middle Ages, Ancona's port continued to thrive as a key port in the Adriatic, benefiting from its autonomy as a maritime republic, though smaller than Venice and Genoa. During the Renaissance, Ancona became a vital trade hub with the East, owing to its connections with the Ottoman Empire and other Mediterranean powers.

To this day the site is an important landmark for the city. It attracts tourists who are captivated by its history, monuments, and maritime traditions. Regular events and activities related to the sea and seafaring help preserve the city's rich maritime heritage and it function as a community hub, providing a beautiful atmosphere where locals usually enjoy walking and spending time.

Figure 1 shows the current state of the port.

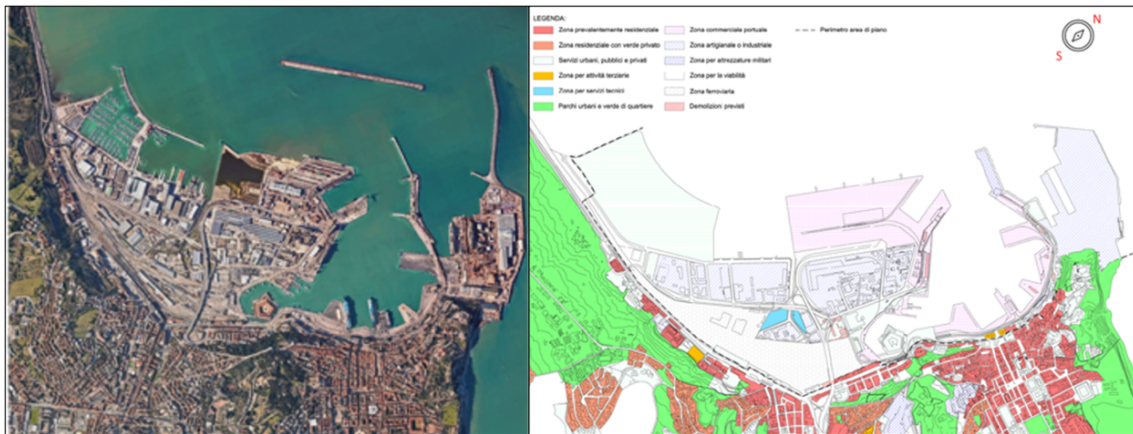


Figure 1: Aerial view (a) and road connection and distribution (b) of the Port of Ancona.

2.2. Acoustic measurements

Noise from transportation infrastructures is regulated by the specific implementing regulations pursuant to Article 11 of L.447/95 [20] such infrastructures are not subject, within their respective bands, to emission and immissions limits and to the attention values provided by the D.P.C.M. November 14, 1997 "Determination of the limit values of sound sources" [12, p. 447] . In addition, Article 4, of that decree, stipulates, that even the differential immissions limit values do not apply to noise produced: by road, rail, airport and maritime infrastructures.

This insufficiency of conventional methods creates the need to experiment with alternative approaches to environmental management, such as those focused on the soundscape.

Information was collected for the geographical and acoustic characterization of the area under study by also taking advantage of the technical report on the acoustic classification plan for the Ancona area approved by the Marche Region.

Therefore, information such as the influence of the port area, the acoustic classification and the measurement points of the noise monitoring network were collected and examined. The seaport currently operates within the limits prescribed by Decree DPCM 14/11/97 in which the environmental noise classification of areas is explained, and the port area of the city of Ancona includes public land of three classes: class IV, affected by vehicular traffic, with high population density, and presence of commercial activities. Classes V and VI cover the industrial area (see figure 2).

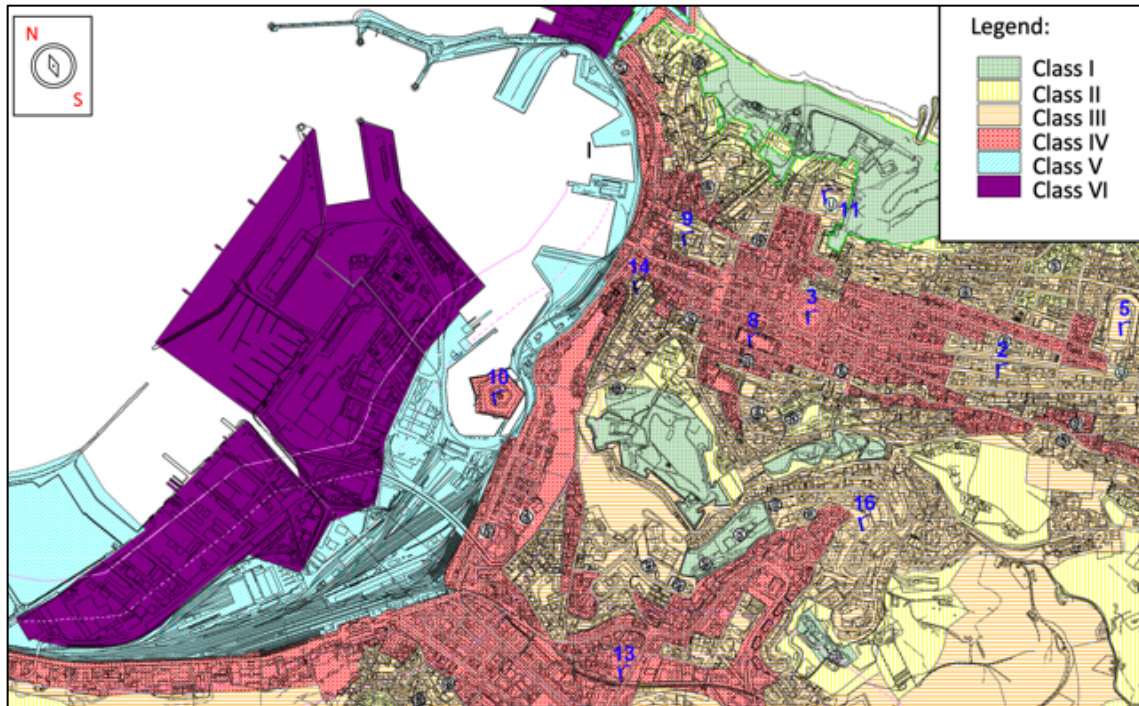


Figure 2: Acoustic classification: framing of the port area.

Experimental measurements allowed the determination of acoustic parameters such as sound pressure levels (SPLs) and the main objective parameters of psychoacoustics: Loudness and relative percentiles, Sharpness, roughness and Fluctuation strength [13].

Sound levels for the calculation of psychoacoustic parameters were measured using a 4100-type head and torso simulator.

The head and torso simulator, equipped with microphones at the entrance of the ear canals allows the shape, size and acoustic impedance of the listener's head and torso to be maintained; it is also capable of maintaining sound directionality.

For headphone playback, the track acquired with the binaural head was chosen. The analysis was performed using Soundbook MK2 software (Spectra) and Sound Quality software (PULSE, B&K).

All data from the site were collected under the same conditions: daytime (11:00 a.m. to 4:00 p.m.) in October 2021 (first week of the month), clear weather, and temperature ranging between 18° C and 21° C.

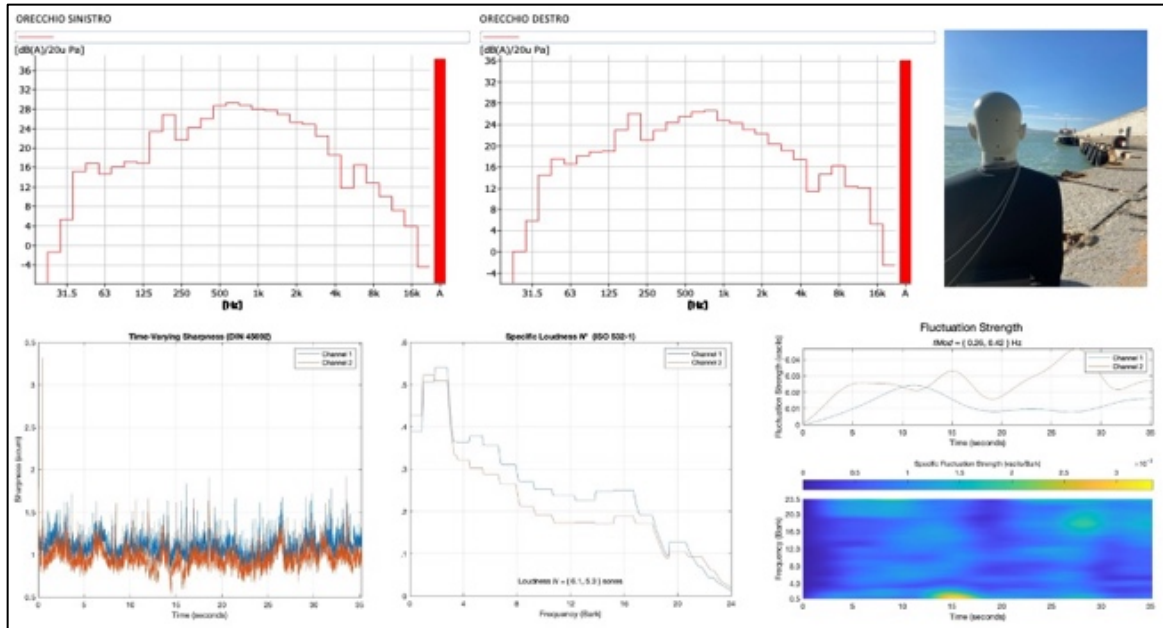
As for the visual data, however, these were recorded in 8k resolution as recommended in ITU-R BT.2020.

Table 1 show the result of acoustic and psychoacoustic measurements in the historical port (named Cluster C1) and figure 3 show the graphic trend of measurements in situ.

Table 1

Acoustic and Psychoacoustic parameters calculated on the binaural recording.

Cluster		SPL (dBA)	Loudness (phon)	Sharpness (acum)	Fluctuation Strength (vacil)	Roughness (asper)
C1	L	34,9	52,33	1,11	0,33	1,17
	R	37,7	54,98	1,10	0,41	1,58

**Figure 3:** Graphic trend of measurements for the C1 Cluster.

2.3. Cluster analysis and pattern recognition

To achieve this goal, an efficient vision system to recognize objects is required, because both smart objects and custom attributes are represented by physical objects that must be precisely detected. Object detection is, indeed, one of the areas that is maturing rapidly thanks to deep learning innovation. Current object detection methods are typically based on Convolutional Neural Network (CNN) models, able to automatically recognize visual features exploiting different architectures [14]. One of the first models featuring convolutions and shared weights was LeNet [15]. However, the spread of the deep learning approach for image and object classification was determined by AlexNet [16], developed in 2012 as an enhanced version of LeNet. ZFNet [17] further improved AlexNet by exploiting deconvolution network, while GoogLeNet introduced the Inception module reducing the number of network parameters. In 2016, the residual network ResNet became the state-of-the-art for the practical use of such models.

These pattern recognition techniques and advanced neural networks are essential for the preservation and environmental monitoring of the historic port of Ancona. By implementing an advanced vision system, we can accurately identify and catalog historical objects and structures, monitor changes over time, and detect any damage or deterioration. The use of CNNs and their variants allows for the analysis of large amounts of visual data, ensuring continuous and detailed monitoring, which is fundamental for the conservation of such a significant historical site. Neural networks, with their capacity for learning and adaptation, offer powerful tools for sound and image recognition, crucial for preserving both tangible and intangible aspects of cultural heritage.

To preserve the unique soundmarks of the port, a sophisticated algorithm was constructed. This algorithm captures, analyzes, and classifies the sound environment of the port, providing valuable insights for heritage preservation and environmental monitoring, ensuring that the port's sonic identity is maintained even in the face of potential changes or disruptions. This tool represents a significant step forward in the holistic conservation of Ancona's maritime heritage, integrating auditory elements into broader conservation and planning efforts.

The development of the pattern recognition algorithm involves in five phases:

- i) **Image Classification:** Image classification was approached as a supervised learning problem where a set of target classes (objects to be identified in images) was defined, and the model was trained to recognize them using labeled video examples.
- ii) **Cluster Analysis:** Cluster analysis involved the application of clustering algorithms with the goal of finding hidden patterns or groupings in a dataset.
- iii) **Convolutional Neural Networks (CNNs):** CNNs were employed due to their effectiveness in image classification and potential for audio classification. For this purpose, an audio classifier for the case study was constructed using Matlab-code, designed specifically to identify the type of noise in the area. The code grants to aggregate local classifications to whole sound column decisions by mimicking the visual video classification of Hershey et al. [18]
- iv) **Training with Binaural Recordings:** The algorithm was trained using binaural recordings collected during the measurement campaign. These recordings provided a comprehensive and immersive representation of the sound environment, capturing the unique acoustic characteristics of the port.
- v) **Temporal Analysis:** After studying several more complex models for combining information over time, a simple average single-frame CNN classification output was found where timestamps correspond to each analyzed region.

The system provided detailed outputs such as:

- *Sounds* - Sounds detected in each region;
- *Average Scores* - Average network scores corresponding to each class of sounds detected in the region;

- *Max Scores* - Maximum network scores corresponding to each sound class detected in the region;

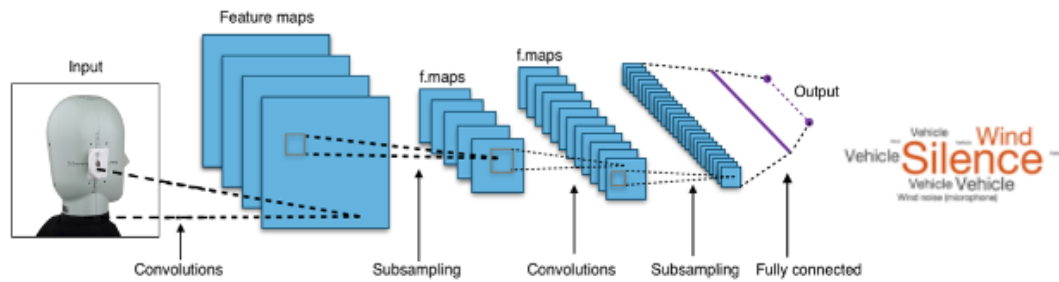


Figure 4: Workflow of the CNN classification.

By capturing and analyzing these soundmarks unique to the port, the system provides a comprehensive understanding of the acoustic environment.

3. Results

The algorithm developed for the preservation and environmental monitoring of the historic port of Ancona effectively categorized the recorded sounds into distinct classes. These classifications were based on the framework provided in [13] and they are crucial for understanding the soundscape dynamics and informing conservation strategies.

The sounds of the historic port were classified into the following categories:

- **TRAFFIC NOISE:** This category includes sounds from boats, cars, and sirens. These sounds are often associated with the daily operations of the port and nearby urban activities.
- **OTHER NOISE:** This includes sounds from construction activities, industrial operations, machinery, and inappropriate music. These are typically considered as intrusive or undesirable noises that can affect the acoustic environment.
- **SOUNDS FROM HUMAN BEINGS:** Sounds in this category come from conversations, laughter, children playing and footsteps. These sounds reflect the human presence and activities within the port area.
- **NATURAL SOUNDS:** These are sounds from the natural environment, such as singing birds, wind in vegetation, flowing water, and sea waves. Natural sounds contribute to the acoustic diversity and aesthetic quality of the soundscape.
- **DESIGNED SOUNDS:** This category includes sounds that are intentionally added to the environment for specific purposes, such as ambient music or public announcements.
- **ACOUSTIC EFFECTS:** These are sounds that result from the acoustic characteristics of the environment, including echoes and reverberations.

To provide a comprehensive analysis of the soundscape, a temporal sound map was created. This map visualizes the distribution and intensity of the different sound categories over

time. Figure 5 illustrates the temporal sound map of cluster C1, highlighting the variations in sound classification within a specific timeslot.

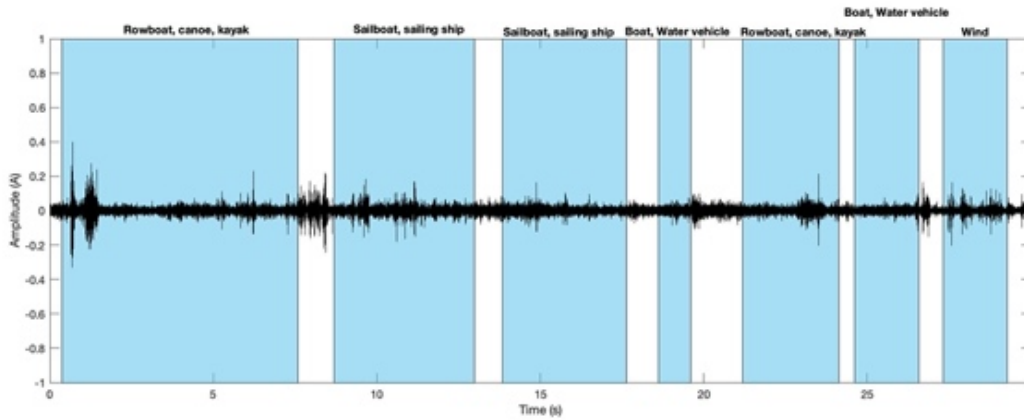


Figure 5: Cloud of cluster C1: Average Scores pari a 0.42 e Maxscores pari a 0.53.

In addition, agreeing with Shafer's theory, keynotes, soundmarks and soundsignals were distinguished and identified according to the following classification for the cluster C1 (see table 2).

Table 2

Keynotes, soundmarks e soundsignal in cluster C1.

Keynotes	Soundsignal	Soundmarks
Traffic	Siren	Rowing boats, canoes, kayaks, Sailboats, Sea waves, Clip-Clop
Air noise	Allarm	
Water	Siren of protection of the boat	

The temporal sound map provides the maximum and minimum network scores: for each sound class detected in the timeslot, the algorithm calculates the maximum and minimum network scores. These scores represent the confidence level of the neural network in classifying the detected sounds into the predefined categories. Higher scores indicate a stronger presence or more frequent occurrence of a particular sound class during the analyzed period.

In addition, the temporal sound map is a crucial tool for several reasons:

- *Monitoring Changes Over Time:* By visualizing sound data over time, it becomes possible to monitor changes in the soundscape. This can help identify patterns related to specific events, activities, or environmental changes that affect the acoustic environment.
- *Detecting Anomalies:* The map helps in detecting anomalies or unexpected changes in the soundscape, such as sudden increases in construction noise or other disturbances. This is vital for timely interventions to preserve the acoustic quality of the historic port.
- *Enhancing Conservation Efforts:* Understanding the temporal dynamics of sounds allows conservationists to develop targeted strategies for preserving the unique sonic identity of the port. For example, measures can be taken to mitigate intrusive noises while enhancing the presence of natural and designed sounds.
- *Informing Urban Planning:* The insights gained from the temporal sound map can inform urban planning and development decisions, ensuring that the acoustic environment is considered alongside other factors.

By utilizing binaural recordings and advanced neural network techniques, the developed algorithm provides a powerful tool for the holistic conservation of the historic port of Ancona. It ensures that both the tangible and intangible aspects of the port's heritage, including its unique soundscape, are preserved and protected for future generations.

Thanks to this investigation, it was possible to identify the actual content of the key notes and sound signals in the track; this was essential for the construction of the quality metrics.

The use of a clustering analysis allowed sound stimuli to be grouped according to common characteristics, thus simplifying the understanding of emerging patterns and recurring trends. This methodology helped to more clearly delineate the complexity of the data and provide a better perspective on the inherent differences among the various sound stimuli examined.

4. Conclusion

The implementation of advanced pattern recognition techniques and neural networks for the preservation and environmental monitoring of the historic port of Ancona has yielded significant and promising results. The developed algorithm successfully categorized the port's sounds into six distinct classes: traffic noise, other noise, sounds from human

beings, natural sounds, designed sounds, and acoustic effects. This classification was made possible using binaural recordings collected during the measurement campaign, which provided a rich and immersive dataset for precise training of the neural network.

The creation of a temporal sound map for cluster C1 offered a detailed understanding of the port's acoustic environment, illustrating the variations in sound classification over time. By calculating the maximum and minimum network scores for each sound class, the system highlighted the dynamic nature of the soundscape, revealing the presence and intensity of different sound sources throughout the day. This comprehensive approach ensures a more thorough protection of the port's cultural heritage by integrating auditory elements into the conservation strategy.

Furthermore, the ability to monitor changes and detect anomalies in the soundscape allows for proactive conservation measures, ensuring the port's sonic identity is maintained amidst environmental and urban changes. The insights gained from the sound classification and temporal sound map can inform urban planning decisions, helping to mitigate intrusive noises and enhance the presence of natural and designed sounds in the port area.

In conclusion, the research demonstrates the effectiveness of combining pattern recognition techniques and neural networks for cultural heritage conservation. The developed system not only provides a detailed understanding of the historic port's soundscape but also offers valuable tools for its preservation.

References

- [1] M. D. Fowler, 'Soundscape as a design strategy for landscape architectural praxis', *Design Studies*, vol. 34, no. 1, pp. 111–128, 2013, doi: 10.1016/j.destud.2012.06.001.
- [2] C.-J. Yu and J. Kang, 'Soundscape in the sustainable living environment: A cross-cultural comparison between the UK and Taiwan', *Science of the Total Environment*, vol. 482–483, no. 1, pp. 501–509, 2014, doi: 10.1016/j.scitotenv.2013.10.107.
- [3] A. Kaplan, 'Landscape architecture's commitment to landscape concept: A missing link?', *Journal of Landscape Architecture*, vol. 4, no. 1, pp. 56–65, 2009, doi: 10.1080/18626033.2009.9723413.
- [4] T. V. Renterghem *et al.*, 'Interactive soundscape augmentation by natural sounds in a noise polluted urban park', *Landscape and Urban Planning*, vol. 194, p. 103705, 2020, doi: <https://doi.org/10.1016/j.landurbplan.2019.103705>.
- [5] R. Pheasant, K. Horoshenkov, G. R. Watts, and B. Barrett, 'The acoustic and visual factors influencing the construction of tranquil space in urban and rural environments: tranquil spaces-quiet places?', *The Journal of the Acoustical Society of America*, vol. 123, pp. 1446–57, Apr. 2008, doi: 10.1121/1.2831735.
- [6] A. M. Jian KANG Francesco ALETTA, Tin OBERMAN, Mercede ERFANIAN, Magdalena KACHLICKA, Matteo LIONELLO, 'Towards soundscape indices', *PROCEEDINGS of the 23rd International Congress on Acoustics*, 2019.
- [7] J. Liu, J. Kang, H. Behm, and T. Luo, 'Effects of landscape on soundscape perception: Soundwalks in city parks', *Landscape and Urban Planning*, vol. 123, pp. 30–40, 2014, doi: <https://doi.org/10.1016/j.landurbplan.2013.12.003>.

- [8] J. W. Smith and B. C. Pijanowski, 'Human and policy dimensions of soundscape ecology', *Global Environmental Change*, vol. 28, no. 1, pp. 63–74, 2014, doi: 10.1016/j.gloenvcha.2014.05.007.
- [9] J. L. Carles, I. L. Barrio, and J. V. de Lucio, 'Sound influence on landscape values', *Landscape and Urban Planning*, vol. 43, no. 4, pp. 191–200, 1999, doi: [https://doi.org/10.1016/S0169-2046\(98\)00112-1](https://doi.org/10.1016/S0169-2046(98)00112-1).
- [10] X. Xu and H. Wu, 'Audio-visual interactions enhance soundscape perception in China's protected areas', *Urban Forestry and Urban Greening*, vol. 61, 2021, doi: 10.1016/j.ufug.2021.127090.
- [11] F. Balducci, P. Buono, G. Desolda, D. Impedovo, and A. Piccinno, 'Improving smart interactive experiences in cultural heritage through pattern recognition techniques', *Pattern Recognition Letters*, vol. 131, pp. 142–149, Mar. 2020, doi: 10.1016/j.patrec.2019.12.011.
- [12] 'Determinazione dei valori limite delle sorgenti sonore', Recommendation Legge 14 novembre 1997, 1997.
- [13] International standard ISO, 'Acoustics — Soundscape — Part 2: Data collection and reporting requirements', International Organization for Standardization, Standard ISO/TS 12913-2:2018, 2018. [Online]. Available: <https://www.iso.org/standard/75267.html>
- [14] F. Balducci, D. Impedovo, and G. Pirlo, 'Detection and Validation of Tow-Away Road Sign Licenses through Deep Learning Methods', *Sensors*, vol. 18, no. 12, 2018, doi: 10.3390/s18124147.
- [15] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, 'Gradient-based learning applied to document recognition', *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.
- [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, 'ImageNet Classification with Deep Convolutional Neural Networks', in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. Burges, L. Bottou, and K. Q. Weinberger, Eds., Curran Associates, Inc., 2012. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf
- [17] K. He, X. Zhang, S. Ren, and J. Sun, 'Deep Residual Learning for Image Recognition', in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.
- [18] Hershey Shawn et al., 'CNN Architectures for Large-Scale Audio Classification', *IEEE International Conference on Acoustics*, pp. 131–35, 2017.