DOI: 10.61782/fa.2025.0585

FORUM ACUSTICUM EURONOISE 2025

MONITORING OF MANUFACTURING PROCESS USING UNSUPERVISED SOUND EVENT CLASSIFICATION WITH APPLICATION OF SOUND DIRECTIVITY

Jurij Prezelj[1*] **Andrej Hvastja**[1] **Tomaž Berlec**[1]
[1] University of Ljubljana, Faculty of Mechanical Engineering, Slovenia

ABSTRACT

This study explores the use of sound directionality (DOA) as a key feature for unsupervised classification of acoustic events in industrial environments. A novel methodology was developed, combining a custom microphone array with SubWindowing to extract directional, spectral, and statistical features from production noise. Acoustic events were clustered using the k-means algorithm, with the optimal number of clusters determined via the elbow and silhouette methods. Feature importance was evaluated using Principal Component Analysis (PCA), which consistently identified sound directionality as the most influential feature. The results showed that integrating DOA significantly improved clustering performance and enabled accurate identification of machine states. The detected clusters aligned well with manually recorded operational states, confirming the method's effectiveness. This approach offers a robust and interpretable framework for real-time monitoring and fault detection in industrial settings. Future work will focus on extending this method to predictive maintenance systems and broader manufacturing environments.

Keywords: *acoustic event classification, sound directionality, unsupervised learning, k-means clustering, industrial monitoring*

[*] ***Corresponding author***: *jurij.prezelj@fs.uni-lj.si*

1. INTRODUCTION

Current machine learning approaches for acoustic event classification in industrial environments typically rely on spectral, statistical, and temporal features. Traditional supervised methods have been widely applied, but unsupervised techniques, such as k-means clustering, are gaining traction due to their ability to classify unlabeled data and reduce dimensionality in high-dimensional datasets [4]. For instance, Ding and Li [4] proposed a feature selection framework for k-means, while Sun and Wang [5] introduced improvements via an ADMM algorithm to enhance clustering performance. Coates and Ng [6] also emphasized the role of k-means in feature learning, particularly due to its simplicity and effectiveness. In audio classification specifically, deep learning methods have been reviewed by Mesaros et al. [7], highlighting the importance of robust feature representation. However, these methods largely overlook spatial acoustic features, such as sound directionality, also known as Direction of Arrival (DOA). Despite its well-known role in human perception, evident in the "cocktail party effect", DOA remains underexplored in unsupervised learning pipelines for acoustic event classification [2]. Base on literature review and our expertise we identified a research gap, placed a hypothesis, detect challenges and placed objectives of this study.

Research Gap: Current machine learning approaches for acoustic event classification typically rely on spectral, statistical, and temporal features. However, the role of sound directionality has been underexplored as a key feature for classification. Despite the relevance of directionality in distinguishing sound sources, it remains inadequately studied within unsupervised learning models, such as k-means clustering, for acoustic event classification [2].

Hypothesis: Incorporating DOA as a core feature can significantly improve clustering performance and enable more accurate detection of machine states and process deviations in industrial environments.

Challenges: Accurate measurement of DOA requires precise sensor arrays and fast-response algorithms (<50 ms). Another challenge lies in combining DOA with time–frequency features into a unified feature set and validating its impact using robust dimensionality reduction methods such as Principal Component Analysis (PCA).

Objective:

This study aims to investigate the role of DOA in the classification of acoustic events using unsupervised learning, specifically focusing on k-means clustering.

2. METHODOLOGICAL OVERVIEW

The classification system developed in this study consists of three main stages: (1) data acquisition using a custom-designed differential microphone array, (2) feature extraction incorporating both time–frequency and directionaly DOA, and (3) unsupervised classification using the k-means algorithm. To enhance temporal resolution, a SubWindowing technique was used to segment the acoustic data. Features were extracted from both time and frequency domains, including Direction of Arrival (DOA) information computed via beamforming algorithms. Dimensionality reduction techniques such as Principal Component Analysis (PCA) was employed to assess feature importance and ensure model interpretability. The proposed system enables precise spatial classification of acoustic events and supports real-time monitoring in industrial environments.

2.1 Feature Extraction

The feature extraction process is a critical step in the overall methodology, converting raw sound signals into meaningful descriptors that can be used for clustering and classification. Key features are extracted from both the time and frequency domains. Spectral features such as spectral centroid, spectral bandwidth, spectral flatness, and spectral roll-off are calculated for each SubWindow segment. Temporal features, including Zero-Crossing Rate (ZCR), Root Mean Square (RMS) energy, and Sound Pressure Level (SPL), are also extracted. Additionally, statistical features like skewness, kurtosis, and entropy are computed to capture distributional properties of the signal. A unique contribution of this study is the inclusion of sound directionality as a feature, which is derived from the microphone array's spatial configuration. This information is combined into a feature vector representing each time segment, enabling a comprehensive representation of the acoustic event.

The feature set is subsequently refined using dimensionality reduction techniques, including Principal Component Analysis (PCA), to identify and retain only the most influential features for clustering

2.2 Time Intervals of feature extraction and their integration

In industrial acoustic environments, sound signals are inherently complex. They are composed of a combination of impulsive, transient events, such as impacts, cutter engagements, or part drops, and longer lasting background or cyclical machine operations. Capturing this full acoustic landscape using a single fixed time resolution would either miss rapid changes or smooth over important long-term patterns. To address this, we implemented a dual-time-constant segmentation strategy, which enables the extraction of both fast-evolving and slowly varying features from the signal.

2.2.1 Motivation for the Approach

The main rationale behind this two-level segmentation was to construct a feature representation that simultaneously captures the detail of transient events and the consistency of long-term operational patterns. Many machine learning models fail to distinguish between these layers of acoustic behavior when using a single time scale, which leads to suboptimal clustering or misclassification of machine states. By layering feature extraction over both short and long intervals, we aimed to provide the unsupervised algorithm with richer and more distinguishable patterns for clustering.

This approach mirrors human auditory perception, where the brain reacts to sudden sounds immediately while also tracking longer sequences and rhythms over time. The short segments allow us to detect quick, local fluctuations in sound energy or frequency, while the longer segments offer context highlighting how those fluctuations develop over time.

2.2.2 Short-Term Segmentation (τ_z)

The sound signal was initially divided into short overlapping windows of duration τ_z. These segments were small enough to preserve detail on the time scale of tens of milliseconds, suitable for detecting tool strikes or abrupt machine transitions. For each τ_z segment, we applied a Short-Time Fourier Transform (STFT) to compute sound pressure level (SPL) values across multiple frequency bands. This frequency-domain representation was essential for identifying spectral structures linked to different machine behaviors. The use of overlapping windows ensured smoother transitions between adjacent segments, preventing the loss of critical information at boundary points and improving temporal continuity in the feature space.

2.2.3 2.2.3 Long-Term Aggregation (τ_k)

To build a higher-level representation, we grouped i consecutive τ_z segments into a longer interval $\tau_k = i \cdot \tau_z$. Within each τ_k block, a set of statistical metrics was computed for every frequency band and acoustic feature:

- Mean: representing the average energy or behavior in the interval,
- Range: capturing variation or dynamics,
- Standard deviation and variance: indicating spread or instability,
- First-order temporal derivatives: highlighting changes and slopes.

This allowed us to extract a macro level picture of the evolving sound environment, revealing whether the machine operation was stable, variable, cyclic, or anomalous.

2.2.4 2.2.4 Temporal Derivatives and Spectral Dynamics

To further enrich the feature set, we computed temporal derivatives (differences between adjacent τz segments) for each SPL value, capturing short-term dynamics like rising or falling spectral power. These were aggregated in the same way as raw SPL values, maintaining the temporal structure of the sound while enabling detection of directional changes in signal characteristics.

We also calculated the spectral length for each frequency band, defined as the cumulative distance (or variation) between successive SPL values. This served as a descriptor of spectral "activity" or instability, providing a useful contrast between steady and fluctuating sounds.

2.2.5 2.2.5 Feature Matrix Construction

The final step involved concatenating all extracted descriptors—spectral, temporal, statistical, and derivative-based—within each τk interval into a composite feature vector. These vectors formed the rows of the feature matrix used for unsupervised clustering with the k-means algorithm. Importantly, we later appended the direction of arrival (DOA) features (see Section 2.3) to this same matrix, enabling the clustering to incorporate both spectral content and spatial origin.

This dual-time approach, illustrated in Figures 1 and 2, ensured that the model could effectively group acoustic events that were similar not only in sound quality, but also in time structure and directional behavior—essential for robust classification in complex industrial soundscapes.

$$L = \sum_{j=1}^{N-1} \sqrt{(\Delta f)^2 + (Lp_{j+1} - Lp_j)^2} \qquad (1)$$

By using this methodology for the analysis of acoustic data, numerous features of the sound can be extracted from the frequency domain. Furthermore, by transforming the data into a new time constant, it is possible to derive features that also describe changes in the time domain. This dual-domain approach ensures that both rapid and gradual changes in sound are effectively captured, thereby improving the clustering and classification of acoustic events.

Frequency bands	New time constant					μ	max-min	...
	$Lp_{1,1}$	$Lp_{2,1}$	...	$Lp_{i,1}$	$\longrightarrow$	z_1	z_{N+1}	...
	$Lp_{1,2}$	$Lp_{2,2}$	...	$Lp_{i,2}$	$\longrightarrow$	z_2	z_{N+2}	...
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\longrightarrow$	$\vdots$	$\vdots$	$\vdots$
	$Lp_{1,N}$	$Lp_{2,N}$	...	$Lp_{i,N}$	$\longrightarrow$	z_N	z_{N+N}	...

Figure 1: Grouping ith time steps into a time constant τk for each frequency band.

Frequency bands	New time constant - 1					μ
	$\Delta Lp_{1,1}$	$\Delta Lp_{2,1}$	...	$\Delta Lp_{i-1,1}$	$\longrightarrow$	z_1
	$\Delta Lp_{1,2}$	$\Delta Lp_{2,2}$	...	$\Delta Lp_{i-1,2}$	$\longrightarrow$	z_2
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\longrightarrow$	$\vdots$
	$\Delta Lp_{1,N}$	$\Delta Lp_{2,N}$	...	$\Delta Lp_{i-1,N}$	$\longrightarrow$	z_N

Figure 2: Grouping (i-1)th derivatives into a time constant τ_k for each frequency band.

2.3 Directionality (DOA) Estimation with SubWindow Beamforming

The process of measuring immission directivity necessitates the use of a microphone array and a beamforming algorithm, tools that aid in localizing sound sources or determining the Direction Of Arrival (DOA) of sound. Differential Microphone Arrays (DMAs), a distinct subset in sound localization, have emerged as the most appropriate beamforming method for a variety of applications that require speech recognition over the past decade. These applications include hands-free systems, mobile phones, and hearing aids. Circular DMAs have undergone extensive examination for speech and audio applications, primarily due to their adaptability in control, the capacity to establish frequency-invariant directivity patterns, and superior directivity factor [8] Their inherent benefits include significant beam alignment, frequency independence, and a compact geometrical configuration Several models have been constructed where all the first-order DMA computations are conducted in the time domain. This approach offers a crucial advantage - the computations exhibit minimal latency. This feature is particularly vital and advantageous in real-time applications [9, 10, 11]. The principle of merging microphone signals from a circular array [12] sparked the development of our unique algorithm [13, 14, 15, 16, 17]. This algorithm, designed to compute the DOA, utilizes a simplified method based on timed delay cascading pairs, as depicted in figure 3 and given by Eq. (2). to form a DMA.

$$argminj = \| \sum_{i=0}^{D} \left[x_{\mathrm{mic1}}(n) - x_{\mathrm{mic2}}\left(n - \frac{D}{2} + i\right)\right]^2 \|_{j=1}^{N} \quad (2)$$

The differential beamforming algorithm based on sub-windowing (SubW-DBA), is developed for this application and is described in our previous work [13] SubW-DBA has an advantage in that it is not limited to the low-frequency range, at least theoretically. In practice, the low-frequency limit is defined by the phase matching of the microphone pairs. The high-frequency limit depends on the nature of the sound. If the sound is random, there is no frequency limit. If the sound is harmonic, the boundary is at the distance between the microphone pairs, like the usage of sound intensity probes.

The algorithm defined in Eq. (2) was used to compute the instantaneous DOA to experimentally verify the hypothesis that the 2D immission directivity pattern can be obtained by associating the instantaneous total sound level (SPL) with the instantaneous dominant direction (DOA) during sub-windowing. The objective is to compile a collection of immission vectors from which an immission directivity pattern can be derived through consistent integration/averaging of the immission vectors. $x_{mic,1}(n)$ and $x_{mic,2}(n)$ in Eq. (2) represents the signals from microphone 1 and microphone 2 respectively. N denotes the number of directions. D is the number of samples in the observed window. The minimum value of D is defined by the speed of sound c, the sampling frequency f_s and the distance between two microphones, which determine the time resolution of direction detection. Index i represents the steering value, i.e., Δn, which is proportional to the time delay $\Delta\tau$.

$$D_{min} = \frac{X_{mic} f_s}{c} \quad (3)$$

For the fastest possible response of detecting the direction of arrival M = N.

$$\Delta t_{min} = \frac{D_{min}}{f_s} \quad (4)$$

Using a differential array with four microphones and a diameter/side length of 0.04 m, it is possible to shorten the length of the direction detection subwindow to 0.232 ms. Shortening the subwindow to 0.232 ms provides 44 signal values for calculating the DOA of the dominant sound source. If the signals from two microphones placed 40 mm apart are recorded at a sampling rate of 192 kHz, then 22 samples are needed to arrange the signals for the maximum delay for each direction, and thus the time resolution of detecting the dominant direction is 0.232 ms.

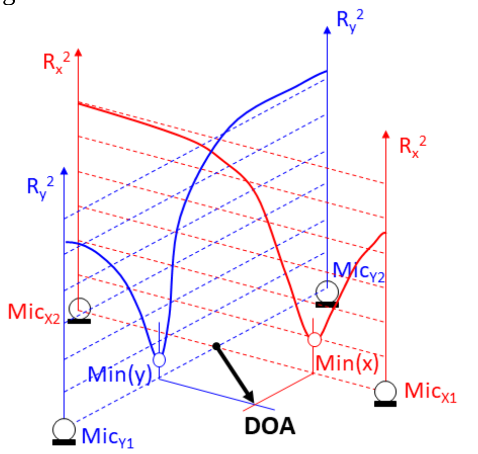

Figure 3: Differential microphone array of the first order, used during the experiment. Two pairs were used for the calculation of DOA on the immission plane.

2.4 Experimental setup

To optimize the quality of the acoustic data, a specific part of the production facility was selected, focusing primarily on the area where aluminum profiles are cut and processed using CNC (Computer Numerical Control) machinery. This decision assumed that these processes would produce distinguishable and characteristic sound signatures suitable for efficient clustering using the k-means algorithm.

The recording site was strategically positioned among seven cutting machines where incoming aluminum profiles are cut and mechanically processed before being automatically stacked into metal containers by handling robots. This location was chosen because it provided a rich source of distinctive mechanical sounds while being sufficiently

distant from the cluster of CNC machines, thereby reducing potential background noise interference.
To further minimize sound reflections and possible echo effects that could distort the acoustic data, the array of microphones was placed in an open area, as far away from wall surfaces as possible. The arrangement of microphones was carefully planned to avoid obstructing factory operations while ensuring the placement did not interfere with worker movement or the transit of forklifts. The selected recording site is depicted in figure. 5, where the observed machines are labeled from 1 to 7. The microphone array was oriented according to the layout shown in the figure.

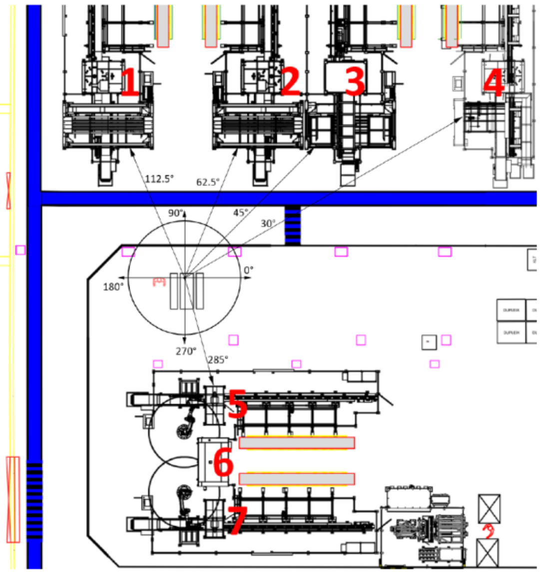

Figure 5: Recording site layout

Based on the microphone configuration, the expected orientation of the dominant sound during the operation of each machine was calculated. For machine 1, the sound direction was expected to be between 110° and 120°, for machine 2 it was estimated between 60° and 65°, and for machine 3, around 45°. Machine 4 was expected to have a dominant sound direction of around 30°, while machines 5, 6, and 7 had dominant sound directions expected to fall between 270° and 290°.
The machines being observed were expected to produce sharp, steady-frequency sounds caused by the milling cutter, as well as loud impulsive sounds generated when aluminum pieces were dropped into metal containers. To capture these sound events effectively, four omnidirectional microphones with polarized 6.35 mm condenser capsules were utilized. These microphones had a frequency range of 20 Hz to 20 kHz and a maximum sound pressure level (SPL) of 132 dB. The microphones were mechanically linked at equal distances to facilitate the calculation of the dominant sound direction.
The system also included a Behringer soundboard for initial pre-amplification and signal merging. For data storage, a simple signal acquisition program was developed using the LabVIEW software environment. This comprehensive setup ensured high-quality data acquisition, enabling the accurate capture of essential sound properties for subsequent feature extraction and clustering analysis.

3. RESULTS

Figure 6 presents a typical 30-second sound recording from the production environment. The upper panel shows the spectrogram, illustrating the spectral energy distribution over time. Several high-energy broadband events are visible, especially between seconds 4–6, 10–12, 17–18, and 23–25, which correspond to distinct mechanical actions such as cutting or part ejection.

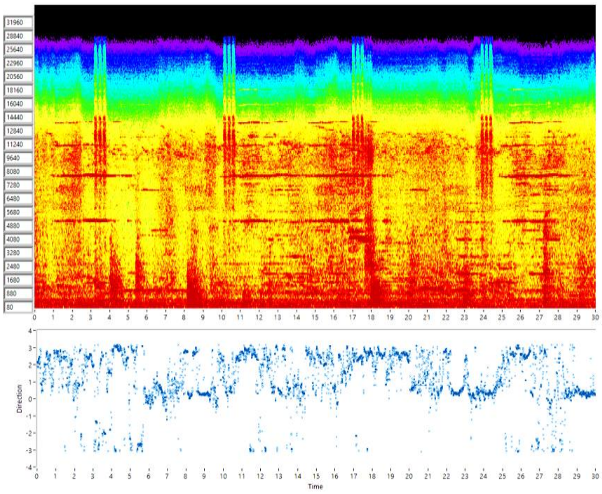

Figure 6. Spectrogram of a typical 30-second recording segment from the manufacturing process (top), with extracted dominant sound directions plotted below (bottom). The spectrogram displays energy distribution across frequencies, while the blue dots in the lower plot represent the instantaneous direction of arrival (DOA) of the sound.

The lower panel of the figure displays the corresponding instantaneous direction of arrival (DOA) of the dominant sound source, extracted using the SubWindow-DBA algorithm. The DOA estimates exhibit clustered patterns in both time and angular space, indicating repeated events

with consistent spatial origin, likely associated with recurring machine actions. Notably, transitions in DOA correspond well with changes in spectral content, affirming the relevance of including directionality as a classification feature.
This visualization supports the hypothesis that directionality provides an orthogonal, highly discriminative feature not captured by traditional spectral descriptors. The events clustered in the spectrogram are reinforced by clear directional patterns in the lower plot, validating the importance of sound directionality for distinguishing machine states. This correlation is consistent with the PCA results (Fig. 7), where DOA was identified as the most influential feature.

3.1 Identification of the Most Influential Features Using PCA

Typica spectrogram, together with extracted direction of Arival (DOA) in blue dots is presented in figure 6 above. To identify the most influential features in the dataset, a PCA analysis was employed. This process provided an assessment of the importance of each feature, which were then ranked according to their significance and the fifteen most important are displayed in Figure 7.

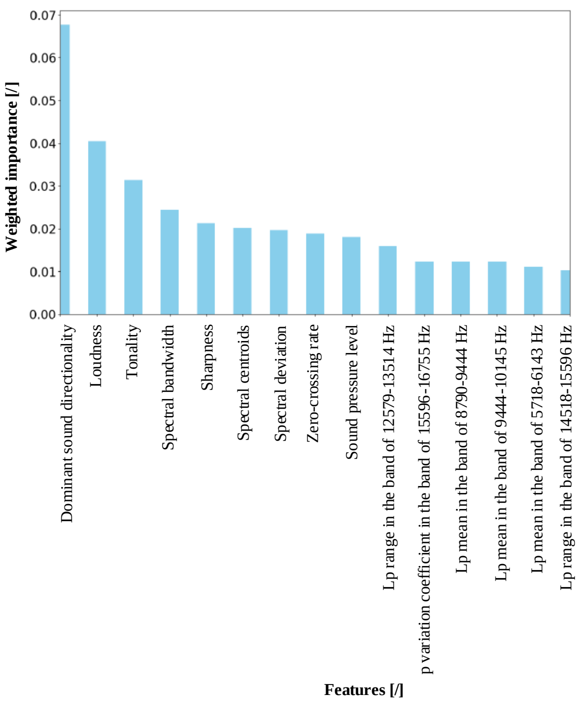

Figure 7: Weighted importance of the 15 most influential features below.

The graph reveals a decline in the weighted importance of features after the first few. This visualization serves as a useful tool for selecting and recognizing features. It helps to assess whether certain features contribute meaningfully to the quality of the data input or if they represent unnecessary noise. Figure 7 further allows for a clear understanding of which features contribute most significantly to the diversity of the data.
The analysis shows that the dominant feature is sound directionality. It has by far the highest weighted importance. This result is expected, as it is the only feature not directly derived from the sound signal itself and the only one that does not describe the characteristics or quality of the sound. Consequently, it has the lowest correlation with the other features, thereby contributing the most to the diversity of the data - a key aspect for successful classification. Following the dominant sound directionality, other key features calculated at the time constant τk=1 second, such as loudness, tonality, and spectral characteristics, also rank highly.
Next, a set of features derived from the frequency domain at a shorter time constant τ_z=50 milliseconds follow in terms of importance. Among these features, those describing higher frequency bands have the most significant influence. While these individual features do not possess high weighted importance on their own, together they still contribute significantly to the diversity of the data. This contribution underlines the relevance of combining features from different time scales and frequency ranges, ensuring the model captures both short-term and long-term acoustic patterns crucial for effective classification.
The optimal number of clusters for the k-means method can be determined in several ways. One common approach is an iterative trial-and-error method, where clustering begins with two clusters and gradually increases the number of clusters, observing changes in the classification results. The process continues until adding more clusters and no longer provides new information about the dataset. Another widely used approach involves the use of formal methods, such as the elbow method and the silhouette coefficient method, to identify the optimal number of clusters.

3.1.1 Elbow Method

The elbow method is visualized through a graph, as shown in Figure 8, which plots the sum of squared distances within each cluster as a function of the number of clusters *k*. The "elbow" of the curve is the point where the rate of decrease in the sum of squared distances slows down significantly, indicating that adding more clusters provides diminishing returns in reducing within-cluster variability. For this particular study, the elbow can be observed around k=3 or k=4, as the addition of clusters beyond this point results in only marginal reductions in the total within-cluster sum of squares. This observation suggests that selecting 3 or 4 clusters would provide an optimal balance between simplicity and classification performance.

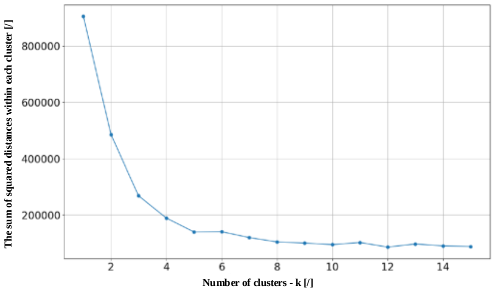

Figure 8. Elbow method graph for finding the optimal number of clusters - *k*.

3.1.2 Silhouette Coefficient Method

The silhouette coefficient method offers another perspective on selecting the optimal number of clusters. The silhouette coefficient measures how similar an object is to its own cluster compared to other clusters. A higher silhouette score indicates well-defined clusters. The silhouette graph for the data used in this study is presented in Figure 9. The peak of the graph is observed at k=2, indicating that, at this point, the clustering achieves the highest coherence of points within clusters. However, dividing the data into only two clusters limits the classification utility, as it would only distinguish between operational and non-operational states of the production process. This segmentation could still be useful for monitoring downtime, including pauses, breaks, and shift changes.

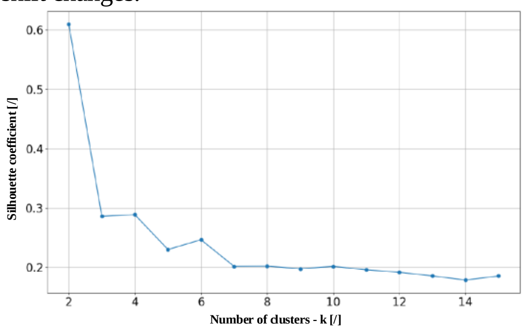

Figure 9. Silhouette coefficient graph for finding the optimal number of clusters - *k*.

As the number of clusters increases, the silhouette score gradually decreases, but it remains above zero, indicating that the cluster structure is still meaningful even with a larger number of clusters. Notably, local peaks are observed at k=4 and k=6, which correspond to improved clustering results relative to other values of k. Given that the elbow method also indicated a clear break at k=4, it can be inferred that selecting k=4 would provide the most interpretable and effective clustering solution for this dataset. This selection ensures a balance between model simplicity and classification accuracy, offering meaningful segmentation of acoustic events within the production process.

4. CONCLUSIONS

Based on the conducted research and the obtained results , the main findings can be summarized as follows:

1. We successfully classified acoustic signals using the k-means algorithm, effectively grouping the data into meaningful clusters based on their features. Initial classification was performed with k=4 clusters, followed by a refined analysis with k=6 clusters.
2. The feature extraction process was based on a methodology where frequency domain characteristics were derived at a lower time constant τ_0=50 ms. These features were then aggregated using various merging methods to form a higher time constant τ_k=1 second, thereby capturing time-domain properties as well.
3. Principal Component Analysis (PCA) was employed. A dimensionality reduction technique facilitated better organization of the data in a multidimensional space, ultimately improving the classification efficiency of the k-means algorithm.
4. It was found that the k-means algorithm was unable to capture all anticipated acoustic patterns, which can be attributed to the algorithm's limited generalization capabilities. Certain specific acoustic patterns, which are critical for a comprehensive understanding of machine operation, remain undetected.
5. The results indicate that the choice of the number of clusters k plays a crucial role in classification quality, but at the same time, it is challenging to determine objectively. The application of PCA to identify the most influential features revealed that the directionality of the dominant sound is the most important feature. This finding was expected since this is the only feature not directly derived from the sound signal and the only one that does not describe the sound's properties or quality. Consequently, it has the lowest correlation with other features and contributes the most to data variability, which is essential for successful classification. By implementing the PCA on the extracted features, we confirmed the hypothesis that the direction of the sound event is a significant feature that

should be included in classification algorithms for acoustic event detection.

The study demonstrated that the k-means algorithm is suitable for basic classification of acoustic signals in production environments, but it has limitations that require further investigation to achieve a more comprehensive capture of all acoustic patterns. Our study contributes to a better understanding of the applicability of classification algorithms in real-world industrial environments.

4.1 Suggestions for Future Work

Based on the findings of this study, it would be beneficial to explore the use of other unsupervised learning methods that were not considered in this research. While this study focused on the classification of specific acoustic patterns, future research could investigate longer periods of production operation and place greater emphasis on macro-level events. Additionally, efforts could be made to quantitatively assess the relative productivity of the production process over time. This broader perspective would provide a more comprehensive understanding of operational efficiency and help identify potential areas for optimization.

5. REFERENCES

[1] J.E.K. Foreman: *Sound Analysis and Noise Control*. New York: Van Nostrand Reinhold, 1990.

[2] J. Prezelj, L. Čurović, T. Novaković, and J. Murovec: "A novel approach to localization of environmental noise sources: Sub-windowing for time domain beamforming," *Applied Acoustics*, vol. 195, 2022.

[3] J. Prezelj, J. Murovec, S. Huemer-Kals, K. Häsler, and P. Fischer: "Identification of different manifestations of nonlinear stick–slip phenomena during creep groan braking noise by using the unsupervised learning algorithms k-means and self-organizing map," *Mechanical Systems and Signal Processing*, vol. 166, 2022.

[4] C. Ding and T. Li: "Unsupervised Feature Selection Ordering in High-Dimensional Data," *Proc. of the 25th International Conference on Machine Learning*, pp. 1033-1040, 2009.

[5] X. Sun, J. Wang: "Improved ADMM Algorithm for K-means-Based Unsupervised Feature Selection". arXiv preprint arXiv:2411.151972024

[6] A. Coates and A.Y. Ng: "Learning feature representations with k-means," In: *Neural Networks: Tricks of the trade,* Lecture Notes in Computer Science, vol. 7700. Springer, Berlin, Heidelberg, pp. 561-580, 2012

[7] A. Mesaros, T. Heittola, T. Virtanen, A review of deep learning methods for acoustic scene classification, *Applied Sciences*, Vol. 10, (6), 2020.

[8] G. Huang, J. Benesty, J. Chen, Design of robust concentric circular differential microphone arrays. J Acoust Soc Am, Vol.141(5), 2017, pp.3236–49. https://doi.org/10.1121/1.4983122.

[9] J. Benesty, C. Jingdong, Study and design of differential microphonearrays, Springer Science & Business Media; 2012. https://doi.org/10.1007/978-3-642-33753-6.

[10] Y. Buchris, I. Cohen, J. Benesty, First-order differential microphone arrays from a time-domain broadband perspective, In: 2016 IEEE international workshop on acoustic signal enhancement (IWAENC). IEEE; 2016. https://doi.org/10.1109/iwaenc.2016.7602886.

[11] Y. Buchris, I. Cohen, J. Benesty, On the design of time-domain differential microphone arrays, Appl Acoust, Vol.148, 2019, pp.212–22. https://doi.org/10.1016/j.apacoust.2018.12.013.

[12] G. Huang, J. Benesty, J. Chen, Design of robust concentric circular differential microphone arrays. J Acoust Soc Am, Vol.141(5), 2017, pp.3236–49. https://doi.org/10.1121/1.4983122.]

[13] J. Prezelj, L. Čurović, T. Novaković, J. Murovec, A novel approach to localization of environmental noise sources: Sub-windowing for time domain beamforming, Applied Acoustics, Volume 195, 30 June 2022, 108836

[14] J. Murovec, et al. Environmental noise event classification based on selforganizing map using psychoacoustic features and spatial filtering. Universitätsbibliothek der RWTH Aachen; 2019.

[15] I. Salom, et al. An implementation of beamforming algorithm on FPGA platform with digital microphone array. Audio Engineering Society Convention 138., Audio Engineering Society, 2015.

[16] D. Todorovic, et al. Implementation and application of FPGA platform with digital MEMS microphone array. Proceedings of the 4th international conference on electrical, electronics and computing engineering, 2017.

[17] J. Murovec, et.al, Microphone array based automated environmental noise measurement system. Appl Acoust Vol. 141, 2018, pp.106–14. https://doi.org/10.1016/j.apacoust.2018.07.004.].