Contrastive Metric Learning Loss-Enhanced Multi-Layer Perceptron for Sequentially Appearing Clusters in Acoustic Emission Data Streams

Oualid Laiadi[1], Ikram Remadna[2,3], El yamine Dris[1], Redouane Drai[1], Sadek Labib Terrissa[2], and Noureddine Zerhouni[4]

[1] *Research Center in Industrial Technologies (CRTI), Cheraga, P.O. Box 64, Algiers 16014, Algeria*
oualid.laiadi@gmail.com

[2] *LINFI Laboratory, University of Biskra*

[3] *National School of Artificial Intelligence (ENSIA) Algiers, Algeria*

[4] *FEMTO-ST Institute, Université Bourgogne Franche-Comté, CNRS, ENSMM*

ABSTRACT

Conventional structural health monitoring methods for interpreting unlabeled acoustic emission (AE) data typically rely on generic clustering approaches. This work introduces a novel approach for analyzing sequential and temporal acoustic emission (AE) data streams by enhancing a Multi-Layer Perceptron (MLP) with a contrastive metric learning loss function (MLP-CMLL)and Time Series K-means (TSKmeans) clustering. This dual approach, MLP-CMLL with TSKmeans, is crafted to refine cluster differentiation significantly. This method is designed to optimize cluster differentiation, bringing similar acoustic patterns closer and distancing divergent ones, thereby improving the MLP's ability to classify acoustic events over time. Addressing the limitations of traditional clustering algorithms in handling the temporal dynamics of AE data, MLP-CMLL with TSKmeans approach provides deeper insights into cluster formation and evolution. It promises enhanced monitoring and predictive maintenance capabilities in engineering applications by capturing the complex dynamics of AE data more effectively, offering a significant advancement in the field of structural health monitoring. Through experimental validation, we apply this method to characterize the loosening phenomenon in bolted structures under vibrations. Comparative analysis with two standard clustering methods using raw streaming data from three experimental campaigns demonstrates that our proposed method not only delivers valuable qualitative information concerning the timeline of clusters but also showcases superior performance in terms of cluster characterization.

Keywords: acoustic emission (AE), sequentially appearing clusters, data streams, structural health monitoring, contrastive metric learning, multi-layer perceptron (MLP)

1. INTRODUCTION

Structural Health Monitoring (SHM) is essential to ensuring the safety, longevity, and efficient maintenance of engineering structures across civil, mechanical, and aerospace fields. This discipline employs advanced technologies to proactively detect and address damages, aiming to avert catastrophic failures and optimize maintenance efforts. Among various SHM applications, the precision monitoring of bolted connections is particularly critical, given its profound impact on the structural integrity and stability of significant constructions such as bridges, aerospace structures, and wind turbines (Bolognani et al., 2018).

The vulnerability of bolted connections to loosening—and the profound implications of such—was dramatically underscored by the 2015 collapse of a 129-meter wind turbine in Sweden (Swedish Accident Investigation Authority, 2017). This incident, attributed to bolt looseness, resulted not only in significant financial loss but also highlighted the urgent need for early detection systems to prevent such disasters. While traditional bolt inspection techniques are effective, they are notably labor-intensive and can significantly interrupt operational workflows. This has led to a shift toward non-destructive testing (NDT) methods (Hoła & Sadowski, 2022), particularly the use of acoustic emission (AE) sensors (Sun, Yang, Li, & Xu, 2023; P. Xu, Zhou, Liu, & Mal, 2021; D. Xu, Liu, Li, & Chen, 2019), as more efficient alternatives.

AE sensors are distinguished by their ability to detect stress-induced changes within materials, offering a sophisticated means of identifying potential damages or loosening. Research, such as that conducted by Wang et al. (Wang, Song, Wang, & Li, 2013), demonstrates a correlation between AE signal energy and the axial load of bolts, enabling precise detection of bolt looseness through analysis of energy dissipation and signal amplitude. However, AE signals' complexity, marked by significant variations in amplitude and energy, coupled with susceptibility to environmental noise and interference, poses a significant challenge (Fu, Zhou, & Guo, 2023). Relying solely on a single AE characteristic often falls short in accurately reflecting bolt tightness. Therefore, there's a pressing need to develop innovative methods capable of quantifying AE signals' nonlinear characteristics and accurately interpreting bolt looseness, underscoring the demand for advanced analytical techniques.

The vast quantities of AE signals within data streams present a significant challenge in identifying ground truth, rendering supervised learning methods impractical for AE data interpretation or anomaly detection (Ramasso, Denoeux, & Chevallier, 2022; Ramasso, Placet, & Boubakar, 2015). This necessitates a pivot towards unsupervised learning techniques, such as K-means, fuzzy C-means (FCM), and Gaussian Mixture Models (GMM), to extract actionable insights from AE data. Among these approaches, Gaussian Mixture Models sequentially (GMMSEQ), introduced by Emmanuel Ramasso et al. (Ramasso, Denoeux, & Chevallier, 2022), stands out by incorporating temporal dynamics into the clustering of unlabeled AE data, thereby significantly enhancing parameter estimation related to damage progression.

Recent advancements highlight the growing significance of unsupervised and self-supervised learning methods, with a notable focus on contrastive metric learning. This approach harnesses the inherent similarities and contrasts within data to facilitate learning without the necessity for explicit labels, marking a pivotal shift toward more efficient representation learning (Saunshi, Plevrakis, Arora, Khodak, & Khandeparkar, 2019). By comparing input samples and manipulating their representations within the embedding space—drawing similar samples closer and distancing dissimilar ones—contrastive representation learning streamlines the learning process. It sidesteps the conventional need for labeling each sample, instead utilizing a pre-established similarity distribution to classify inputs into positive or negative pairs (Hassani & Khasahmadi, 2020).

Building on these insights, we propose a novel method that leverages the power of a Multi-Layer Perceptron (MLP) enhanced with a contrastive metric learning loss (MLP-CMLL), to adeptly handle AE data streams, particularly those exhibiting sequentially appearing clusters. The proposed MLP-CMLL approach, rooted in the principles of contrastive metric learning, aims to differentiate between similar and dissimilar features within the AE data, generating robust feature embeddings without the need for explicit labels. These embeddings serve as a powerful foundation for clustering, enabling our system to dynamically identify and group sequentially appearing clusters of AE data. By applying time series k-means clustering algorithm (TSKMean) (Huang et al., 2016), we can effectively cluster AE events based on both their feature similarities and their temporal characteristics. This integration enables the detection of sequentially appearing clusters, a common occurrence in AE data streams, thereby providing deeper insights into the material's behavior and the efficacy of the monitoring system.

The remainder of this paper is organized as follows: Section 2 introduces the proposed MLP-CMLL method, along with their respective data preprocessing methods. Section 3 describes the dataset and provides an analysis of experimental results. Finally, the main findings of this study are summarized in Section 4 along with a description of future work perspectives.

2. Proposed method

In this section, we delineate the architecture of the proposed framework, which aims to classify bolt tightening levels through the analysis of acoustic emission data streams. Figure 1 illustrates the overarching architecture of our proposed approach, specifically designed for clustering bolt tightening levels based on acoustic emission data streams. Subsequently, we will elaborate on the intricacies and functional components of the proposed framework, detailing each block's contribution to the overall system.

2.1. AE signal Preprocessing & Feature Extraction

The preprocessing and feature extraction of AE signals is a critical step in analyzing the raw data stream, performed through a three steps, initially outlined in (Kharrat, Ramasso, Placet, & Boubakar, 2016). The process begins with the data stream undergoing an initial filtration stage, employing a fifth-order high-pass filter with a cutoff frequency of 10 kHz and a passband ripple of 0.2 dB, effectively eliminating the DC component from the data.

- **Step 1**: **Wavelet filtering** Utilizing wavelet denoising on 250,000 sample frames achieves an optimal balance between computational efficiency and denoising quality. The chosen Daubechies "dB45" wavelet, featuring 90 coefficients and 14 decomposition levels, effectively identifies AE signal onsets (Kharrat et al., 2016). This step includes applying the soft Donoho-Johnstone universal threshold to the wavelet coefficients and adjusting for level-dependent noise, alongside correcting for any group delay introduced by the filtering process. Figure 2 displays the raw signal and denoised signal.

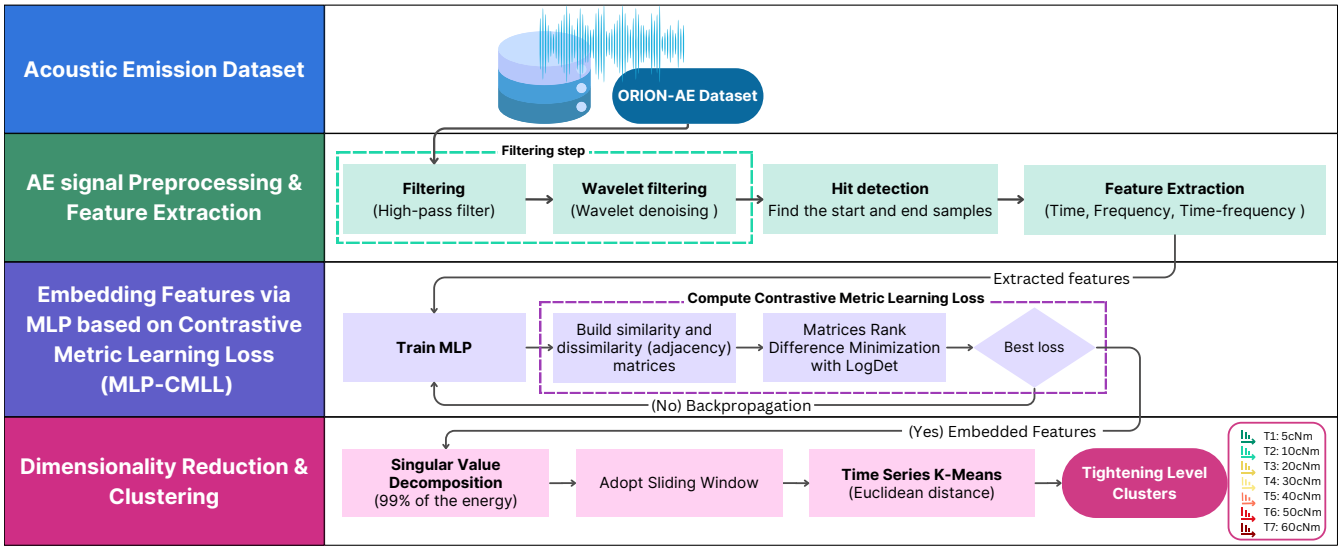

Figure 1. The general architecture of the proposed approach for bolt tightening level clustering.

- **Step 2**: **Hit Detection Procedure** aims to identify the start and end of each AE signal post-filtering based on amplitude thresholds (1.2 mV in this case). This step ensures that only relevant AE events are selected for analysis, utilizing specific counters ("HDT" 1100 μs and "HLT" 80 μs) to accurately demarcate signal boundaries (Kharrat et al., 2016).

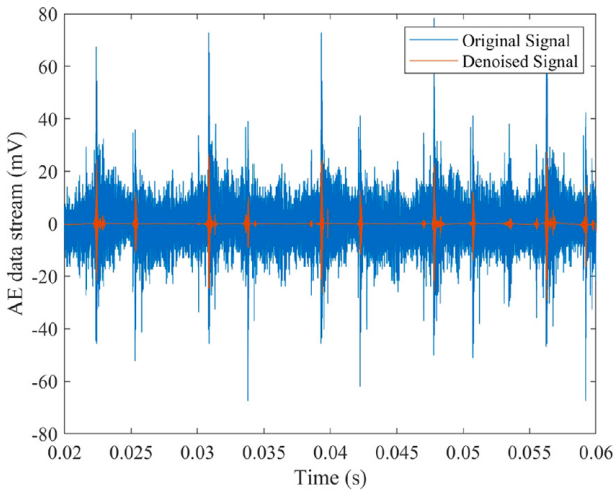

Figure 2. Raw signal and denoised signal.

- **Step 3 Feature Extraction**: Each detected AE signal is then analyzed to extract an extensive set of features, encompassing time-based and frequency-based characteristics (Kharrat et al., 2016; Sause, Gribov, Unwin, & Horn, 2012; Gonzalez Andino et al., 2000) such as rise time, counts, PAC-energy, amplitude, frequency metrics, signal strength, and energy distributions across specified frequency intervals. Figure 3 shows an AE signal and some typical features. Additional features include the Renyi number from the scalogram analysis using a Morlet wavelet and the frequency at maximum energy, providing a detailed signal characterization suitable for further analysis.

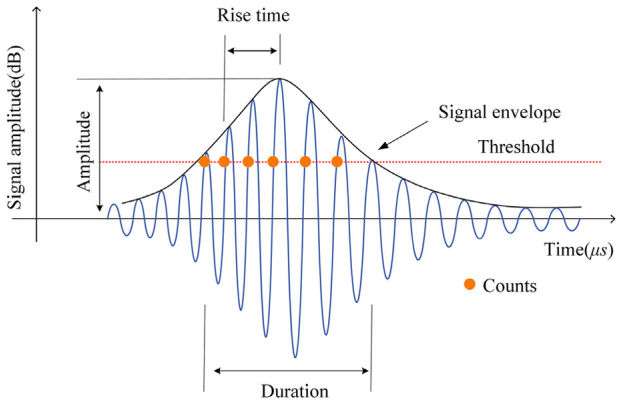

Figure 3. AE signal and some typical characteristics.

2.2. Embedding Features via MLP based on Contrastive Metric Learning Loss (MLP-CMLL)

This subsection will describe the proposed MLP based on the Contrastive Metric Learning Loss (MLP-CMLL) method. In the following, we will mention the details of our algorithm for learning a best loss metric based on an unsupervised metric learning with unlabeled data. The proposed contrastive metric learning framework is based on the combination of two methods, unsupervised EASE metric learning (Zhu & Koniusz, 2022a) and Generalized Laplacian Eigenmaps (Zhu & Koniusz, 2022b).

Let $\mathbf{X} \in \mathbb{R}^{n \times m}$ be an unlabeled AE data of n samples and m features. We propose a new MLP based on Contrastive Metric Learning Loss (MLP-CMLL) framework for unsupervised network embedding. To compute the loss function of the MLP framework, we calculate the logdet of scatter matrices based on the similarity and the dissimilarity (adjacency):

$$\mathbf{\Theta}^* = \arg\min_{\mathbf{\Theta}} \operatorname{Rank}\left(\mathbf{S}_{sim}\left(\mathbf{X}\right)\right) - \operatorname{Rank}\left(\mathbf{S}_{dis}\left(\mathbf{X}\right)\right) \quad (1)$$

Eq. (1) aims to compute a metric loss $\mathbf{\Theta}$ for each epoch that maximizes the similarity between similar features and minimizes the dissimilarity between dissimilar features.

Let:
$\mathbf{S}_{dis} = f_{\mathbf{\Theta}}(\mathbf{X})^\top \mathbf{L}_{dis} f_{\mathbf{\Theta}}(\mathbf{X})$ and $\mathbf{S}_{sim} = f_{\mathbf{\Theta}}(\mathbf{X})^\top \mathbf{L}_{sim} f_{\mathbf{\Theta}}(\mathbf{X})$
Then the LogDet relaxation becomes:

$$\begin{aligned}\mathbf{\Theta}^* &= \arg\min_{\mathbf{\Theta}} \log\det\left(\mathbf{I} + \alpha f_{\mathbf{\Theta}}(\mathbf{X})^\top \mathbf{L}_{sim} f_{\mathbf{\Theta}}(\mathbf{X})\right) \\ &\quad - \log\det\left(\mathbf{I} + \alpha f_{\mathbf{\Theta}}(\mathbf{X})^\top \mathbf{L}_{dis} f_{\mathbf{\Theta}}(\mathbf{X})\right) \\ &= \arg\min_{\mathbf{\Theta}} \log\det(\mathbf{I} + \alpha \mathbf{S}_{\text{sim}}) - \log\det(\mathbf{I} + \alpha \mathbf{S}_{\text{dis}})\end{aligned} \quad (2)$$

where $\mathbf{I}$ ensures $\mathbf{I} + \alpha f_{\mathbf{\Theta}}(\mathbf{X})^\top \mathbf{L} f_{\mathbf{\Theta}}(\mathbf{X}) > 0$ as $f_{\mathbf{\Theta}}(\mathbf{X})^\top \mathbf{L} f_{\mathbf{\Theta}}(\mathbf{X})$ may be $\mathbb{S}_+^m$ leading to $\det\left(f_{\mathbf{\Theta}}(\mathbf{X})^\top \mathbf{L} f_{\mathbf{\Theta}}(\mathbf{X})\right) = 0$. Thus, we use $\log\det(\mathbf{I} + \alpha\mathbf{S})$ as a smooth surrogate for $\operatorname{Rank}(\mathbf{S})$.

$$\begin{aligned}\mathbf{L}_{\text{sim}} &= \mathbf{I} - \tilde{\mathbf{A}}_{\text{sim}} \in \mathbb{S}_+^n, \\ \mathbf{L}_{\text{dis}} &= \mathbf{I} - \tilde{\mathbf{A}}_{\text{dis}} \in \mathbb{S}_+^n.\end{aligned} \quad (3)$$

Let us also define normalized graph Laplacian matrices in Eq. (2) as in Eq. (3). Let $\mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2} = \tilde{\mathbf{A}}$ and $\mathbf{D} = \operatorname{diag}\left(d_1, \cdots, d_n\right)$, where $d_i = \sum_j \mathbf{A}_{ij}$. We explain how we obtain $\mathbf{A}_{sim}$ and $\mathbf{A}_{dis}$ later in the text. From equations Eq. (3) and Eq. (2) we have:

$$\begin{aligned}\mathbf{L}_{\text{sim}} - \mathbf{L}_{\text{dis}} &= (\mathbf{I} - \tilde{\mathbf{A}}_{\text{sim}}) - (\mathbf{I} - \tilde{\mathbf{A}}_{\text{dis}}) \\ &= \tilde{\mathbf{A}}_{\text{dis}} - \tilde{\mathbf{A}}_{\text{sim}},\end{aligned} \quad (4)$$

As $\mathbf{L}_{\text{sim}} - \mathbf{L}_{\text{dis}} = \tilde{\mathbf{A}}_{\text{dis}} - \tilde{\mathbf{A}}_{\text{sim}}$, we obtain:

$$\begin{aligned}\mathbf{\Theta}^* = \arg\min_{\mathbf{\Theta}} &\log\det\left(\mathbf{I} + \alpha f_{\mathbf{\Theta}}(\mathbf{X})^\top \tilde{\mathbf{A}}_{\text{dis}} f_{\mathbf{\Theta}}(\mathbf{X})\right) \\ &- \log\det\left(\mathbf{I} + \alpha f_{\mathbf{\Theta}}(\mathbf{X})^\top \tilde{\mathbf{A}}_{\text{sim}} f_{\mathbf{\Theta}}(\mathbf{X})\right),\end{aligned} \quad (5)$$

where $\mathbf{A}_{\text{sim}}$ and $\mathbf{A}_{\text{dis}}$ are two different measurements with the opposite effect. Thus, we introduce parameter $\alpha > 0$ to balance the impact of these both terms.

Dissimilarity Matrix. Although one might design a linear projection based on the similarity relationship alone, we use both the dissimilarity information and the similarity matrix for learning a metric loss. Intuitively, in the context of a K-clustering task with n unlabeled samples and M_i queries for each cluster, we are addressing a problem where ($\mathrm{n} = \mathrm{K} \times \mathrm{M_i}$) samples are to be clustered into K groups. Here, off-diagonal entries are understood to signify distinct entities, whereas on-diagonal entries indicate identical entities. Thus, we form a dissimilarity matrix as the adjacency matrix of a densely connected graph:

$$\mathbf{A}_{dis} = \frac{1}{n}\mathbf{e}\mathbf{e}^\top - \mathbf{I}, \quad (6)$$

where $\mathbf{e}$ is an (n)-dimensional all-ones vector and $\mathbf{I}$ is the identity matrix.

Similarity Matrix. To measure the similarity between the pairs of samples, one has to choose a distance (or similarity measure) that will perform well in the clustering setting.

The typical choice for the measure of similarity is the RBF function $Z_{ij} = \exp\left(-\left\|f_\theta\left(\mathbf{x}_i\right) - f_\theta\left(\mathbf{x}_j\right)\right\|_2^2 / \sigma\right), \sigma > 0$ but the RBF function alone does not capture the structure of data. In this work, we claim that for the K-cluster learning task, the expected similarity matrix should be a K-block diagonal matrix. However, the similarity matrix based on the RBF kernel has no blockdiagonal structure.

Low-Rank Representation (LRR) (Liu, Lin, & Yu, 2010) expresses each data point $\mathbf{x}_i$ as a linear combination of other points, $\mathbf{x}_i = \sum_{j \neq i} Z_{ij}\mathbf{x}_j$, and uses the representational coefficient $\left(|Z_{ij}| + |Z_{ji}|\right)/2$ to measure the similarity between $\mathbf{x}_i$ and $\mathbf{x}_j$. LRR takes the correlation structure of data into account, and finds a low-rank representation instead of a sparse representation. In this work, the LRR is applied in the following rank minimization problem:

$$\arg\min_{\mathbf{Z}} \left\|f_\theta(\mathbf{X}) - f_\theta(\mathbf{X})\mathbf{Z}\right\|_F^2 \quad \text{s.t. } \operatorname{rank}(\mathbf{Z}) = K. \quad (7)$$

Eq. (7) is solved in two stages: 1) $\mathbf{Z} = \mathbf{V}^\top\mathbf{V}$, where $\mathbf{V}$ is obtained from the skinny SVD of $f_\theta(\mathbf{X}) = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$, and 2) for each row of $\mathbf{V}$, one only keeps top-K absolute largest entries of $\mathbf{\Sigma}$. Given the feature matrix $f_\theta(\mathbf{X})$, we obtain the representation matrix $\mathbf{Z}$ by solving Eq. (7). The similarity matrix is defined as $\mathbf{W}_{\text{sim}} = |\mathbf{Z}| - \operatorname{diag}(|\mathbf{Z}|)$.

We provide our implementation in Alg. 1. The proposed algorithm targets unsupervised network embedding by employing contrastive metric learning loss to enhance similarity among similar features while reducing dissimilarity among different ones. Central to this framework are the LogDet relaxation and Low-Rank Representation (LRR), both aimed at achieving an embedding that accurately captures the inherent structure of unlabeled data. This structured approach outlines a comprehensive step-by-step methodology for implementing the MLP-CMLL method, specifically designed to optimize metric learning loss in scenarios involving unlabeled datasets.

2.3. Time Series K-Means for clustering

Following the generation of high-dimensional embedded features via MLP-CMLL, with each feature vector comprising

Algorithm 1 MLP based on Contrastive Metric Learning Loss (MLP-CMLL)

Input: $\mathbf{X} \in \mathbb{R}^{n \times m}$: Unlabeled AE data of n samples and m features.
Initialize: Predefined MLP architecture, $\alpha > 0$.

Compute Similarity and Dissimilarity Matrices based on Laplacian Matrices:
$\mathbf{S}_{\text{dis}} = f_{\mathbf{\Theta}}(\mathbf{X})^\top \tilde{\mathbf{A}}_{\text{dis}} f_{\mathbf{\Theta}}(\mathbf{X})$
$\mathbf{S}_{\text{sim}} = f_{\mathbf{\Theta}}(\mathbf{X})^\top \tilde{\mathbf{A}}_{\text{sim}} f_{\mathbf{\Theta}}(\mathbf{X})$

Optimization:
while not converged **do**
    Solve for $\mathbf{\Theta}^*$ minimizing:

$$\begin{aligned}\mathbf{\Theta}^* &= \arg\min_{\mathbf{\Theta}} \log\det\left(\mathbf{I} + \alpha f_{\mathbf{\Theta}}(\mathbf{X})^\top \tilde{\mathbf{A}}_{\text{dis}} f_{\mathbf{\Theta}}(\mathbf{X})\right) \\ &\quad - \log\det\left(\mathbf{I} + \alpha f_{\mathbf{\Theta}}(\mathbf{X})^\top \tilde{\mathbf{A}}_{\text{sim}} f_{\mathbf{\Theta}}(\mathbf{X})\right) \\ &= \arg\min_{\mathbf{\Theta}} \log\det(\mathbf{I} + \alpha \mathbf{S}_{\text{sim}}) - \log\det(\mathbf{I} + \alpha \mathbf{S}_{\text{dis}})\end{aligned}$$

    Update MLP parameters.

    Adjust Matrices ($\mathbf{S}_{\text{sim}}$ and $\mathbf{S}_{\text{dis}}$) Based on MLP-embedded features.
end while

Output: MLP-embedded features $\mathbf{F}$ (transform $\mathbf{X}$ into feature-embedded space using best $f_{\mathbf{\Theta}}(.)$).

1024 dimensions, the next crucial step involves dimensionality reduction and the application of time series k-means for effective clustering. Singular Value Decomposition (SVD) (Wall, Rechtsteiner, & Rocha, 2003; Furnas et al., 2017) is employed to reduce the dimensionality of these embeddings, enhancing computational efficiency and preserves the essential characteristics of the embedded features.

Upon completing the dimensionality reduction, we employ a sliding window technique to integrate the time series k-means algorithm, a pivotal step for clustering AE data streams that exhibit temporal dependencies. This method involves segmenting the reduced feature set into overlapping windows, allowing for the dynamic nature of AE data to be captured over time. The sliding window approach (SW) organizes the data into sequences of a specified window size. We empirically choose the SW size as 50, with a step size (1) dictating the overlap between consecutive windows. This structuring is essential for maintaining the temporal continuity of AE events, facilitating the identification of clusters that evolve over time.

By applying time series k-means (Huang et al., 2016) to these windowed sequences, we can effectively cluster AE events based on both their feature similarities and their temporal characteristics. This integration enables the detection of sequentially appearing clusters, a common occurrence in AE data streams, thereby providing deeper insights into the material's behavior and the efficacy of the monitoring system.

3. Experimentation and Results

3.1. Acoustic emission dataset Description

The ORION-AE dataset (Ramasso, Verdin, & Chevallier, 2022) was obtained through a test rig known as ORION. The ORION is specifically designed to mimic the loosening phenomena commonly observed in bolted joints of structures in various industries, including aeronautics, automotive, and civil engineering. It is composed of two metallic plates linked together by three M4 bolts (as shown in Figure 4, enabling the simulation of bolt loosening under vibrational stress.

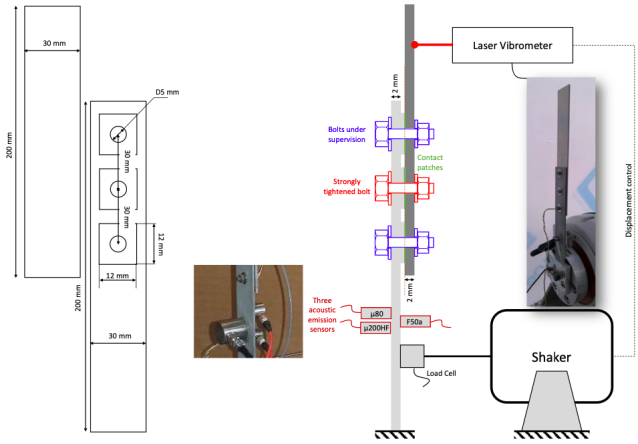

Figure 4. Setup description: part dimensions, sensors positions, bolts positions

The ORION-AE data are dynamically loaded with a vibration shaker and monitored with a laser vibrometer for velocity measurements and three AE sensors (micro80, F50A, micro200HF). The sensors sampled data at a rate of 5 MHz, producing datasets ranging from approximately 1.4 to 1.9 GB.

The ORION-AE dataset was generated by manually loosening a bolt on a test assembly and then subjecting it to 120 Hz harmonic vibrations, to simulate operational conditions. The experiment explored seven levels of bolt tightness (T1: 5cNm, T2: 10cNm, T3: 20cNm, T4: 30cNm, T5: 40cNm, T6: 50cNm, T7: 60cNm), with AE transients recorded for 10 seconds at each level. This procedure was repeated five times, resulting in five campaigns/datasets (B, C, D, E, and F), each with seven classes with 70 s of data for different sensors. Each campaign recorded varying numbers of signals, totaling 10,866; 9,461; 9,285; 15,628; and 17,810 signals, respectively. Note that, for campaign C, the level of bolt tightness 20 cNm is missing.

The seven tightening levels can be used as a ground truth when designing learning methods. This makes this dataset useful for developing and testing clustering and classification

methods for interpreting acoustic emission data.

For the purposes of this paper, analysis was focused exclusively on the micro-200-HF sensor, and only campaigns B, C and E were utilized to measure the performance of the clustering method. Figure 5 displays the tightening levels, acoustic emission and laser vibrometer data superimposed for measurements ”B” and sensor micro-200-HF (variable C).

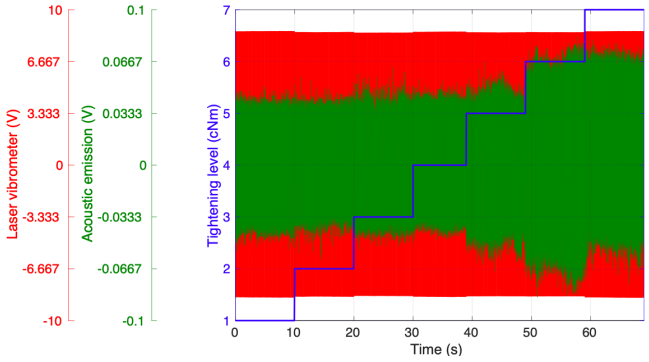

Figure 5. Tightening levels, acoustic emission and laser vibrometer data superimposed for measurements ”B” and sensor micro-200-HF (variable C).

3.2. Evaluation metrics

To properly evaluate the performance of clustering algorithms, such as TimeSeriesKMeans, on our test dataset, we use a variety of metrics. These metrics, as suggested by literature (Maulik & Bandyopadhyay, 2002), include:

- **Silhouette Score** evaluates cohesion within clusters and separation between them.
- **Davies-Bouldin Index** measures the average similarity between each cluster and its most similar cluster.
- **Adjusted Rand Index**, **Normalized Mutual Information (NMI)**, **Homogeneity**, **Completeness**, and **V-Measure** compare the clustering results to a ground truth, providing a measure of how well the clustering matches actual categories in the data.

3.3. Performance analysis

To demonstrate the effectiveness of the proposed unsupervised MLP-CMLL, we conduct numerous experiments to show the effectiveness of our embedded features compared to three different features, including raw data (AE signal Preprocessing & Feature Extraction), PCA (Kherif & Latypova, 2020) and SVD (Wall et al., 2003; Furnas et al., 2017). Tables 1, 2, and 3 show performance metrics for Campaigns B, C and E using different features over the TSKMeans cluster with sliding window.

Table 1. Performance metrics for Campaign B using different features over TSKMeans cluster with sliding window.

Method	ARI	Silhouette	DBI	NMI	Completeness
RAW	0.818	0.296	**1.394**	0.842	0.844
SVD	0.818	0.296	**1.394**	0.842	0.844
PCA	0.818	0.296	**1.394**	0.842	0.844
MLP-CMLL	**0.875**	**0.335**	1.278	**0.884**	**0.884**

Table 2. Performance metrics for Campaign C using different features over TSKMeans cluster with sliding window.

Method	ARI	Silhouette	DBI	NMI	Completeness
RAW	0.700	0.317	1.244	0.784	0.786
SVD	0.700	0.317	1.244	0.784	0.786
PCA	0.700	0.320	1.230	0.784	0.786
MLP-CMLL	**0.949**	**0.418**	**1.079**	**0.929**	**0.929**

Table 3. Performance metrics for Campaign E using different features over TSKMeans cluster with sliding window.

Method	ARI	Silhouette	DBI	NMI	Completeness
RAW	0.738	0.228	**1.743**	0.800	0.805
SVD	0.738	0.228	**1.743**	0.800	0.805
PCA	0.738	0.228	**1.743**	0.800	0.805
MLP-CMLL	**0.854**	**0.300**	1.700	**0.866**	**0.867**

In the three tables 1, 2, and 3, the consistent outperformance of MLP-CMLL across all campaigns underscores the potential of sophisticated neural network-based feature extraction methods in enhancing clustering performance. It suggests that MLP-CMLL can adaptively learn and highlight the most relevant features for clustering, outpacing traditional dimensionality reduction techniques in capturing the essential structures of various datasets. Furthermore, the relatively close performance of SVD, PCA, and RAW methods across the campaigns might reflect their limitations in dealing with complex data structures or their potential redundancy when the raw data is already amenable to effective clustering. Analyzing three campaigns using various feature extraction methods within a TSKMeans cluster with a sliding window approach reveals consistent trends across performance metrics. The MLP-CMLL method consistently outperforms the other methods (RAW, SVD, PCA) in all evaluated metrics — Adjusted Rand Index (ARI), Silhouette score, Davies-Bouldin Index (DBI), Normalized Mutual Information (NMI), and Completeness—indicating superior clustering effectiveness. The RAW, SVD, and PCA methods display nearly identical performance across most metrics and campaigns, suggesting similar capabilities in handling clustering tasks. MLP-CMLL's higher scores across all metrics highlight its ability to capture more complex patterns and nonlinearities that linear methods might miss, resulting in better-defined and more accurately clustered data groups. This underlines the importance of method selection in data clustering to achieve optimal re-

Table 4. Performance metrics of different clustering methods for Campaign B using MLP-CMLL embedded features with (w/) and without (w/o) sliding window.

Method	ARI	Silhouette	DBI	NMI	Completeness
GMM (w/o)	0.676	0.406	1.007	0.810	0.809
Kmeans (w/o)	0.623	**0.436**	1.000	0.786	0.795
MLP-CMLL+TSKMeans (w/o)	0.657	0.410	0.997	0.796	0.795
MLP-CMLL+TSKMeans (w/)	**0.875**	0.335	**1.278**	**0.884**	**0.884**

Table 5. Performance metrics of different clustering methods for Campaign C using MLP-CMLL embedded features with (w/) and without (w/o) sliding window.

Method	ARI	Silhouette	DBI	NMI	Completeness
GMM (w/o)	**0.975**	0.420	1.050	**0.962**	**0.962**
Kmeans (w/o)	0.919	0.423	1.036	0.904	0.904
MLP-CMLL+TSKMeans (w/o)	0.919	**0.424**	1.037	0.904	0.905
MLP-CMLL+TSKMeans (w/)	0.949	0.418	**1.079**	0.929	0.929

sults based on specific campaign characteristics and objectives. Therefore, these observations suggest that while traditional methods like SVD and PCA have their merits, especially in contexts where computational simplicity and interpretability are key, advanced neural network-based approaches like MLP-CMLL offer a promising avenue for tackling more complex clustering challenges. Future work could explore further optimizations of the MLP-CMLL architecture, comparisons with other advanced machine learning techniques, and applications to a broader range of data types and clustering scenarios.

For Campaign B, MLP-CMLL shows the best performance across almost all metrics, highlighting its ability to extract meaningful embedded features that contribute to effective clustering. This suggests that the MLP-CMLL approach, with its presumably more nuanced understanding of the data structure, is particularly well-suited for the types of datasets represented in Campaign B. RAW, SVD, and PCA show similar performance in terms of ARI, Silhouette score, and other metrics. This could indicate that for Campaign B's dataset, the simpler dimensionality reduction techniques (SVD and PCA) do not provide significant advantages over using RAW data. This might be due to the nature of the data where the intrinsic data structure is either too complex for simple linear transformations to capture or perhaps is already in a form where raw data clustering is relatively effective.

For Campaign C, MLP-CMLL again outperforms other methods significantly in ARI and Completeness, reinforcing the value of advanced feature extraction methods in improving clustering outcomes. The improvement in the Silhouette score and DBI suggests that MLP-CMLL leads to more distinct, well-separated clusters than other methods. The performance gap between MLP-CMLL and other methods (SVD, PCA, and RAW) is notable, especially in terms of ARI and Completeness. This could imply that the Campaign C dataset contains complex patterns or high-dimensional structures that are better captured by the MLP-CMLL's feature extraction capabilities.

For Campaign E, MLP-CMLL's superiority is evident but less pronounced compared to Campaign C. It still leads in Adjusted Rand Index and Completeness, indicating its consistent effectiveness across different datasets. The similarity in performance between SVD, PCA, and RAW methods suggests that for Campaign E's data, the simple dimensionality reduction does not significantly impact the clustering performance, similar to Campaign B. However, the overall lower scores compared to Campaign B could indicate that Campaign E's dataset is inherently more challenging to cluster effectively, possibly due to noise, less distinct groupings, or more complex data structures.

Tables 4, 5, and 6 show performance metrics using different clustering methods for Campaigns B, C, and E with (w/) and without (w/o) sliding window.

Table 4 shows the performance metrics for Campaign B. The performance metrics for Campaign B provide a nuanced view of algorithm effectiveness. The Gaussian Mixture Model (GMM) showcases strong performance with an ARI of 0.676, suggesting a high degree of accuracy in clustering with respect to the true classifications. This is supported by an NMI of 0.810 and a Completeness score of 0.809, indicating a robust alignment between cluster assignments and actual data labels. The introduction of a sliding window with Time Series K-Means enhances its performance significantly, as evidenced by a jump in ARI to 0.875 and NMI to 0.884, underscoring the method's ability to capture temporal dependencies within the data. The Silhouette Score and DBI provide additional insights; despite a lower Silhouette Score (0.335) with the sliding window, indicating less clear separation between clusters, the method's overall effectiveness is not notably diminished, suggesting that the sliding window compensates by capturing temporal patterns not evident in spatial metrics alone.

Table 5 shows the performance metrics for Campaign C. Campaign C's analysis reveals the exceptional capability of the

Table 6. Performance metrics of different clustering methods for Campaign E using MLP-CMLL embedded features with (w/) and without (w/o) sliding window.

Method	ARI	Silhouette	DBI	NMI	Completeness
GMM (w/o)	0.626	0.311	1.453	0.716	0.718
Kmeans (w/o)	0.546	0.336	1.284	0.642	0.644
MLP-CMLL+TSKMeans (w/o)	0.598	**0.341**	1.371	0.671	0.673
MLP-CMLL+TSKMeans (w/)	**0.854**	0.300	**1.700**	**0.866**	**0.867**

GMM algorithm, achieving near-perfect ARI (0.975) and NMI (0.962) scores, which imply an almost flawless clustering outcome compared to true labels. This campaign highlights the impact of using a sliding window with Time Series K-Means, which achieves an ARI of 0.949 and an NMI of 0.929. These results suggest that the temporal structure captured by the sliding window significantly enhances clustering fidelity. The Silhouette Score (0.418 with the sliding window) and DBI (1.079 with the sliding window) indicate a balance between cluster cohesion and separation, affirming the effectiveness of incorporating temporal context in clustering analysis.

Table 6 shows the performance metrics for Campaign E. In Campaign E, the stark contrast in performance metrics between methods with and without sliding windows becomes even more pronounced. The use of the sliding window with Time Series K-Means propels its ARI to 0.854 and NMI to 0.866, suggesting a high degree of clustering accuracy that leverages temporal information effectively. Despite a lower Silhouette Score (0.300) with the sliding window, indicating potential overlap among clusters, the high NMI and Completeness scores (0.866 and 0.867, respectively) with the sliding window imply a successful capture of the intrinsic data structure. This campaign showcases the critical role of temporal analysis in clustering, especially for data where temporal patterns significantly influence the underlying structure.

Across campaigns B, C, and E, the analysis underscores the nuanced performance of GMM and Time Series K-Means, particularly when enhanced with a sliding window technique, across various clustering quality metrics. While simpler algorithms like Kmeans show competitive performance in specific metrics such as the Silhouette Score, the added complexity and temporal awareness of the sliding window modification in Time Series K-Means generally translate into superior clustering outcomes, especially in terms of aligning with true cluster structures and maintaining class completeness.

Advantages of MLP-CMLL Time Series K-Means (MLP-CMLL with TSKMeans)

From all tables 4, 5 and 6, the introduction of Contrastive Metric Learning Loss-Enhanced Multi-Layer Perceptron with Time Series K-Means (MLP-CMLL with TSKMeans) marks a significant advancement in clustering complex time-series data. This novel approach leverages the strength of contrastive learning to fine-tune the feature representation, significantly enhancing the clustering capability of TSKMeans by ensuring that similar instances are brought closer while dissimilar ones are distanced in the feature space. Our results underscore the efficacy of this method, particularly in achieving superior clustering performance metrics across all campaigns when compared to traditional approaches. Notably, the MLP-CMLL with TSKMeans exhibits remarkable improvements in metrics such as ARI and NMI, indicating not only an enhanced alignment with the true cluster structures but also a comprehensive capture of the intrinsic data relationships. This methodological enhancement introduces a powerful tool for time-series analysis, offering robustness against the challenges posed by the dynamic nature of temporal data and paving the way for more accurate, interpretable clustering solutions.

Our MLP-CMLL with TSKMeans vs. GMMSEQ (Ramasso, Denoeux, & Chevallier, 2022). In a comparative analysis between the novel MLP-CMLL+TSKmeans method and the GMMSEQ (Ramasso, Denoeux, & Chevallier, 2022) method across three experimental campaigns labeled B, C, and E, the performance is quantitatively measured using the Adjusted Rand Index (ARI). The ARI scores indicate the similarity between the clustering results and the true classifications, with a range from -1 to 1, where 1 denotes perfect agreement. For Campaign B, the MLP-CMLL+TSKmeans method significantly outperforms GMMSEQ, achieving an ARI of 0.875 compared to GMMSEQ's 0.772. This suggests a superior ability of the MLP-CMLL+TSKmeans to accurately match the true cluster structures. In Campaign C, both methods exhibit exceptional performance with MLP-CMLL+TSKmeans slightly leading (0.949 vs. 0.947), indicating that both are very capable but MLP-CMLL+TSKmeans shows a slight edge in capturing the clustering structure accurately. Campaign E again sees MLP-CMLL+TSKmeans outperforming GMMSEQ (0.854 vs. 0.799), reinforcing the method's robustness and accuracy in analyzing the complex dynamics of acoustic emission data streams. Overall, MLP-CMLL+TSKmeans consistently surpasses GMMSEQ in clustering performance across all campaigns, evidencing its effectiveness and the significant benefits it offers for structural health monitoring applications through better differentiation and handling of temporal dynamics within AE data.

4. CONCLUSION

This work introduced a new method for the analysis of acoustic emission (AE) data streams, which are inherently sequen-

tial and temporal. The study proposes a unique approach by enhancing a Multi-Layer Perceptron (MLP) with a contrastive metric learning loss function (MLP-CMLL) and time series kmeans, aiming to efficiently identify and analyze sequentially appearing clusters within the data. This novel loss function is meticulously designed to optimize the MLP by improving the differentiation between distinct clusters. The approach primarily concentrates on embedding sequences in a manner that clusters with similar acoustic patterns are brought closer together, while those with divergent patterns are distanced, thereby augmenting the MLP's capability to recognize and classify acoustic events based on their emission signatures over time.

The importance of this work lies in its ability to address the challenges associated with the precise characterization of dynamically forming clusters within AE data streams. Traditional clustering algorithms often falter in handling the temporal dynamics of AE data, where the sequencing and timing of events are crucial for a comprehensive understanding of the phenomena being monitored. By integrating a contrastive metric learning loss with an MLP architecture tailored to the specifics of sequentially appearing clusters in AE data streams, our method aims to unveil deeper insights into the formation and evolution of clusters. This approach promises to enhance monitoring and predictive maintenance in engineering applications by capturing the complex dynamics of AE data more effectively.

Through extensive experimentation and comparative analysis against conventional techniques, we validate the superiority of our proposed method in discerning the intricate dynamics of AE data. This work presents a robust analytical tool for the investigation of sequential clusters and their implications in the domain of structural health monitoring, offering significant advancements over existing methods in terms of cluster detection, characterization, and temporal analysis.

REFERENCES

Bolognani, D., Verzobio, A., Tonelli, D., Cappello, C., Glisic, B., Zonta, D., & Quigley, J. (2018). Iwshm 2017: Quantifying the benefit of structural health monitoring: what if the manager is not the owner? *Structural Health Monitoring*, *17*(6), 1393–1409.

Fu, W., Zhou, R., & Guo, Z. (2023). Automatic bolt tightness detection using acoustic emission and deep learning. In *Structures* (Vol. 55, pp. 1774–1782).

Furnas, G. W., Deerwester, S., Durnais, S. T., Landauer, T. K., Harshman, R. A., Streeter, L. A., & Lochbaum, K. E. (2017). Information retrieval using a singular value decomposition model of latent semantic structure. In *Acm sigir forum* (Vol. 51, pp. 90–105).

Gonzalez Andino, S., Grave de Peralta Menendez, R., Thut, G., Spinelli, L., Blanke, O., Michel, C., & Landis, T. (2000). Measuring the complexity of time series: an application to neurophysiological signals. *Human brain mapping*, *11*(1), 46–57.

Hassani, K., & Khasahmadi, A. H. (2020). Contrastive multi-view representation learning on graphs. In *International conference on machine learning* (pp. 4116–4126).

Hoła, J., & Sadowski, Ł. (2022). *Non-destructive testing in civil engineering* (Vol. 12) (No. 14). MDPI.

Huang, X., Ye, Y., Xiong, L., Lau, R. Y., Jiang, N., & Wang, S. (2016). Time series k-means: A new k-means type smooth subspace clustering for time series data. *Information Sciences*, *367-368*, 1-13. doi: https://doi.org/10.1016/j.ins.2016.05.040

Kharrat, M., Ramasso, E., Placet, V., & Boubakar, M. (2016). A signal processing approach for enhanced acoustic emission data analysis in high activity systems: Application to organic matrix composites. *Mechanical Systems and Signal Processing*, *70*, 1038–1055.

Kherif, F., & Latypova, A. (2020). Principal component analysis. In *Machine learning* (pp. 209–225). Elsevier.

Liu, G., Lin, Z., & Yu, Y. (2010). Robust subspace segmentation by low-rank representation. In *Proceedings of the 27th international conference on machine learning (icml-10)* (pp. 663–670).

Maulik, U., & Bandyopadhyay, S. (2002). Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on pattern analysis and machine intelligence*, *24*(12), 1650–1654.

Ramasso, E., Denoeux, T., & Chevallier, G. (2022). Clustering acoustic emission data streams with sequentially appearing clusters using mixture models. *Mechanical Systems and Signal Processing*, *181*, 109504.

Ramasso, E., Placet, V., & Boubakar, M. L. (2015). Unsupervised consensus clustering of acoustic emission time-series for robust damage sequence estimation in composites. *IEEE Transactions on Instrumentation and Measurement*, *64*(12), 3297–3307.

Ramasso, E., Verdin, B., & Chevallier, G. (2022). Monitoring a bolted vibrating structure using multiple acoustic emission sensors: A benchmark. *Data*, *7*(3), 31.

Saunshi, N., Plevrakis, O., Arora, S., Khodak, M., & Khandeparkar, H. (2019). A theoretical analysis of contrastive unsupervised representation learning. In *International conference on machine learning* (pp. 5628–5637).

Sause, M. G., Gribov, A., Unwin, A. R., & Horn, S. (2012). Pattern recognition approach to identify natural clusters of acoustic emission signals. *Pattern Recognition Letters*, *33*(1), 17–23.

Sun, J., Yang, H., Li, D., & Xu, C. (2023). Experimental investigation on acoustic emission in fretting friction and wear of bolted joints. *Journal of Sound and Vibration*, *558*, 117773. doi: https://doi.org/10.1016/j.jsv.2023.117773

Swedish Accident Investigation Authority. (2017). *Windactionjvestas wind turbine collapse in lemnhult.*

Wall, M. E., Rechtsteiner, A., & Rocha, L. M. (2003). Singular value decomposition and principal component analysis. In *A practical approach to microarray data analysis* (pp. 91–109). Springer.

Wang, T., Song, G., Wang, Z., & Li, Y. (2013). Proof-of-concept study of monitoring bolt connection status using a piezoelectric based active sensing method. *Smart Materials and Structures*, *22*(8), 087001.

Xu, D., Liu, P., Li, J., & Chen, Z. (2019). Damage mode identification of adhesive composite joints under hygrothermal environment using acoustic emission and machine learning. *Composite structures*, *211*, 351–363.

Xu, P., Zhou, Z., Liu, T., & Mal, A. (2021). Determination of geometric role and damage assessment in hybrid fiber metal laminate (fml) joints based on acoustic emission. *Composite Structures*, *270*, 114068. doi: https://doi.org/10.1016/j.compstruct.2021.114068

Zhu, H., & Koniusz, P. (2022a). Ease: Unsupervised discriminant subspace learning for transductive few-shot learning. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 9078–9088).

Zhu, H., & Koniusz, P. (2022b). Generalized laplacian eigenmaps. *Advances in Neural Information Processing Systems*, *35*, 30783–30797.