An Isolation Forest Approach for Robust Anomaly Detection in Industrial Machines Using Out-of-Distribution Acoustic Data

Cristofer Silva*, João Campos†, Leonardo Afonso Ferreira Bortoni‡, Pablo Andretta Jaskowiak‡, Diego Pinheiro†
*Universidade de Pernambuco (UPE), Recife, Brazil
cgss@ecomp.poli.br
†Universidade Católica de Pernambuco (UNICAP), Recife, Brazil
joao.2021104704@unicap.br, diego.silva@unicap.br
‡Universidade Federal de Santa Catarina (UFSC), Joinville, Brazil
leonardo.bortoni@posgrad.ufsc.br, pablo.andretta@ufsc.br

***Abstract*—Anomaly detection, essential for predictive maintenance as Industry 4.0 advances, has become paramount in industrial environments, enabling early fault detection, reducing potential financial losses, and mitigating safety risks. Deep learning methods have attracted more attention for anomaly detection in industrial machines, given their convenient capability of automatic feature extraction. This convenience, however, often leads us to overlook simpler and more explainable models. Despite this consequence, the use and creation of simpler and more explainable machine learning pipelines are underexplored and underanalyzed compared to the pre-considered use of deep learning, especially in *out-of-distribution* scenarios. We hypothesized that a simple and explainable model using handcrafted extracted features can be at least non-inferior to deep learning models for anomaly detection in multivariate time series from industrial machinery. To validate the hypothesis, we compared the *Isolation Forest*—a simple and explainable model—combined with Mel-Frequency Cepstral Coefficients (MFCCs) as handcrafted features—with state-of-the-art deep learning models. Furthermore, we employed the *Malfunctioning Industrial Machine Investigation and Inspection* (MIMII) acoustic dataset, which provides sounds from valves, pumps, fans, and slide rails under normal and faulty conditions. The proposed methodology, based on Isolation Forest, is compared against state-of-the-art approaches, already implemented within the MTSA framework, namely, Hitachi, GANF, and RANSynCoders, considering two distinct scenarios: *in-distribution* (ID) and *out-of-distribution* (OOD). While *in-distribution* scenarios allow us to analyze the performance of machine learning models on data with distributions that almost mirror those of the training data, *out-of-distribution* scenarios allow us to go further and analyze the performance of models on data with distributions that deviate significantly from those of the training data, which brings us closer to the real-world application of the machine learning models. Our results support that Isolation Forest based on MFCCs is non-inferior to deep learning models. This showcases the effectiveness of a less complex method based on handcrafted features in industrial acoustic data and its capability of dealing with real-world scenarios.**

***Index Terms*—anomaly detection, multivariate time series, acoustic data, MFCC, isolation forest, out-of-distribution**

979-8-3503-7457-5/24/$31.00 ©2024 IEEE

I. Introduction

Anomaly detection can be defined as the process of recognizing patterns that deviate from those expected within a particular application, scenario, or dataset [1]. Research in this area is constantly evolving due to its importance in various contexts, such as medicine [2], security [3], and industry [4].

In the industrial context, operational anomalies serve as an indicator of potential machine failure. If left unattended, such failures might result in life-threatening situations to workers [5] and substantial financial losses [6]. Even though these risks can be reduced with predictive maintenance, reliance on human inspection usually results in a high dependency on experienced professionals, whose demand often exceeds availability [7]. Not surprisingly, industrial automation has gained prominence alongside machine learning [8], allowing not only for equipment monitoring, but also the identification of anomalous behaviors before critical outcomes.

As pointed out by Saufi et al. [9], the first manifestations of anomalous machinery behavior occur through acoustic emissions. Due to the high dimensionality of multivariate acoustic time-series, state-of-the-art anomaly detection approaches in this context often adopt increasingly complex pipelines [8]. Nonetheless, the absence of standardization and design patterns makes it difficult to train, evaluate, and compare models outside the environments in which they were originally developed, aggravating the reproducibility crisis [10].

Motivated by these observations, in this work, we explore anomaly detection in industrial machines through acoustic data, with a focus on the reproducibility and effectiveness of a simpler model. Regarding the reproducibility aspect, we adopt the framework MTSA (Multiple Time Series Analysis) [11], a framework for anomaly detection, which facilitates the development of machine learning pipelines and enables the statistical and comparative analysis of state-of-the-art approaches. Using MTSA, we implemented our approach in the framework and employed a feature extraction pipeline that uses Mel Frequency Cepstral Coefficients (MFCCs)—a standard for acoustic data [12], [13].

Contrary to the prevailing trend of increasing complexity in anomaly detection models, we also integrate into MTSA Isolation Forest [14], a simple yet effective anomaly detection method. This allows us to fairly investigate and compare its performance to that of complex methods, all based on features derived from MFCCs. The state-of-the-art competitors taken into account are Hitachi [7], GANF [15], and RANSynCoders [16]. These are also available in the MTSA framework.

Through a systematic experimental evaluation considering the MIMII Benchmark Dataset [7], our results demonstrate that Isolation Forest achieves performance comparable to that of more complex methods. Moreover, we find that Isolation Forest is less sensitive to variations in hyperparameters, demonstrating great stability compared to the other state-of-the-art models. This enhanced stability reinforces the robustness of Isolation Forest, ensuring consistent results without extensive hyperparameter tuning.

Finally, we take into account both in-distribution (ID) and out-of-distribution (OOD) scenarios during evaluation. These are essential to determine the generalization and robustness of a model in anomaly detection. In brief, in-distribution accounts for model evaluation within data from the same distribution adopted during training (e.g., the same machine type). Conversely, out-of-distribution employs data that differs considerably (e.g., a different machine type) from that of training during evaluation. One might argue that good performance on in-distribution data is the bare minimum expected and, therefore, a model truly stands out when it remains effective on out-of-distribution data, demonstrating robustness and the ability to generalize beyond training conditions.

The work is organized as follows. In Section II we discuss related work, with focus on the competitor methods. In Section III, we provide a review of Isolation Forest. In Section IV we introduce the experimental setup, whereas in Section V we present and discuss the main findings of the work. Section VI provides the conclusions and possible future work directions.

II. Related work

A multivariate time series $\mathbf{X} = \{X^1, \ldots, X^n\}$ is a set of n univariate time-series, $X^i = (X^i_1, \ldots, X^i_m)$, each of which comprised of m of measurements. Measurements across series are within the same time frame and, ideally, taken at the very same time steps, allowing for the characterization of complex behaviors that would otherwise be overlooked if each series were interpreted or analyzed independently. In the context of industrial anomaly detection, variables are usually related to sensor readings from machines that record information such as vibration, temperature, and acoustic emissions [17]. This data characterizes the operational state of a machine.

Given that anomalous acoustic emissions are among the first manifestations of machinery abnormalities, an increasing number of works have focused on multivariate time series consisting exclusively of acoustic data. To address their complexity, machine learning models have been widely explored, usually considering a reduced feature space derived from Mel Frequency Cepstral Coefficients (MFCCs) [12]. In the following, we describe three state-of-the-art approaches from the literature that are considered in our work.

A. Hitachi

The method referred to here as Hitachi performs anomaly detection using an Autoencoder that takes log-Mel spectrograms as input [7]. Despite their compact representation of acoustic signals, log-Mel spectrograms effectively capture relevant features for identifying anomalous patterns.

The Autoencoder is composed of fully connected layers, configured to extract and reconstruct representative data patterns. During training, the model is optimized to minimize the error between the input and the output (reconstructed input), thereby learning to efficiently represent normal data. In the detection phase, objects that result in a high reconstruction error are considered anomalous, as they significantly deviate from the underlying phenomenon represented by the model. The encoder consists of three layers: $FC(\text{Input}, 64, \text{ReLU})$, $FC(64, 64, \text{ReLU})$, and $FC(64, 8, \text{ReLU})$, where $FC(a, b, f)$ denotes a fully connected layer with a input neurons, b output neurons, and activation function f. The decoder mirrors this structure with layers $FC(8, 64, \text{ReLU})$, $FC(64, 64, \text{ReLU})$, and $FC(64, \text{Output}, \text{None})$, reconstructing the given input.

B. GANF

The Graph-Augmented Normalizing Flows (GANF) [15] is an approach for anomaly detection in multivariate time series that employs Graph Neural Networks (GNNs). The model combines a Bayesian Network—a directed acyclic graph that models the dependencies among time series—with LSTM Autoencoders to capture the temporal dependencies of each time series in an individual fashion.

After modeling the inter-series dependencies and learning the temporal dependencies, GANF estimates the probability density of the data using Normalizing Flows. The Normalizing Flows are models that transform a simple distribution into a more complex one through invertible and differentiable transformations, enabling flexible modeling of the normal data distribution. The GANF model identifies as anomalies those data points with low probability under the learned probability density distribution.

C. RANSynCoders

RANSynCoders [16] is an approach for anomaly detection in multivariate time series that employs a pipeline composed of several sequential data processing steps. Initially, dimensionality reduction is performed using an autoencoder, enabling data representation in a compressed yet relevant feature space. Next, the method estimates the latent spectral density using the Fast Fourier Transform (FFT), identifying the dominant frequencies in the time series. These frequencies act as prior information to construct a synchronized representation of the data, in which asynchronous time series are transformed into a unified representation. After synchronization, the bootstrapping method is applied to reconstruct the upper and lower bounds of the multivariate distribution for each encoder. During the detection phase, anomaly detection is carried out using majority voting among the bootstrap encoders previously discussed. Thus, observations that deviate significantly from the learned patterns are considered anomalous.

III. *Isolation Forest*

Isolation Forest [14] is a method for anomaly detection based on the isolation of anomalous observations, rather than the modeling of normal ones. In brief, the method builds a forest of binary trees (iTrees). In order to obtain a diverse forest, a sub-sampling strategy (without replacement) is employed to select objects before the induction of each iTree. Once a subset of objects has been obtained, an iTree can be obtained by recursively partitioning objects based on a randomly selected feature. As proposed by its authors [14], the partitioning process of an iTree terminates once a maximum desired height is reached, a node with a single object is obtained, or all objects within the node are identical.

The method is based on the premise that anomalies are easier to separate from regular data points. In the context of the random trees, this means that anomalies are isolated in shallower depths than regular objects, on average. The idea is exemplified by Fig. 1, for a 2D toy dataset.

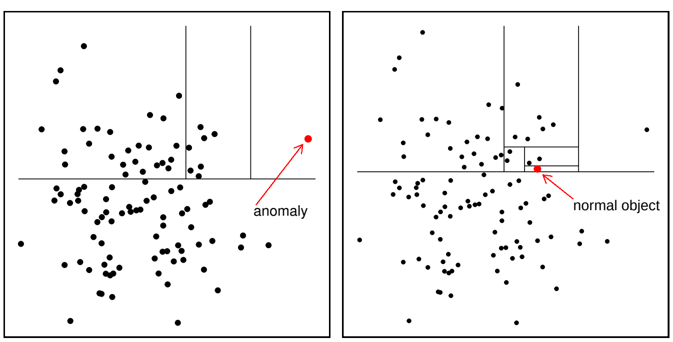

Fig. 1. Anomalies are often isolated after just a few random partitions of the data (left). Normal objects, on the other hand, typically require a larger number of intricate nested partitions to be isolated (right). Adapted from [14].

Taking that into account and considering an Isolation Forest derived from a dataset with n objects, The path length $h_i(x)$ of an object x in a given isolation tree (iTree) i within the Isolation Forest is the number of edges x traverses from the root to a leaf node. The anomaly score $s(x,n)$ of an object x in an isolation forest trained with n objects is given by the normalized average over all iTrees as follows:

$$s(x,n) = 2^{-\frac{E[h_i(x)]}{c(n)}} \quad , \qquad (1)$$

in which $E[h_i(x)]$ is the average path length that x takes in all iTrees within the Isolation Forest. The term $c(n)$ is used as a normalization factor and is the average path length of a binary tree with n nodes (refer to [14] for complete formulation).

In brief, values of $s(x,n)$ close to 1 can be regarded as anomalies, whereas values smaller than 0.5 can be safely considered normal observations, according to the method.

IV. Experimental Setup

A. *Dataset*

Our experiments were performed using the MIMII database [7], which comprises acoustic signals of industrial machines operating under normal and abnormal conditions. The database includes recordings of four machine types: valve, pump, fan, and slide rail. Within each machine type, there are seven distinct machine instances (distinguished by IDs), which account for different models of that machine type. Each acoustic signal was recorded for 10 seconds at 16kHz by a TAMAGO-03 microphone that was placed 50cm from the machines (except for the valve machine type, for which the distance was 10cm). This particular microphone has eight channels; therefore, one recording session produces eight acoustic signals as a result. In this work, only data from the microphone in channel zero was employed. Table I presents an overview of the acoustic segments of the dataset employed.

TABLE I
Distribution for normal and abnormal data regarding the four distinct machine types from the MIMII database.

Machine Type	Normal	Anomalous
Valve	991	119
Pump	1,006	143
Fan	1,011	407
Slide	1,068	356
Total	4,076	1,025

B. *Mel-Frequency Cepstral Coefficients (MFCC)*

Before applying any of the anomaly detection methods, features were extracted based on the Mel-Frequency Cepstral Coefficients (MFCCs) from the raw acoustic signals. MFCCs feature extraction is widely employed in acoustic data, as it reduces the signal dimensionality while preserving meaningful information. It has been employed for music information retrieval [18], rail track fault detection [19], and analysis of vocal cord signals for vocal disorder [20], just to mention a few examples. Given its compelling characteristics, the adoption of MFCCs as a feature extraction technique leverages shallow machine learning models, such as *Isolation Forest*, to perform anomaly detection on time series data. The technique extracts the signal's mel-spectrogram, which is the result of the Short-Time Fourier Transform processed by mel-filter banks. The Discrete Cosine Transform is then applied to the mel-spectrogram. The result can be regarded as the spectrum of a spectrum, yielding what is termed as *Cepstrum*. Finally, a low-pass filter extracts the most relevant harmonics from the MFCCs matrix. This final matrix comprises stacked cepstrum coefficients into time bins. For details, see [21]–[23].

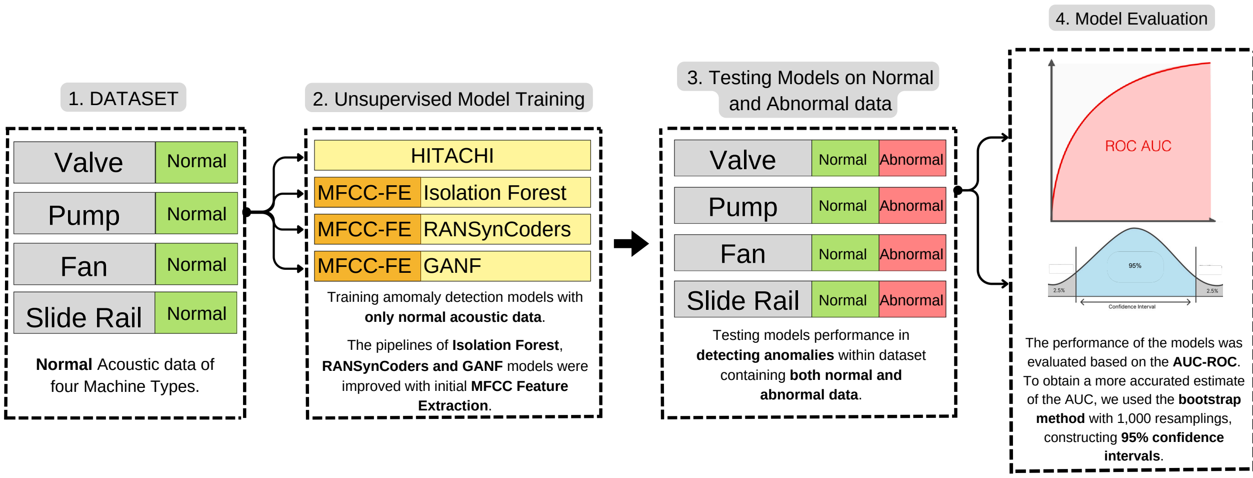

Fig. 2. Overview of the evaluation procedure.

C. Hyperparameter Tunning

Given its simplicity and low training and testing time, hyperparameters for the *Isolation Forest* were tuned for each machine type. The hyperparameters considered were *n_estimators*, which accounts for the number of iTrees in the forest; *max_samples*, which is the number of samples randomly extracted from the dataset to build a tree (small values may improve tree diversity, but might jeopardize generalization capacity); and *max_features*, which defines the number of features randomly selected to build each iTree. This value is given as a percentage of the total amount of features available in the dataset. Table II shows the best hyperparameters for the Isolation Forest method as adjusted for each MIMII machine type according to its performance at the validation set.

TABLE II
ISOLATION FOREST HYPERPARAMETERS ACCORDING TO MACHINE TYPE.

Machine Type	*n_estimators*	*max_samples*	*max_features*
Valve	100	512	1.0
Pump	70	128	1.0
Fan	100	256	1.0
Slider	10	128	1.0

D. Training and Testing

The training and testing processes followed usual practices from the unsupervised anomaly detection literature. All anomalous acoustic segments were reserved for testing, with an equal number of normal acoustic segments selected at random to compose the test set. The remaining normal segments were used for training the models, thus reproducing a real-world application scenario. In order to obtain sound estimates of evaluation metrics, we carried out a 5-fold cross-validation.

In-distribution (ID) and out-of-distribution (OOD) experiments were conducted. The MIMII dataset is comprised of four machine types (valve, pump, fan, and slide rail), each of which includes four distinct product models (given by *ids 00*, *02*, *04*, and *06*). Thus, there is a total of 16 machine-model pairs. In the ID setting, both training and testing were performed using data from the very same machine model. Conversely, in the OOD setting, models were trained considering a specific machine model (*id-02*) and tested on a different one (*id-00*). This approach simulates a scenario in which the machine learning model needs to generalize to equipment not seen during training, a usual condition faced in industrial environments.

The anomaly detection models were evaluated in the test sets by the Area Under the ROC Curve (AUC-ROC) metric, in order to assess their capabilities in distinguishing anomalous acoustic segments from normal ones. Since the 5-fold cross-validation yields five trained models, it is possible to perform a bootstrap of their corresponding ROC-AUC values and compute their confidence intervals. The bootstrap procedure was based on 1,000 ROC-AUC samples from the five ROC-AUC values, which allowed us to determine a 95% confidence interval. Fig. 2 gives an overview of the evaluation pipeline.

All anomaly detection approaches considered in this work, namely Isolation Forest, Hitachi, RANSynCoders, and GANF, were evaluated using the MTSA framework, allowing for reproducibility and fair evaluation of the methods.

V. RESULTS

All methods were evaluated considering ID and OOD scenarios, considering four different machine types, namely valve, pump, fan, and slide rail. Our first analysis, however, considers all these experimental factors in an aggregated fashion. These results are depicted in Fig. 3, in which bars account for average ROC-AUC values across different evaluations and error bars account for standard deviation. On average, the best results were obtained with Isolation Forest (0.82), followed by Hitachi (0.71), GANF (0.67), and RANSynCoders (0.63).

In addition to a better overall performance, considering the AUC-ROC metric, Isolation Forest proves to be an even more competitive approach if we consider the training time of this model, which is considerably lower than that of the complex approaches. Fig. 4 shows the confidence interval for the overall average training time for each model.

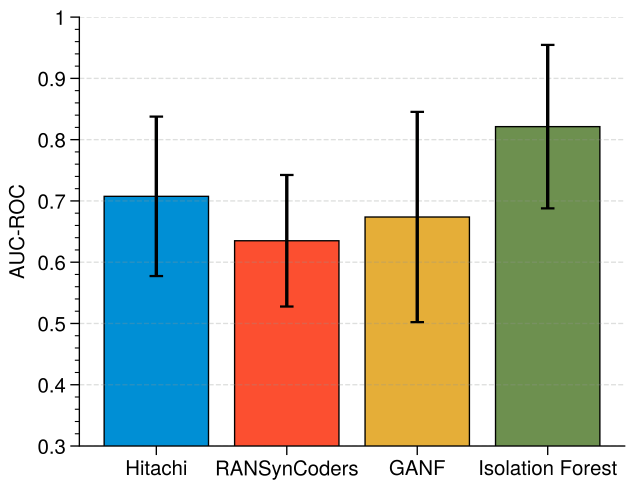

Fig. 3. Aggregated results for all evaluation factors.

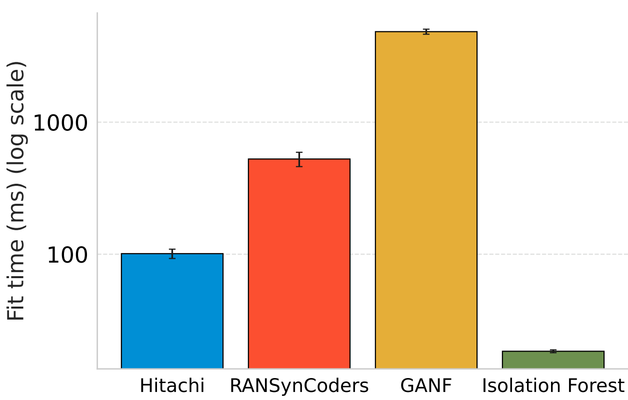

Fig. 4. Aggregated results for fit time (milliseconds).

The parameters for the models Hitachi, RANSynCoders, and GANF were chosen based on previous publications that compared these three models [24]. For Isolation Forest, some tuning of hyperparameters was performed, even though not systematically; it was very useful for finding critical values. A more systematic Hyperparameter tuning can be efficiently carried out for Isolation Forest, due to its simplicity and low computation cost, which is not observed in the competitor methods. In brief, we argue that, if computational resources are scarce, it is better to employ a simpler method with some hyperparameter tuning than a more complex one with default values.

Apart from this overview, it is possible to take into consideration the results for different evaluation factors separately. These are shown in Fig. 5, which depicts the AUC-ROC results for each model under evaluation, considering all machine types (valve, pump, fan, and slide rail) and both evaluation scenarios (ID and OOD). Bar values account for the average value of the ROC-AUC, whereas error bars account for the confidence interval estimated with the bootstrap procedure.

It is possible to contrast both in-distribution (ID) and out-of-distribution (OOD) scenarios to better understand model robustness. While all models demonstrated reasonably high performance under ID conditions, a clear drop in AUC-ROC scores is observed in the OOD scenario for most methods. This gap highlights the challenge of generalizing beyond the data seen during training. Notably, Isolation Forest consistently maintained more stable performance across both scenarios, showing only a modest decrease under OOD conditions. In contrast, complex models like GANF and RANSynCoders exhibited sharper performance declines—especially on machine types such as valve and pump—indicating higher sensitivity to distribution shifts. These results reinforce the notion that model complexity does not guarantee generalization, and that simpler models like Isolation Forest can offer more robust behavior when deployed in the real world with unseen variations.

Even though data for the four distinct machine types were collected within the same setup, it becomes evident from Fig. 5 that the difficulty of detecting anomalies varies for each machine, with valve and fan arising as the most challenging ones. Indeed, in the case of fan, the best results in terms of AUC-ROC were around 0.65, with little distinction in performance for the methods between ID and OOD scenarios.

Table III presents the results obtained using the best set of hyperparameters found under the in-distribution condition for each model. A performance drop is observed for Isolation Forest when applied to out-of-distribution data, as reflected by the decrease in AUC-ROC values compared to the tests conducted on in-distribution data. Despite this reduction, Isolation Forest maintained a reasonable performance, demonstrating its robustness and generalization capability even when exposed to data from different machines.

A. *Remarks on out-of-distribution evaluation*

The results presented in Table III show that Isolation Forest maintained a more consistent performance in out-of-distribution scenarios than state-of-the-art models. Although Isolation Forest exhibited a moderate drop in AUC-ROC values, GANF and RANSynCoders models experienced more significant decreases, particularly for valve and pump. Similarly, the Hitachi model also showed a pronounced reduction in ROC-AUC values, especially for the pump and slide rail.

These results indicate that, despite its simplicity, Isolation Forest can be more robust when handling data outside the training distribution. Its lower complexity and reduced reliance on fine-tuning of hyperparameters contribute to a more

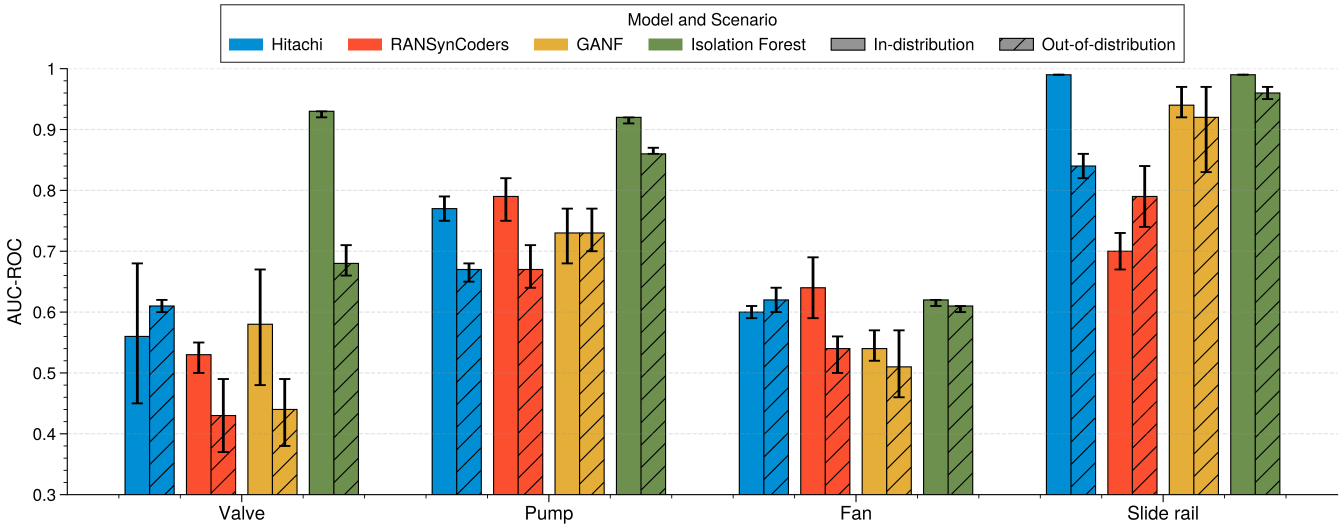

Fig. 5. In-distribution and Out-of-distribution bootstrap AUC values of Hitachi, RANSynCoders, GANF, and Isolation Forest approaches. The error bars stand for a 95% confidence interval.

effective generalization in real-world scenarios. In contrast, more complex models such as GANF and RANSynCoders showed greater sensitivity to data variations, resulting in poorer performance under out-of-distribution conditions.

VI. Conclusions

Anomaly detection plays a crucial role in the predictive maintenance of industrial machines, helping to mitigate both economic and safety risks. This work advocates for the use of Isolation Forest—a simple yet effective model—combined with MFCCs as an efficient approach for detecting anomalies in multivariate acoustic time series. The use of MFCCs for feature extraction was essential for reducing the dimensionality of the acoustic data and extracting relevant features, thus facilitating analysis and improving model performance. All methods considered in this study were integrated within the MTSA framework for their comparison for anomaly detection in multivariate acoustic time series. Such an integration allowed for a detailed performance analysis of each model, further highlighting the importance of MFCC-based feature extraction early in the data processing pipeline.

Additionally, the evaluation of methods considering in-distribution and out-of-distribution factors was a central aspect of this work, indicating the degree of robustness of each model when exposed to data diverging from the distribution of data considered during training. This analysis is essential for understanding how models generalize to real-world situations, where machine behavior can be unpredictable. This comparison revealed that, although methods such as Hitachi, GANF, and RANSynCoders rely on complex architectures, a simple model such as Isolation Forest can outperform them—suggesting that simpler models may offer superior generalization capabilities. Of course, this superiority comes with some hyperparameter tuning, which is not always feasible for the complex methods, due to the actual problem or limitations in computational resources.

Finally, by enabling full reproducibility of the experiments and comparative evaluation of different approaches, the MTSA framework offers a solid foundation for the development and enhancement of anomaly detection models. Future work may explore improved hyperparameter tuning, the use of synthetic data to increase test diversity, and the expansion of MTSA to integrate, evaluate, and further refine models and techniques for anomaly detection in industrial acoustic data.

Acknowledgements

Leonardo Afonso Ferreira Bortoni acknowledges the financial support from CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil). Diego Pinheiro acknowledges the financial support by the STIC-AmSud program (STIC AmSud Nro. 23-STIC-06) and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

References

[1] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.
[2] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE Access*, vol. 7, pp. 81542–81554, 2019.
[3] A. Vikram and Mohana, "Anomaly detection in network traffic using unsupervised machine learning approach," in *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, pp. 476–479, IEEE, 2020.
[4] R. de Paula Monteiro, M. C. Lozada, D. R. C. Mendieta, R. V. S. Loja, and C. J. A. B. Filho, "A hybrid prototype selection-based deep learning approach for anomaly detection in industrial machines," *Expert Systems with Applications*, vol. 204, p. 117528, 2022.
[5] Y. Chinniah, "Analysis and prevention of serious and fatal accidents related to moving parts of machinery," *Safety science*, vol. 75, pp. 163–173, 2015.
[6] S. AG, "Senseye predictive maintenance: The true cost of downtime 2022," tech. rep., Siemens AG, Erlangen, Germany, 2022.

TABLE III
AUC-ROC FOR IN-DISTRIBUTION (ID) AND OUT-OF-DISTRIBUTION (OOD) EVALUATIONS ACROSS DIFFERENT MACHINE TYPES.

Machine	Model	Scenario	AUC (95%CI)
Valve	Hitachi	ID	.56 (.45, .68)
		OOD	.61 (.60, .62)
	RANSynCoders	ID	.53 (.50, .55)
		OOD	.43 (.37, .49)
	GANF	ID	.58 (.48, .67)
		OOD	.44 (.38, .49)
	Isolation Forest	ID	**.93 (.92, .93)**
		OOD	**.68 (.66, .71)**
Pump	Hitachi	ID	.77 (.75, .79)
		OOD	.67 (.65, .68)
	RANSynCoders	ID	.79 (.75, .82)
		OOD	.67 (.64, .71)
	GANF	ID	.73 (.68, .77)
		OOD	.73 (.70, .77)
	Isolation Forest	ID	**.92 (.91, .92)**
		OOD	**.86 (.86, .87)**
Fan	Hitachi	ID	.60 (.59, .61)
		OOD	**.62 (.60, .64)**
	RANSynCoders	ID	**.64 (.59, .69)**
		OOD	.54 (.50, .56)
	GANF	ID	.54 (.52, .57)
		OOD	.51 (.46, .57)
	Isolation Forest	ID	.62 (.61, .62)
		OOD	.61 (.60, .61)
Slide Rail	Hitachi	ID	**.99 (.99, .99)**
		OOD	.84 (.82, .86)
	RANSynCoders	ID	.70 (.67, .73)
		OOD	.79 (.74, .84)
	GANF	ID	.94 (.92, .97)
		OOD	.92 (.83, .97)
	Isolation Forest	ID	**.99 (.99, .99)**
		OOD	**.96 (.95, .97)**

[7] H. Purohit, R. Tanabe, K. Ichige, T. Endo, Y. Nikaido, K. Suefusa, and Y. Kawaguchi, "Mimii dataset: Sound dataset for malfunctioning industrial machine investigation and inspection," *arXiv preprint arXiv:1909.09347*, 2019.

[8] H. Zhou, K. Yu, X. Zhang, G. Wu, and A. Yazidi, "Contrastive autoencoder for anomaly detection in multivariate time series," *Information Sciences*, vol. 610, pp. 266–280, 2022.

[9] S. R. Saufi, Z. A. B. Ahmad, M. S. Leong, and M. H. Lim, "Challenges and Opportunities of Deep Learning Models for Machinery Fault Detection and Diagnosis: A Review," *IEEE Access*, vol. 7, pp. 122644–122662, 2019.

[10] J. Pineau, P. Vincent-Lamarre, K. Sinha, V. Larivière, A. Beygelzimer, F. d'Alché Buc, E. Fox, and H. Larochelle, "Improving reproducibility in machine learning research (a report from the neurips 2019 reproducibility program)," 2020.

[11] D. PINHEIRO, "mtsa — pypi.org." https://pypi.org/project/mtsa/, 2024. [Accessed 29-05-2024].

[12] Z. K. Abdul and A. K. Al-Talabani, "Mel frequency cepstral coefficient and its applications: A review," *IEEE Access*, vol. 10, pp. 122136–122158, 2022.

[13] L. Bortoni and P. A. Jaskowiak, "Acoustic features and autoencoders for fault detection in rotating machines: A case study," in *34th Brazilian Conference on Intelligent Systems (BRACIS)*, Springer Nature Switzerland, 2025.

[14] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *2008 Eighth IEEE International Conference on Data Mining*, pp. 413–422, IEEE, 2008.

[15] E. Dai and J. Chen, "Graph-augmented normalizing flows for anomaly detection of multiple time series," in *2022 The International Conference on Learning Representations (ICLR)*, 2022.

[16] A. Abdulaal, Z. Liu, and T. Lancewicki, "Practical Approach to Asynchronous Multivariate Time Series Anomaly Detection and Localization," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, (Virtual Event Singapore), pp. 2485–2494, ACM, Aug. 2021.

[17] L. F. M. Filho, R. d. P. Monteiro, D. Pinheiro, P. T. Endo, and A. M. N. C. Ribeiro, "Forecasting imminent failures in electrical industrial centrifuge using machine learning," in *2023 IEEE Latin American Conference on Computational Intelligence (LA-CCI)*, pp. 1–6, 2023.

[18] K. M. Rezaul, M. Jewel, M. S. Islam, K. N. e. A. Siddiquee, N. Barua, M. A. Rahman, M. Shan-A-Khuda, R. B. Sulaiman, M. S. I. Shaikh, M. A. Hamim, *et al.*, "Enhancing audio classification through mfcc feature extraction and data augmentation with cnn and rnn models," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 15, no. 7, 2024.

[19] F. Rustam, A. Ishaq, M. S. A. Hashmi, H. U. R. Siddiqui, L. A. D. López, J. C. Galán, and I. Ashraf, "Railway track fault detection using selective mfcc features from acoustic data," *Sensors*, vol. 23, no. 16, p. 7018, 2023.

[20] M. S. Sidhu, N. A. A. Latib, and K. K. Sidhu, "Mfcc in audio signal processing for voice disorder: a review," *Multimedia Tools and Applications*, 2024.

[21] D. O'Shaughnessy, *Speech Communications: Human and Machine*. Wiley-IEEE Press, 2 ed., 1999.

[22] L. Rabiner and R. Schafer, *Theory and applications of digital speech processing*. Prentice Hall Press, 2010.

[23] Z. K. Abdul and A. K. Al-Talabani, "Mel frequency cepstral coefficient and its applications: A review," *IEEE Access*, vol. 10, 2022.

[24] C. Silva, M. Coelho, U. Barros, A. M. N. C. Ribeiro, R. de Paula Monteiro, and D. Pinheiro, "A unified framework for replicability of anomaly detection in multiple time series: Enhancing ganf and ransyncoders pipelines with mfcc feature extraction for acoustic data," in *2024 IEEE Latin American Conference on Computational Intelligence (LA-CCI)*, pp. 1–6, 2024.