RESEARCH PAPER

DOI: 10.22080/cste.2026.30790.1102

Contributions of Science and Technology for Engineering

Journal Home page: https://cste.journals.umz.ac.ir/

A VMD–mRMR–Random Forest Framework for Intelligent Detection and Severity Assessment of Broken Rotor Bar Faults in Induction Motors

Mohammad-Ebrahim Moazzen [1], Ali Akbar Abdoos [2]*, Seyyed Asghar Gholamian [2]

[1] *Ph.D student, Department of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran, ORCID:0000-0003-0843-3532.*

[2] *Associate Professor, Department of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran, ORCID:0000-0003-0973-9349, ORCID:0000-0001-7654-7668*

**Corresponding Author: a.abdoos@nit.ac.ir*

Article Info

Received: 21 December 2025

Accepted:10 February 2026

Available online: 25 August 2026

Keywords:

Induction Motors;

Broken Bar;

Pattern Recognition;

Signal Analysis;

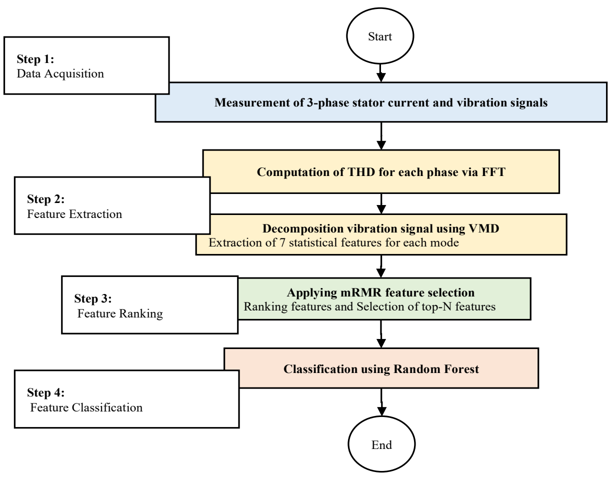
Feature Selection;

Vibration Signals.

Abstract:

Induction motors are critical and expensive components in industrial sectors; therefore, providing a reliable condition-monitoring scheme is a serious issue. However, mechanical maloperation due to broken rotor bar (BRB) faults can significantly degrade motor performance. If these severe faults go undetected, secondary damage is inevitable. This paper presents a new diagnosis method that uses time-frequency analysis to distinguish between healthy and BRB faults and to assess fault severity. To achieve this, vibration signals are first decomposed into intrinsic mode functions using Variational Mode Decomposition (VMD), and features are extracted from the decomposed signals. Then, the minimum Redundancy Maximum Relevance (mRMR) method selects the most effective features to improve the accuracy and generalizability of the detection scheme. Finally, these features are fed into a Random Forest (RF) classifier to detect the faulty condition and determine its severity. To implement and evaluate the proposed method, we use real-world vibration data collected from a 380 V, 4-pole induction motor. The results indicate that the proposed method is more accurate than conventional methods in detecting the number of broken bars, despite its simplicity. Therefore, the proposed intelligent method can be effectively utilized in industrial applications.

Graphical Abstract



ISSN 3060-6578

Please cite this paper as: M. E. Moazzen, A.-A. Abdoos, and S. A. Gholamian, "A VMD–mRMR–Random Forest Framework for Intelligent Detection and Severity Assessment of Broken Rotor Bar Faults in Induction Motors," Contrib. Sci. Technol. Eng., vol. 3, no. 3, pp. 1–10, 2026, doi:10.22080/cste.2026.30790.1102.

1. Introduction

Induction motors play an important role in transportation and industry as prime movers; therefore, ensuring their continuous operation is a serious issue. They account for a major share of global electricity consumption because of their favorable attributes. However, these machines are routinely exposed to thermal, mechanical, and electrical stresses, which may lead to maloperation. If degradation of key components such as stator windings, rotor bars, and bearings goes undetected, abrupt failures and severe economic losses are inevitable. Faults in induction motors are generally classified as electrical or mechanical. Broken rotor bars are among the most prevalent severe rotor-related failures and are typically triggered by repeated direct-on-line starting, sustained overloads, or unbalanced magnetic forces. These conditions induce localized heating, arcing, and spurious vibrations, leading to gradual performance deterioration [1-3].

Various diagnostic approaches have been proposed to discriminate broken rotor bars. For instance, stray flux-based methods using spectral analysis are employed [4]; however, their reliability is compromised under supply voltage unbalance. In contrast, [5] presents a technique based on air-gap magnetic field monitoring, but its reliance on intrusive sensors limits its practicality. Thermal monitoring has also been used, often alongside identification frameworks. Moreover, vibration analysis remains a well-established method for diagnosing mechanical issues [6]. Infrared thermography has been applied for fault detection [7]; however, the high computational burden is a major drawback for real-time implementation.

Moreover, broken rotor bar faults can manifest in vibration signatures [8]. Even incipient faults alter the machine's dynamic response, producing spurious vibrational characteristics. Rotor bar discontinuities directly influence stator currents; hence, Motor Current Signature Analysis (MCSA) is employed to identify the severity of such faults [9]. Fusing stator current and vibration data has been shown to improve diagnostic accuracy significantly. This is because both signals are non-intrusive and are simultaneously sensitive to electrical and mechanical maloperation. Moreover, their combined use aids in fault discrimination and is compatible with online condition monitoring [2]. Recently, researchers have proposed data-driven intelligent systems for fault detection [3]. These systems typically involve feature extraction and classification stages. Time-domain or frequency-domain techniques are used to transform raw signals. However, the non-stationary nature of fault transients has led to the adoption of advanced time–frequency representations. Methods such as wavelet transform [10-16] and modified empirical mode decomposition (EMD) [17] capture transient dynamics. The classification accuracy depends on the quality and relevance of the extracted features. To mitigate redundancy, feature selection techniques such as Principal Component Analysis (PCA) and Genetic Algorithms (GA) are employed [1-3].

In the classification phase, Artificial Neural Networks (ANNs) are widely used because of their nonlinear mapping capabilities [13, 16, 18]; however, their performance is sensitive to architectural choices. Moreover, Support Vector Machines (SVMs) construct optimal separating hyperplanes; however, their success depends on kernel selection [14, 15, 19, 20]. Neuro-fuzzy systems have also been presented to handle uncertainty in fault patterns, but their design demands expert knowledge [21]. Recently, deep learning models have been used to learn hierarchical fault features directly, thereby reducing reliance on manual engineering [22, 23]. In Schreier et al. [24], the Hidden Markov Model (HMM) was used to detect rotor-bar faults by analyzing stator and rotor currents.

Despite these advances, developing a robust diagnostic system is a serious issue. Existing studies indicate that the signal analysis method strongly influences diagnostic accuracy [1, 3]. Furthermore, the degree of class separability significantly affects the reliability of detection schemes.

This paper presents a new diagnostic framework for discriminating broken rotor bar faults. To do so, it decomposes vibration signals using Variational Mode Decomposition (VMD) [25]. Unlike traditional methods, VMD operates in both the time and frequency domains and is highly immune to noise. The method then extracts a set of statistical and spectral features to capture fault patterns. To improve model efficiency, the minimum Redundancy Maximum Relevance (mRMR) criterion identifies features relevant to the fault condition while minimizing redundancy [26]. Finally, a Random Forest (RF) classifier is utilized for fault diagnosis. It offers high classification accuracy and is well-suited for severity assessment in noisy conditions [27]. To implement and evaluate the proposed algorithm, we use real experimental data collected from a 4-pole, 60 Hz induction motor [28]. The obtained results demonstrate that the proposed method provides significant benefits, as outlined below:

- It is worth noting that VMD decomposes the signal effectively, isolating fault components without spurious mode-mixing.
- On the other hand, mRMR selects the most meaningful features; hence, model generalization is improved without sacrificing accuracy.
- Moreover, the suggested framework discriminates the fault severity levels accurately based on the number of broken bars.
- It should be noted that the computational burden of the proposed method is low, making it suitable for real-time monitoring.
- Finally, the method exhibits strong robustness and is quite superior for practical industrial applications, despite its simplicity.

2. Required Tools

2.1. Variational Mode Decomposition (VMD)

VMD is a time-frequency signal processing technique that decomposes a real-valued input signal $x(t)$ into K intrinsic mode functions (IMFs) $\{u_k(t)\}_{k=1}^{K}$, each assumed to be narrow-band around a center frequency ω_k. The decomposition is formulated as a constrained variational problem [25]:

$$\min_{\{u_k\},\{\omega_k\}} \left\{\sum_{k=1}^{K} \left\|\partial_t \left[\left(\delta(t)+\frac{j}{\pi t}\right) * u_k(t)\right] e^{-j\omega_k t}\right\|_2^2\right\} \text{ subject to } \sum_{k=1}^{K} u_k(t) = x(t) \quad (1)$$

Here, $\delta(t)$ denotes the Dirac delta function, and the term inside the norm represents the analytic signal of $u_k(t)$ using the Hilbert transform, demodulated to baseband by multiplication with $e^{-j\omega_k t}$. The L^2-norm of its derivative estimates the bandwidth of mode $u_k(t)$. Minimizing the sum of these bandwidths forces each mode to be spectrally compact. To solve this constrained problem, an augmented Lagrangian is introduced [25]:

$$\mathcal{L}(\{u_k\},\{\omega_k\},\lambda) = \alpha \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta + \frac{j}{\pi t} \right) * u_k \right] e^{-j\omega_k t} \right\|_2^2 + \left\| x(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 + \left\langle \lambda(t), x(t) - \sum_{k=1}^{K} u_k(t) \right\rangle \quad (2)$$

where α is the penalty parameter controlling the trade-off between bandwidth minimization and reconstruction accuracy, and is the Lagrange multiplier enforcing the equality constraint. Using the alternating direction method of multipliers (ADMM), the solution is iteratively updated in the frequency domain. Denoting the Fourier transform of a signal $f(t)$ as $\hat{f}(\omega)$, the update rules for the n-th iteration are [25]:

$$\begin{aligned} \hat{u}_k^{n+1}(\omega) &= \frac{\hat{x}(\omega) - \sum_{i\neq k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \\ \omega_k^{n+1} &= \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \\ \hat{\lambda}^{n+1}(\omega) &= \hat{\lambda}^n(\omega) + \tau[\hat{x}(\omega) - \sum_{k=1}^{K} \hat{u}_k^{n+1}(\omega)] \end{aligned} \quad (3)$$

where $\tau \in (0,2]$ is the dual-ascent step size and is often set to 1 for stability.

2.2. Minimum Redundancy Maximum Relevance (mRMR) Feature Selection

After decomposing the vibration signal using VMD and extracting a high-dimensional feature vector (e.g., from statistical moments, energy, entropy, or spectral peaks of each mode), the resulting feature space often contains redundant or weakly informative descriptors. Including all of them in a classifier not only increases computational cost but also degrades generalization due to the curse of dimensionality. To address this, we adopt the minimum Redundancy Maximum Relevance (mRMR) criterion. This filter-based feature selection method balances two competing objectives: maximizing relevance to the fault class while minimizing redundancy among selected features.

Let $\mathcal{F} = \{f_1, f_2, \dots, f_N\}$ denote the original set of N candidate features, and let C represent the fault class label (e.g., healthy, 1 broken bar, 2 broken bars). The relevance of a feature f_i to the class C is typically measured by the mutual information $I(f_i; C)$, while the redundancy between two features f_i and f_j is quantified by $I(f_i; f_j)$.

The mRMR algorithm incrementally constructs a subset $S \subset F$ of size $m \ll N$ by solving the following optimization at each selection step [26]:

$$\max_{f_i \in \mathcal{F}\setminus S} \left[I(f_i; C) - \frac{1}{|S|} \sum_{f_j \in S} I(f_i; f_j) \right] \quad (4)$$

This expression embodies the mRMR criterion: it favors features that are highly informative about the fault condition (large $I(f_i; C)$)) but minimally correlated with already-selected features (small average $I(f_i; f_j)$).

In practice, mutual information for continuous variables is estimated using kernel density estimation (KDE) or k-nearest neighbors (k-NN) methods, as proposed by Kraskov et al. For discrete or discretized features (e.g., after binning), it can be computed directly from joint histograms [26]:

$$I(X;Y) = \sum_{x\in X} \sum_{y\in Y} p(x,y) \log\left(\frac{p(x,y)}{p(x)p(y)} \right) \quad (5)$$

where $p(x,y)$, $p(x)$ and $p(y)$ are estimated from data samples.

In our BRB diagnostic framework, after extracting features from VMD modes, mRMR selects the most discriminative ones. Empirical observations show that features related to kurtosis and mode energy of mid-frequency IMFs often rank highest, consistent with the impulsive nature of BRB-induced vibrations.

The resulting compact feature vector not only accelerates RF training and inference but also improves robustness by eliminating noisy or correlated dimensions that could mislead the model, which is especially critical when the number of training samples is limited, as is common in industrial fault datasets.

Unlike wrapper methods (e.g., recursive feature elimination), mRMR is computationally efficient and classifier-independent, making it ideal for real-time monitoring systems where simplicity and speed are as important as accuracy.

2.3. Random Forest (RF) Classifier

Once the feature vector is obtained from VMD signal decomposition and mRMR feature selection, the next step is to discriminate discrete fault conditions such as healthy, one broken bar, or two broken bars. To do this, the Random Forest (RF) classifier is employed. RF is an ensemble learning method that combines multiple decision trees to achieve high accuracy. Moreover, it has inherent immunity against spurious overfitting; hence, it is quite superior for practical applications. RF constructs a collection of T decision trees $\{\tau_t\}_{t=1}^{T}$, each trained on a bootstrap sample (i.e., random sampling with replacement) of the original training dataset. During the construction of each tree, at every node split, only a random subset of $m \ll p$ features (where p is the total number of input features) is considered. This two-level randomization, in data sampling and feature selection, ensures low correlation among individual trees, which is key to the ensemble's generalization performance [27].

Given a test sample $x \in \mathbb{R}^p$, the final class prediction $\hat{y}$ is obtained by majority voting:

$$\hat{y} = \arg\max_{c \in \mathcal{C}} \sum_{t=1}^{T} \mathbb{I}(\mathcal{T}_t(\mathbf{x}) = c) \tag{6}$$

where c is the set of fault classes, and $\mathbb{I}(.)$ is the indicator function. In probabilistic form, the predicted class probability for class c is:

$$P(\hat{y} = c \mid \mathbf{x}) = \frac{1}{T}\sum_{t=1}^{T} \mathbb{I}(\mathcal{T}_t(\mathbf{x}) = c) \tag{7}$$

Beyond classification, RF provides feature importance scores without additional computation. The importance of feature f_j is measured by the average reduction in impurity (e.g., Gini index or entropy) across all trees whenever f_j is used for splitting:

$$Importance\left(f_j\right) = \frac{1}{T}\sum_{t=1}^{T} \sum_{n \in \mathcal{N}_t(f_j)} \frac{N_n}{N_{total}} . \Delta\text{Impurity}_n \tag{8}$$

where $\mathcal{N}_t\left(f_j\right)$ is the set of nodes in tree t that split on f_j, N_n is the number of samples reaching node n, and $\Delta\text{Impurity}_n$ is the impurity decrease at that node [27].

We found that RF consistently outperforms single decision trees in our BRB diagnostic setup, particularly when training data is limited. Its insensitivity to feature scaling, robustness to irrelevant features, and ability to model non-linear decision boundaries make it particularly well-suited for industrial fault diagnosis, where data is often noisy and class distributions may be slightly imbalanced. Moreover, the out-of-bag (OOB) error, computed using samples not included in a tree's bootstrap sample, provides an unbiased estimate of generalization performance without requiring a separate validation set, a practical advantage in data-scarce scenarios [29, 30].

3. Experimental Setup and Data Acquisition

The currents and vibration signals have been collected from the experimental tests were conducted at the Laboratory of Intelligent Automation of Processes and Systems and the Laboratory of Intelligent Control of Electrical Machines at the University of São Paulo, Brazil. The test bench features a 1150 VA, 4-pole, 60 Hz, three-phase squirrel-cage induction motor (220 V Δ / 380 V Y, 1715 rpm, rated torque: 4.1 N·m) with a 34-bar rotor. A control panel enables selection of the connection configuration (star or delta) and power supply mode (direct mains or three-phase inverter) [28]. A DC generator provides mechanical loading, with resistive torque adjusted manually via the field-winding voltage. This voltage is controlled using an 1800 W single-phase Variac and a rectifier to convert AC to DC. Shaft torque is measured using a Transtec MT-103 rotary torque transducer (up to 2000 rpm, 2 mV/N·m sensitivity, Wheatstone bridge-based), allowing realistic simulation of varying operational loads. Vibration data were captured using five Vibrocontrol PU 2001 uniaxial accelerometers. These sensors, with a sensitivity of 10 mV/mm/s and a frequency range of 5–2000 Hz, are mounted at key locations, including the drive and non-drive ends, along the axial housing, and on the support base. This placement enables the measurement of radial, axial, and tangential vibration components. Electrical signals are acquired using Yokogawa model 96033 current probes and Yokogawa oscilloscope voltage probes, which are connected directly to the motor terminals [28].

To emulate broken rotor bar (BRB) faults, the original healthy rotor (intact bars) was first tested, then replaced with modified rotors containing drilled bars to simulate single- and multiple-BRB conditions. All tests used direct-on-line starting under balanced 60 Hz three-phase supply. Current and voltage signals were sampled at 50 kHz, while vibration signals were recorded at 7.6 kHz [28].

A comprehensive dataset is constructed by applying eight torque levels ranging from 0.5 to 4.0 N·m in 0.5 N·m steps. Ten repeated runs were performed per condition to cover transient-to-steady-state operation. For each run, we recorded three-phase voltages and currents. Moreover, five vibration velocity signals, including tangential, axial, and radial components, were acquired. This process ensures a rich, multi-dimensional dataset for fault diagnosis research.

4. Proposed Diagnostic Framework

This section describes the proposed intelligent method for detecting RBR faults in induction motors. The methodology is structured into four consecutive stages, as described below.

Stage 1: Multi-Modal Signal Acquisition and Preprocessing

In the first step, the motor under test outputs three-phase stator current and vibration signals simultaneously. Then, the current signals are analyzed in the frequency domain using Fast Fourier Transform (FFT). It is worth noting that the Total Harmonic Distortion (THD) of each phase is computed from the resulting spectrum as a global indicator of waveform distortion, which is a known symptom of severe BRB faults. These three THD values are preserved as low-dimensional, high-level features to capture electrical maloperation.

Stage 2: Time–Frequency Decomposition and Feature Extraction

On the other hand, each vibration signal is decomposed into a set of band-limited intrinsic mode functions using VMD. The number of decomposition levels (K) is treated as a tunable hyperparameter and optimized during validation to balance resolution and redundancy. For each resulting mode, we extract a comprehensive set of statistical descriptors. These features include minimum and maximum amplitude, mean and standard deviation, Root Mean Square (RMS), variance, skewness, and kurtosis. With K modes and 8 features per mode, this stage yields 8*K features. With 5 vibration signals, the total number of extracted features, including the THD of the three-phase stator current, is 8*5*K+3. In Figure 1, eight decomposed modes of a tangential vibration signal are illustrated.

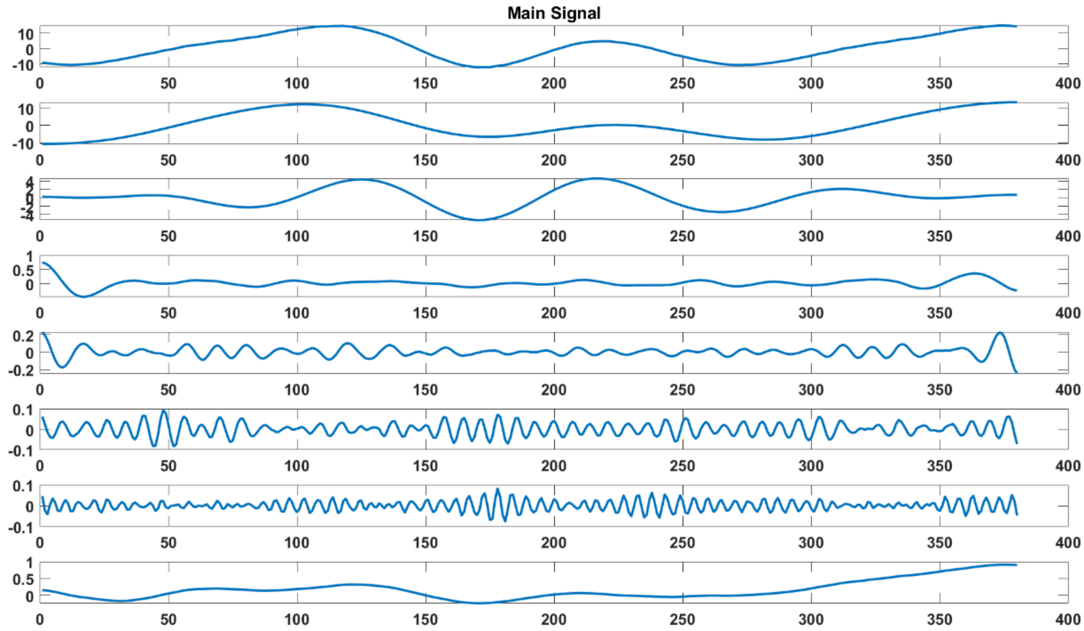


Figure 1. The tangential vibration signal as well as decomposed modes using VMD

Stage 3: Feature Selection via mRMR

To mitigate the curse of dimensionality and improve model generalization, we apply the mRMR criterion. This filter-based method ranks candidate features by their relevance to the fault class, measured by mutual information, and their redundancy with other selected features. Hence, only the top-ranked features are retained to strike an optimal trade-off between discriminative power and compactness. Finally, cross-validation determines the number of selected features to maximize classification performance.

Stage 4: Fault Classification with Random Forest

Finally, the refined feature vector is fed into the RF classifier. This classifier uses an ensemble of decorrelated decision trees trained on bootstrapped samples. The model outputs a probabilistic prediction for the five target classes: healthy, 1 broken bar, 2 broken bars, 3 broken bars, and 4 broken bars. To ensure robustness and avoid spurious overfitting, we optimize key hyperparameters such as the number of trees. Moreover, we fine-tune the decomposition levels (K) and the number of selected features through systematic experimentation to achieve the best trade-off between accuracy and computational efficiency. Figure 2 shows the implementation steps of the proposed detection method.

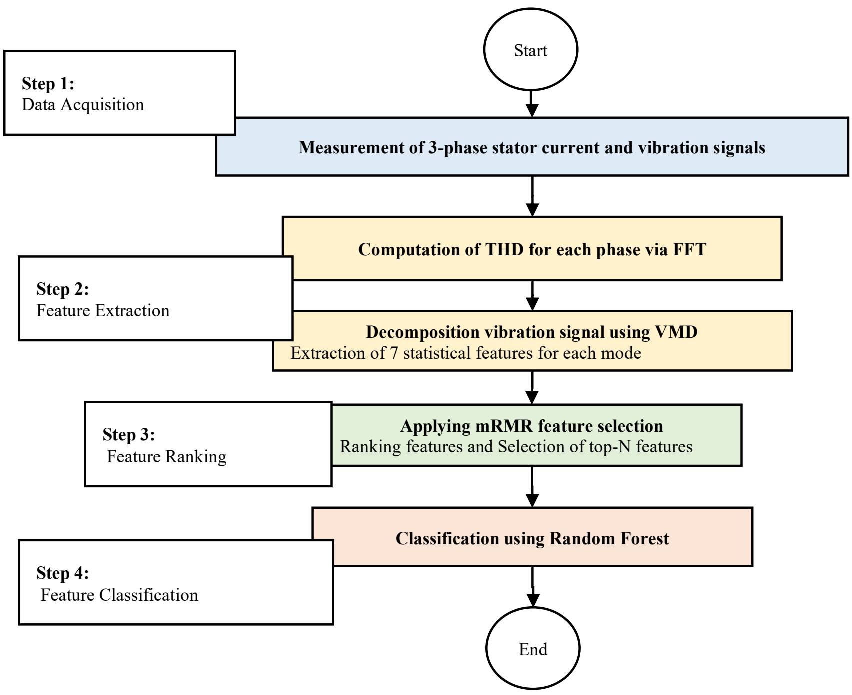


Figure 2. Flowchart of the proposed detection scheme

4.1. Finding Optimal Number of Intrinsic Mode Functions

To identify the optimal number of intrinsic mode functions (K) for VMD, a sensitivity analysis is performed. To do so, K is varied from 1 to 10 while keeping all other parameters fixed, including a random forest classifier with 100 decision trees in the detection core without applying feature selection. Figure 3 shows the resulting classification accuracy for each K value. As shown, the detection accuracy exhibits a clear non-monotonic trend. It rises sharply from 82% at K=1 to approximately 87.5% at K=4; however, it plateaus between K=5 and K=6, reaches its maximum of 88.0% at K=7, and then gradually declines to 85.0% at K=10. This behavior reflects a fundamental trade-off in signal decomposition. At low K values (e.g., K<4), the decomposition lacks sufficient resolution to isolate subtle fault-related transients; hence, underfitting occurs. In contrast, at moderate K (around 7), the modes effectively capture both the dominant frequency components and transient fault signatures without introducing excessive redundancy or spurious noise. It should be noted that at high K (e.g., K>8), over-decomposition occurs where some modes become spectrally overlapping. This degrades the classifier performance due to increased dimensionality and reduced generalization. Therefore, we select K=7 as the optimal decomposition level for subsequent experiments. This choice maximizes detection accuracy while ensuring computational efficiency.

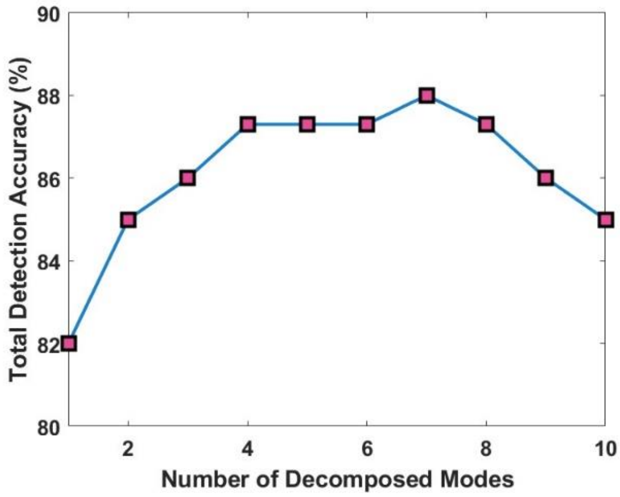


Figure 3. The detection accuracy versus different number of decomposed modes

4.2. Determining the Optimal Number of Features

To determine the optimal number of features for classification, we perform a comprehensive evaluation. To do this, the feature count is varied from 10 to 283 (5 vibration signals × 7 VMD modes × 8 features + 3 THD values = 283 initial features). In all cases, the VMD decomposition level is fixed at K=7, and the RF classifier is configured with 100 decision trees. As shown in Figure 4, detection accuracy follows a distinct pattern. It remains low for small feature sets (below 40); however, it rises sharply between 40 and 60 features, reaching a plateau around 91.3% at 60 selected features. Beyond this point, the accuracy stabilizes near 90–91% with minor fluctuations before slightly declining as the number of features exceeds 200.

This behavior highlights a critical insight:

- Too few features (e.g., < 40) fail to capture sufficient information for discrimination, resulting in underfitting.
- Around 60 features, the model achieves its best balance. This means enough relevant information is retained to discriminate fault severities, while the mRMR criterion effectively filters out redundant or spurious features.
- On the other hand, beyond 60 features, adding more dimensions does not improve performance. It may even degrade performance slightly due to increased overfitting risk or the inclusion of less informative descriptors.

Hence, the peak at 60 features is selected as the optimal dimensionality for the final diagnostic system. This choice maximizes classification accuracy while ensuring computational efficiency and robustness. This result confirms that feature selection is not merely a preprocessing step but a core component of the diagnostic procedure. Blindly using all available features can be counterproductive; however, intelligent pruning via mRMR enables the model to focus on the most physically meaningful indicators of severe BRB faults.

4.3. Determining the Optimum Number of Decision Trees

The performance of the RF classifier is highly sensitive to its hyperparameters; therefore, determining the number of decision trees (T) is a key challenge in balancing accuracy and computational burden. To do this, a grid search is conducted over the range of T=50 to T=500, while keeping other parameters fixed. It should be noted that VMD decomposition is set at K=7 modes and feature selection is limited to 60 determining features via mRMR.

As shown in Figure 5, the total detection accuracy increases steadily from approximately 85.5% at 50 trees to a peak of 96.7% at 250 trees; however, it plateaus briefly before gradually declining to around 92.5% at 500 trees. This trend is consistent with the theoretical behavior of ensemble methods. Increasing the number of trees initially reduces variance and improves stability. However, beyond a certain point, adding more trees yields diminishing returns and may introduce redundancy or spurious overfitting due to correlated base learners.

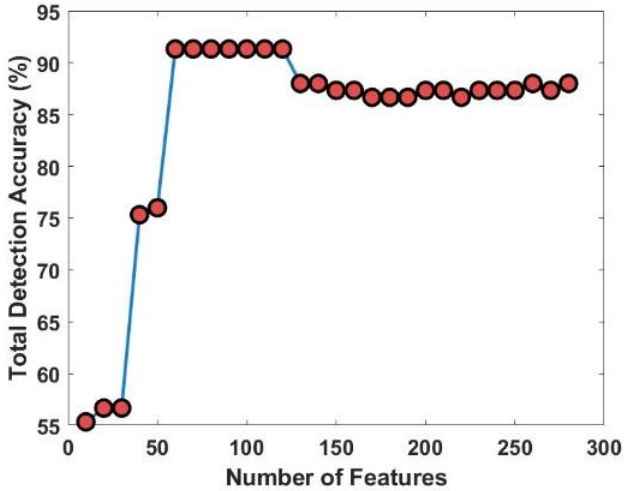


Figure 4. **The detection accuracy versus different numbers of selected features**

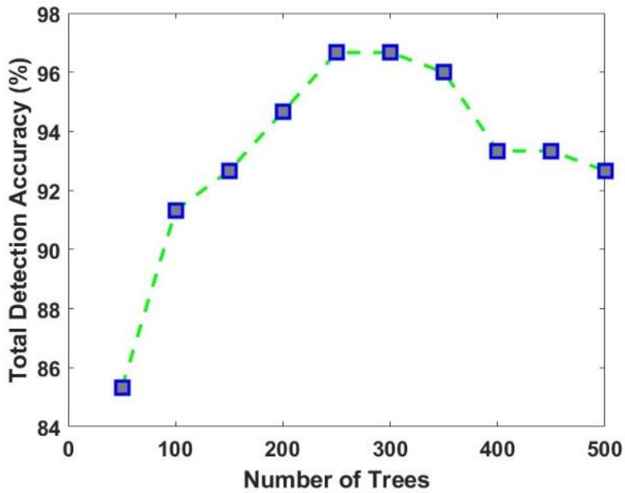


Figure 5. **The detection accuracy versus different number of trees**

Hence, the peak at 250 trees is selected as the optimal configuration for the final model. This choice offers the best trade-off. It ensures high accuracy (96.7%) in discriminating faults across all severity levels. Moreover, it has a reasonable computational load, which makes it suitable for real-time deployment. The system achieves stable performance, preventing the marginal degradation typically seen at higher tree counts. While RF is generally robust to the choice of T, this experiment confirms that blindly using default values can lead to suboptimal results. Empirical tuning is essential to maximize immunity and reliability in industrial fault diagnosis systems; therefore, the proposed method is desirable for practical applications.

5. Results

This paper presents a diagnostic framework that combines VMD, mRMR feature selection, and RF classification. It achieves 96.7% accuracy in discriminating five operational states: Healthy, 1 Broken Bar, 2 Broken Bars, 3 Broken Bars, and 4 Broken Bars. This performance is quite superior to many conventional methods and demonstrates the effectiveness of the integrated approach for multi-level fault severity assessment. As shown in the confusion matrix (Figure 6), the model performs near-perfectly for moderate to severe fault conditions. Notably, both precision and recall reach 100% for 2, 3, and 4 broken bars. This indicates that the method reliably identifies critical faults; hence, early intervention can prevent catastrophic failure and severe maloperation. However, a notable challenge remains in distinguishing the healthy state from mild faults. While the model correctly classifies all samples labeled "healthy," it misclassifies 13.3% of actual healthy samples as having one broken bar. Conversely, one sample from the 1 Broken Bar class is misclassified as 3 Broken Bars, though this error is minor. It should be noted that the vibration signatures of early-stage BRB faults are spectrally similar to those of normal operation. Despite this, the high precision for the Healthy class minimizes spurious alarms; moreover, the near-perfect detection of severe faults ensures reliability. The training and test times are about 1 and 0.03 s, respectively, using a computer with an Intel Core i7 @ 2.5 GHz CPU and 12 GB of RAM. In summary, the proposed method is a robust solution for intelligent condition monitoring, which is suitable for industrial environments, despite its simplicity.

Figure 6. Confusion matrix for optimized intelligent structure

5.1. Comparison

To demonstrate the proposed algorithm's competence, we compare it against several state-of-the-art intelligent methods for broken rotor bar (BRB) fault detection, as shown in Table 1. In Zgarni et al. [20], wavelet packet transform (WPT) combined with support vector data description (SVDD) was used to detect BRB faults using stator current signals. The study reported that SVDD performs better than conventional SVM. In Bessam et al. [18], the Hilbert transform (HT) is utilized to extract harmonic components from current signals, followed by a neural network (NN) to estimate the number of broken bars. In contrast, Keskes et al. [15] integrate a Daubechies wavelet-based discrete wavelet transform (WT) with SVM. Notably, the one-versus-one SVM strategy outperformed the one-versus-all variant. In Keskes and Braham [14], a pitch-synchronous WT (PSWT) coupled with a directed acyclic graph (DAG)-SVM architecture achieved 99% accuracy; however, it did not identify the exact number of faulty bars.

In Sadeghian et al. [16], WT is applied to current signals, and an artificial neural network (ANN) is used for classification. Although it performs robustly even under low-load conditions, it is limited to cases involving one or two broken bars. In Schreier et al. [24], frequency spectrum analysis of current signals is combined with a hidden Markov model (HMM) to discriminate BRB faults from stator winding faults. The accuracy is 98%, which is desirable; however, it does not quantify severity (i.e., the number of broken bars). In Defdaf et al. [13], a WT–ANN framework using vibration signals is proposed where RMS and kurtosis from decomposed levels served as dominant features. The accuracy is reported as 98.66%; however, it does not identify the number of BRBs.

Most existing approaches rely exclusively on stator current signals, with only a few exploring vibration-based diagnostics. In contrast, this paper investigates both stator current and multiple vibration signal channels. The proposed method appears less accurate than others, but they don't specify the number of broken bars. Our method's advantage is that the number of broken bars can determine the fault's severity. Our analysis shows that vibrational signals contain richer discriminative information for BRB fault characterization.

Table 1. Performance comparison with existing intelligent BRB detection methods

Reference	Data	Signal Processing Technique	Classifier	Accuracy (%)	Determining Number of BRB
Zgarni et al. [20]	Current signals	WPT	SVDD	100	Yes
Bessam et al. [18]	Current signals	HT	NN	-	Yes
Keskes et al. [15]	Current signals	WT	SVM	99	No
Keskes and Braham [14]	Current signals	PSWT	DAG-SVM	99	No
Sadeghian et al. [16]	Current signals	WT	ANN	96	Yes
Schreier et al. [24]	Current signals	HMM	-	-	No
Defdaf et al. [13]	Vibration signals	WT	ANN	98.66%.	No
Proposed Method	Vibration signals	VMD	RF	96.7	Yes

6. Conclusions

This paper presents a new data-driven diagnostic framework to discriminate and assess the severity of BRB faults in induction motors. To do so, it leverages vibration-based analysis to address critical challenges in industrial condition monitoring. VMD effectively decomposes vibration signals, overcoming limitations of traditional decomposition techniques such as mode mixing. In contrast, mRMR reduces dimensionality while preserving the most informative attributes. Finally, the RF classifier demonstrates exceptional robustness in handling class imbalance. To evaluate the proposed method, experimental validation on a real 380 V induction motor is performed. The approach discriminates between healthy and faulty conditions and accurately identifies the exact number of broken rotor bars. Obtained results indicate that the proposed framework is quite superior to existing methods for deployment in harsh industrial environments, despite its simplicity. Future work will extend the framework to multi-fault scenarios.

7. Statements & Declarations

7.1. Author Contributions

Mohammad Ebrahim Moazzen: Conceptualization, Formal analysis, Investigation, Methodology, Software, Writing - original draft. **Ali Akbar Abdoos:** Supervision, Data curation, Validation, Writing - review & editing. **Sayyed Asghar Gholamian:** Supervision, Data curation, Validation, Writing - review & editing.

7.2. Data Availability Statement

Data available in a publicly accessible repository: The data presented in this study are openly available in https://ieee-dataport.org/open-access/experimental-database-detecting-and-diagnosing-rotor-broken-bar-three-phase-induction at doi.10.21227/4f7v-8c57, reference number [28].

7.3. Ethics Approval

Not applicable.

7.4. Informed Consent

Not applicable.

7.5. Consent for Publication

Not applicable.

7.6. Conflicts of Interest

The author(s) declare no conflict of interest.

8. References

[1] S. Halder, S. Bhat, D. Zychma, and P. Sowa, "Broken Rotor Bar Fault Diagnosis Techniques Based on Motor Current Signature Analysis for Induction Motor—A Review," Energies, vol. 15, no. 22, p. 8569, 2022, doi:10.3390/en15228569.

[2] M. E. E. D. Atta, D. K. Ibrahim, and M. I. Gilany, "Broken Bar Fault Detection and Diagnosis Techniques for Induction Motors and Drives: State of the Art," IEEE Access, vol. 10, pp. 88504–88526, 2022, doi:10.1109/ACCESS.2022.3200058.

[3] O. E. Hassan, M. Amer, A. K. Abdelsalam, and B. W. Williams, "Induction Motor Broken Rotor bar Fault Detection Techniques based on Fault Signature Analysis – a review," IET Electric Power Applications, vol. 12, no. 7, pp. 895–907, May 2018, doi:10.1049/iet-epa.2018.0054.

[4] P. A. Panagiotou, I. Arvanitakis, N. Lophitis, J. A. Antonino-Daviu, and K. N. Gyftakis, "A New Approach for Broken Rotor Bar Detection in Induction Motors Using Frequency Extraction in Stray Flux Signals," IEEE Trans. Ind. Appl., vol. 55, no. 4, pp. 3501–3511, 2019, doi:10.1109/TIA.2019.2905803.

[5] Y. Soleimani, S. M. A. Cruz, and F. Haghjoo, "Broken Rotor Bar Detection in Induction Motors Based on Air-Gap Rotational Magnetic Field Measurement," IEEE Trans. Instrum. Meas., vol. 68, no. 8, pp. 2916–2925, 2019, doi:10.1109/TIM.2018.2870265.

[6] C. Morales-Perez, J. Rangel-Magdaleno, H. Peregrina-Barreto, J. P. Amezquita-Sanchez, and M. Valtierra-Rodriguez, "Incipient Broken Rotor Bar Detection in Induction Motors Using Vibration Signals and the Orthogonal Matching Pursuit Algorithm," IEEE Trans. Instrum. Meas., vol. 67, no. 9, pp. 2058–2068, 2018, doi:10.1109/TIM.2018.2813820.

[7] D. Lopez-Perez and J. Antonino-Daviu, "Application of Infrared Thermography to Failure Detection in Industrial Induction Motors: Case Stories," IEEE Trans. Ind. Appl., vol. 53, no. 3, pp. 1901–1908, 2017, doi:10.1109/TIA.2017.2655008.

[8] A. Prudhom, J. Antonino-Daviu, H. Razik, and V. Climente-Alarcon, "Time-Frequency Vibration Analysis for the Detection of Motor Damages Caused by Bearing Currents," Mechanical Systems and Signal Processing, vol. 84, pp. 747–762, Feb. 2017, doi:10.1016/j.ymssp.2015.12.008.

[9] Naha, A. K. Samanta, A. Routray, and A. K. Deb, "A Method for Detecting Half-Broken Rotor Bar in Lightly Loaded Induction Motors Using Current," IEEE Trans. Instrum. Meas., vol. 65, no. 7, pp. 1614–1625, 2016, doi: 10.1109/TIM.2016.2540941.

[10] M. A. Mohamed, A. A. A. Mohamed, M. Abdel-Nasser, E. E. M. Mohamed, and M. A. M. Hassan, “Induction Motor Broken Rotor Bar Faults Diagnosis Using ANFIS-based DWT,” Int. J. Model. Simul., vol. 41, no. 3, pp. 220–233, 2021, doi: 10.1080/02286203.2019.1708173.

[11] J. Rangel-Magdaleno, H. Peregrina-Barreto, J. Ramirez-Cortes, R. Morales-Caporal, and I. Cruz-Vega, “Vibration Analysis of Partially Damaged Rotor Bar in Induction Motor under Different Load Condition Using DWT,” Shock Vib., vol. 2016, p. 3530464, 2016, doi: 10.1155/2016/3530464.

[12] G. R. Agah, A. Rahideh, H. Khodadadzadeh, S. M. Khoshnazar, and S. Hedayatikia, “Broken Rotor Bar and Rotor Eccentricity Fault Detection in Induction Motors Using a Combination of Discrete Wavelet Transform and Teager-Kaiser Energy Operator,” IEEE Trans. Energy Convers., vol. 37, no. 3, pp. 2199–2206, 2022, doi: 10.1109/TEC.2022.3162394.

[13] M. Defdaf, F. Berrabah, A. Chebabhi, and B. D. E. Cherif, “A New Transform Discrete Wavelet Technique based on Artificial Neural Network for Induction Motor Broken Rotor Bar Faults Diagnosis,” Int. Trans. Electr. Energy Syst., vol. 31, no. 4, p. 12807, 2021, doi: 10.1002/2050-7038.12807.

[14] H. Keskes and A. Braham, “DAG SVM and Pitch Synchronous Wavelet Transform for Induction Motor diagnosis,” 7th IET International Conference on Power Electronics, Machines and Drives (PEMD 2014), pp. 0166–0166, 2014, doi:10.1049/cp.2014.0321.

[15] H. Keskes, A. Braham, and Z. Lachiri, “Broken Rotor Bar Diagnosis in Induction Machines through Stationary Wavelet Packet Transform and Multiclass Wavelet SVM,” Electric Power Systems Research, vol. 97, pp. 151–157, Apr. 2013, doi:10.1016/j.epsr.2012.12.013.

[16] Sadeghian, Z. Ye, and B. Wu, “Online Detection of Broken Rotor Bars in Induction Motors by Wavelet Packet Decomposition and Artificial Neural Networks,” IEEE Trans. Instrum. Meas., vol. 58, no. 7, pp. 2253–2263, 2009, doi: 10.1109/TIM.2009.2013743.

[17] M. S. Arifin, W. Wang, and M. N. Uddin, “A Modified EMD Technique for Broken Rotor Bar Fault Detection in Induction Machines,” Sensors, vol. 24, no. 16, p. 5186, 2024, doi: 10.3390/s24165186.

[18] B. Bessam, A. Menacer, M. Boumehraz, and H. Cherif, “Detection of Broken Rotor Bar Faults in Induction Motor at Low Load Using Neural Network,” ISA Trans., vol. 64, pp. 241–246, 2016, doi: 10.1016/j.isatra.2016.06.004.

[19] Hs. Halder, S. Bhat, And B. Dora, “Prediction of Broken Rotor Bar in Induction Motor Using Spectral Entropy Features and TLBO Optimized SVM,” Turkish Journal of Electrical Engineering and Computer Sciences, vol. 30, no. 5, pp. 1962–1979, Jul. 2022, doi:10.55730/1300-0632.3916.

[20] S. Zgarni, H. Keskes, and A. Braham, “Nested SVDD in DAG SVM for Induction Motor Condition Monitoring,” Eng. Appl. Artif. Intell., vol. 71, pp. 210–215, 2018, doi: 10.1016/j.engappai.2018.02.019.

[21] M. S. Eddine, K. Djallel, G. Abdelhalim, and D. Selman, “Fuzzy Diagnostics of Rotor Bar Breakage in Induction Motor Based on Time Synchronous Averaging and Wavelet Transform,” Arab. J. Sci. Eng., vol. 50, no. 14, pp. 10779–10794, 2025, doi: 10.1007/s13369-024-09850-5.

[22] Y. Shu, W. Zhang, X. Song, G. Liu, and Q. Jiang, “DBF-CNN: A Double-Branch Fusion Residual CNN for Diagnosis of Induction Motor Broken Rotor Bar,” IEEE Trans. Instrum. Meas., vol. 72, pp. 1–10, 2023, doi: 10.1109/TIM.2023.3325863.

[23] P. Kumar and A. S. Hati, “Dilated Convolutional Neural Network based Model for Bearing Faults and Broken Rotor bar Detection in Squirrel Cage Induction Motors,” Expert Syst. Appl., vol. 191, 2022, doi: 10.1016/j.eswa.2021.116290.

[24] L. Schreier, J. Bendl, and M. Chomat, “Analysis of Stator and Rotor Currents and Torque of Induction Machine with Rotor-Bar Faults,” Electr. Eng., vol. 103, no. 1, pp. 519–528, 2021, doi: 10.1007/s00202-020-01096-2.

[25] K. Dragomiretskiy and D. Zosso, “Variational Mode Decomposition,” IEEE Trans. Signal Process., vol. 62, no. 3, pp. 531–544, 2014, doi: 10.1109/TSP.2013.2288675.

[26] X. Wang, Y. Tao, and K. Zheng, “Feature Selection Methods in the Framework of mRMR,” 2018 Eighth International Conference on Instrumentation & Measurement, Computer, Communication and Control (IMCCC), pp. 1490–1495, Jul. 2018, doi:10.1109/imccc.2018.00307.

[27] H. A. Salman, A. Kalakech, and A. Steiti, “Random Forest Algorithm Overview,” Babylonian Journal of Machine Learning, vol. 2024, pp. 69–79, Jun. 2024, doi:10.58496/bjml/2024/007.

[28] E. Treml, R. A. Flauzino, M. Suetake, and N. A. R. Maciejewski, “Experimental Database for Detecting and Diagnosing Rotor Broken Bar in a Three-Phase Induction Motor,” IEEE Dataport, p. 1, 2020, doi: 10.21227/fmnm-bn95.

[29] H. Yousefpour and J. Ghasemi, “Ensemble-Based Detection and Classification of Liver Diseases Caused by Hepatitis C,” Contrib. Sci. Technol. Eng., vol. 1, no. 1, pp. 32–42, 2024, doi: 10.5937/cste2401033Y.

[30] M. Mehregan, A. Rezasoltani, and A. Khani, “A Novel Hybrid Machine Learning Model for Defect Prediction in Industrial Manufacturing Processes,” Contrib. Sci. & Tech Eng, vol. 2, no. 4. p. 16, 2025, doi: 10.22080/cste.2025.29099.1037.