# Severity Classification of Progressive Rotor Electrical Unbalance in Wind Turbine DFIGs Using Deep Temporal Learning


Mohamed Salah Channouf
University of Tunis, National Higher Engineering School of Tunis (ENSIT)
LISIER Lab
Tunis 1008, Tunisia
mohamedsalah.channouf@ensit.rnu.tn

Lotfi Saidi
National Higher Engineering School of Tunis (ENSIT),
University of Tunis, SIME Lab, Tunis 1008, Tunisia
lotfi.saidi@esstt.rnu.tn

khmais BACHA
University of Tunis, National Higher Engineering School of Tunis (ENSIT)
LISIER Lab
Tunis 1008, Tunisia
khemaies.bacha@isste.u-carthage.tn



***Abstract*—Rotor electrical unbalance (REU) in doubly-fed induction generators (DFIG) used in wind energy conversion systems is a progressive fault whose early detection remains challenging under variable-speed operating conditions. This study presents a comprehensive hybrid framework integrating classical signal processing with advanced deep learning for robust, multi-severity REU detection. The framework operates in three stages. In the first stage, spectral analysis via the Fast Fourier Transform (FFT) extracts fault-indicative sideband amplitudes at $(1\pm2ks)f_s$; Hilbert-transform envelope analysis captures amplitude modulation patterns; and four statistical indicators (kurtosis, crest factor, skewness, RMS) quantify signal impulsiveness. In the second stage, the Robust Local Mean Decomposition (RLMD) decomposes the stator current into Product Functions (PFs), from which Fault Energy Ratios (FER) are computed as dimensionless, load-invariant descriptors. These stages form a compact eleven-dimensional feature vector. The third stage introduces temporal modelling via LSTM networks and a hybrid CNN-LSTM. Experimental validation on a 30 kW laboratory DFIG test bench with four REU severity levels (Healthy, 150%, 225%, 300%) under 5-fold cross-validation yields 98.0% accuracy and 97.7% macro-F1 for the CNN-LSTM. Inference time is 4.2 ms/sample on CPU.**



***Keywords—Doubly-fed induction generator, rotor electrical unbalance, FFT, Hilbert envelope, kurtosis, robust local mean decomposition, fault energy ratio, LSTM, CNN-LSTM, wind turbine, condition monitoring, predictive maintenance.***


## I. INTRODUCTION

Wind energy has emerged as one of the most rapidly expanding renewable energy sources worldwide. Global installed capacity surpassed 900 GW in 2023 and is projected to reach 2.0 TW by 2030 [2]. The doubly-fed induction generator (DFIG) is the dominant topology for variable-speed wind turbines in the 1 MW to 10 MW range, accounting for over 60% of newly installed capacity, owing to decoupled active/reactive power control and a partially rated back-to-back converter [3], [4].

Despite these advantages, DFIGs are subject to electrical and mechanical faults. Electrical faults account for 23% of all failure events in a survey of 35 000 onshore turbines over 13 years [5]. REU—caused by asymmetric impedance in the rotor windings due to insulation degradation, inter-turn short circuits, or manufacturing tolerances—is particularly insidious because it develops gradually and its spectral signature overlaps with converter-generated harmonics [7], [8].

Early REU detection is critical: a missed mild fault escalating to severe unbalance requires full rotor rewinding, with repair costs of EUR 30 000–EUR 80 000 per turbine and a 2–4 week outage [9]. Under variable speed, fault sidebands at $(1\pm2ks)f_s$ shift in frequency and are masked by converter harmonics. Classical FFT-based Motor Current Signature Analysis (MCSA) degrades significantly at low slip ($s < 0.02$) [8].

Three fundamental limitations characterize the state of the art: Feature poverty: Most MCSA systems use only one or two spectral amplitudes, discarding envelope, statistical, and energy-decomposition information [7]. Temporal blindness: Classical ML classifiers treat each signal window independently, ignoring fault evolution [31]. Decomposition fragility: Empirical Mode Decomposition (EMD)-based approaches suffer from mode mixing in industrial noise environments [19], [21].

This study makes three principal contributions to the field of DFIG condition monitoring. First, it introduces an 11-dimensional feature vector that, for the first time, jointly combines FFT sideband amplitudes, Hilbert envelope analysis, four time-domain statistical indicators (RMS, kurtosis, crest factor, and skewness), and three RLMD-based Fault Energy Ratios for REU characterisation in DFIGs — a multi-source fusion that no prior work has assembled for this fault type. Second, a three-tier temporal modelling pipeline is proposed, progressing from lag-augmented classical machine learning (Random Forest (RF), Support Vector Machines (SVM), and Multilayer Perceptron. (MLP)) through a standalone LSTM sequential model, to a fully end-to-end CNN-LSTM architecture that requires no manual feature engineering and captures both local spectral patterns and long-range temporal fault evolution simultaneously. Third, the framework is validated with rigorous experimental discipline — including 5-fold stratified cross-validation with strict data-leakage prevention at the recording level, Synthetic Minority Over-sampling Technique (SMOTE) -based class balancing applied exclusively within training folds, and a systematic ablation study quantifying the individual contribution of each feature group — conducted on a 30 kW laboratory DFIG test bench across four REU severity levels. This work extends the authors' preliminary conference study [20] and prior RLMD-based diagnostic framework [1] into a comprehensive, temporally-aware, and experimentally rigorous end-to-end system.

Thomson and Fenger [11] established the theoretical basis for MCSA. A comprehensive review of induction motor signature analysis is given in [18]. FFT-based MCSA is effective under constant-load, fixed-speed operation but degrades at variable speed where fault frequencies smear across bins [12]. The STFT partially addresses non-stationarity via a sliding window but introduces a time-frequency resolution trade-off [13]. Envelope analysis via the Hilbert transform detects amplitude modulations [14]; kurtosis of the envelope indicates impulsive faults [15]. EMD [19] decomposes signals adaptively but suffers from mode mixing; RLMD [21] overcomes this through stable local mean estimation. Multiple fault detection in grid-connected DFIG wind turbines was recently studied in [6].

SVM with RBF kernels achieve 85–93% accuracy on small laboratory datasets [22], [23]; a comprehensive review of SVM for fault diagnosis is given in [24]. RFs provide ensemble robustness and feature importance [20]. Both methods ignore temporal dynamics. SMOTE [26] is standard for balancing training distributions.

1D-CNNs applied to raw signals learn features automatically and exceed 98% accuracy on public bearing datasets [28], [29], [30]. LSTM [27] consistently outperforms static classifiers on fault data with gradual temporal evolution [31]. Transformer-based models achieve competitive accuracy but their quadratic complexity makes them impractical for nacelle edge hardware [32], [33]. Hybrid CNN-LSTM architectures combine local spectral extraction with sequential temporal modelling [34].

The remainder of this paper is organized as follows. Section II reviews the relevant literature and identifies the research gaps. Section III provides the theoretical background of the signal processing methods employed. Section IV describes the DFIG model and REU fault characterization. Section V details the multi-source feature extraction methodology. Section VI describes the experimental setup and dataset. Section VII reports and discusses the results. Finally, Section VIII draws conclusions and outlines future research directions.

## II. THEORETICAL BACKGROUND

### *A. Fast Fourier Transform*

The discrete Fourier transform (DFT) of a signal $x[n]$ over $N$ samples is given by:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N}, k = 0, \dots, N-1 \quad (1)$$

The FFT amplitude at $f_r$ is:

$$A = \frac{2}{N}|X[k]| \quad (2)$$

$k$ corresponds to $f_r$.

A Hanning window

$$w[n] = 0.5\left(1 - \cos(\frac{2\pi n}{N-1}\right) \quad (3)$$

reduces spectral leakage. For non-stationary signals the STFT applies a sliding window:

$$X(m,k) = \sum_{-\infty}^{+\infty} x[n]\, w[n - mH] e^{-j2\pi kn/N} \quad (4)$$

### *B. Hilbert Envelope Analysis*

The analytic signal

$$\hat{x}(t) = x(t) + j\mathcal{H}\{x(t)\} \quad (6)$$

and the instantaneous envelope $e(t)$ is given by:

$$e(t) = |\hat{x}(t)| = \sqrt{x^2(t) + \mathcal{H}^2\{x(t)\}} \quad (7)$$

The envelope highlights amplitude modulations caused by mechanical faults, particularly in scenarios where transient events introduce sidebands around fr [14]. Randall and Antoni [17] provide a comprehensive tutorial on envelope-based diagnostics.

### *C. Statistical Indicators*

Four time-domain metrics quantify signal non-Gaussianity [15], [16] are given in Table I.

### *D. Summary of Research Gaps and Main Contributions*

Although various fault diagnostic methods exist, they often suffer from feature poverty in traditional MCSA, temporal blindness in static classifiers, or mode-mixing issues in standard EMD. To overcome these limitations, the main contributions of this work are:

- Multi-Source Feature Fusion: Proposing an 11-dimensional feature vector combining FFT, Hilbert Envelope, time-domain statistics, and an RLMD-based Fault Energy Ratio (FER).
- Deep Temporal Architecture: Developing a hybrid CNN-LSTM framework capable of extracting local spatial features and capturing long-range temporal fault progression.
- Robust Severity Tracking: Achieving highly accurate 4-class severity classification (Healthy, 150%, 225%, 300%) under variable wind speed and switching converter noise conditions.

## III. DFIG MODEL AND FAULT CHARACTERISATION

To effectively quantify the non-Gaussianity and impulsiveness of the stator current signals under progressive fault conditions, four time-domain statistical metrics are utilized. The mathematical definitions of these indicators :namely Root Mean Square (RMS), Kurtosis (K), Crest Factor (CF), and Skewness (Sk) are detailed in Table I.

TABLE I : STATISTICAL HIS EXTRACTED FROM VIBRATION SIGNALS

| Statistical indicator | Mathematical definition |
|---|---|
| Root Mean Square (RMS) | $x_{RMS} = \sqrt{\frac{1}{N}\sum_{n=0}^{N-1} x[n]^2}$ |
| Kurtosis (K) | $K = \frac{\mathbb{E}[(x-\mu)^4]}{\sigma^4}$ |

| | |
|---|---|
| Crest Factor (*CF*) | $CF = \frac{max\lvert x(t)\rvert}{x_{RMS}}$ |
| Skewness (Sk) | $Sk = \frac{E[(x-\mu)^3]}{\sigma^3}$ |

### A. REU Physical Mechanism

REU arises from asymmetric impedance ΔZ in rotor windings. By symmetrical components, the antisymmetrical fault component generates a backward MMF at slip frequency $sf_1$, producing stator current components at:

$$f^{\pm}(k) = (1 \pm 2ks)f_1 \,, k = 1,2,3, \ldots \quad (5)$$

Sideband amplitude is proportional to fault severity [7]:

$$A\pm \propto \left(\frac{\Delta R}{R+\Delta R}\right). I_r \quad (6)$$

### B. Severity Classification

Four classes are defined:

- Class 0 (Healthy): $\frac{\Delta R}{R_r}$ = 0%.
- Class 1 (150% REU): $\frac{\Delta R}{R_r}$ = 50%. Sideband +3–8 dB.
- Class 2 (225% REU): $\frac{\Delta R}{R_r}$ = 125%. Sideband +8–14 dB.
- Class 3 (300% REU): $\frac{\Delta R}{R_r}$ = 200%. Sideband +14–20 dB.

Separating Class 1 from Healthy is the primary diagnostic challenge, since at $s < 0.04$ the sideband amplitude difference can be as small as 3 dB above the noise floor.

### C. Converter Harmonic Interference

The back-to-back converter generates PWM switching harmonics at $\hat{f}(m,n) = mf_sw \pm nf_1$ (fsw: 2–5 kHz) [8]. These can mask low-amplitude REU sidebands at Class 1.

The multi-source feature approach mitigates this: RLMD-based FER features are extracted in complementary frequency bands, less sensitive to PWM interference.

## IV. Feature Extraction Methodology

### A. Pipeline Overview

Four parallel paths operate on each $N$ = 1024-sample window: (*i*) FFT spectral analysis; (*ii*) Hilbert envelope demodulation; (*iii*) time-domain statistics; (*iv*) RLMD decomposition. Their outputs form the 11-dimensional feature vector $x \in \mathbb{R}^{11}$, z-score normalised on the training set only.

### B. FFT Sideband Features

Lower and upper sideband amplitudes at:

$$A\pm = \frac{2}{N}\lvert X[k^{\pm}]\rvert \,,\; k^{\pm} = \left[\frac{(1\pm 2s)f_1}{\Delta f} + 0.5\right] \quad (7)$$

### C. Hilbert Envelope Feature

The mean-subtracted envelope peak at rotor frequency $f_r$:

$$A_{env} = \frac{2}{N}\lvert E'[k_r]\rvert \,,\; k_r = \left[\frac{f_r}{\Delta f} + 0.5\right] \quad (8)$$

### D. RLMD and Fault Energy Ratio

RLMD iteratively extracts PFs using adaptive-window local mean estimation, avoiding EMD mode mixing [21]. The FER:

$$FER_k = \frac{EE_k}{\sum_{i=1}^{m} E_i} \,,\; E_k = \sum_{n=0}^{n-1} \lvert PF_k[n]\rvert^2 \quad (9)$$

FER is load-invariant: amplitude changes cancel in the normalisation, a key advantage under variable wind loading [21].

### E. Complete Feature Vector

The complete feature vector (Table II) is given by: $x = [A^-, A^+, A_{env}, x_{RMS}, K, CF, Sk, P2P, FER_1, FER_2, FER_3]^T$.

TABLE II: Feature Vector Definition

| # | Symbol | Method | Physical Meaning |
|---|---|---|---|
| 1 | $A^-$ | FFT | Lower sideband at $(1-2s)f_1$ |
| 2 | $A^+$ | FFT | Upper sideband at $(1+2s)f_1$ |
| 3 | *Aenv* | Hilbert | Envelope peak at $f_r$ |
| 4 | *xRMS* | Stat. | Signal energy level |
| 5 | *K* | Stat. | Impulsiveness (kurtosis) |
| 6 | *CF* | Stat. | Peak-to-RMS ratio |
| 7 | *Sk* | Stat. | Distribution asymmetry |
| 8 | *P2P* | Stat. | Peak-to-peak amplitude |
| 9 | $FER_1$ | RLMD | Dominant PF energy fraction |
| 10 | $FER_2$ | RLMD | Secondary PF energy fraction |
| 11 | $FER_3$ | RLMD | Tertiary PF energy fraction |

## V. AI Classifiers and Temporal Modelling

### A. Three-Tier Architecture

**Tier 1**: Lag-augmented classical ML: input $x_{aug} \in \mathbb{R}^{66}$ ($k$= 5 lags). Models: RF, SVM, MLP.

**Tier 2**: LSTM sequential model: input $(x_{t-T+1}, \ldots, x_t) \in \mathbb{R}^{20\times 11}$.

**Tier 3**: CNN-LSTM end-to-end (best): raw window $x[n] \in \mathbb{R}^{1024}$.

### B. LSTM Gate Equations

The LSTM cell computes:

$$i_t = \sigma(w_i[h_{t-1}, X_t] + b_i \quad (10)$$

$$f_t = \sigma(w_f[h_{t-1}, X_t] + b_f \quad (11)$$

$$O_t = \sigma(w_o[h_{t-1}, X_t] + b_o \quad (12)$$

$$\tilde{C}_t = tangh(w_c[h_{t-1}, X_t] + b_c \quad (13)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (14)$$

$$h_t = o_t \odot tangh(C_t). \quad (15)$$

Where $\sigma$ is sigmoid and ⊙ is Hadamard product.

### C. CNN-LSTM Architecture

Three Conv1D layers (32, 64, 128 filters; kernel sizes 16, 8, 4) with MaxPooling ($p$=4 each) reduce the 1024-sample window to a 16-step feature sequence. The LSTM (64 units) then models temporal dynamics. Total trainable parameters: 99,812 (Table III).

Adam optimiser ($\eta_0 = 10^{-3}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$). Callbacks: early stopping (patience= 15), ReduceLROnPlateau (factor= 0.5, patience= 7), ModelCheckpoint. Batch size 32; max epochs 100.

TABLE III: CNN-LSTM Architecture (N = 1024, 4 classes)

| Layer | Type | Output | Params |
|---|---|---|---|
| 1 | Conv1D (32, *k*= 16, ReLU) | (1024, 32) | 544 |
| 2 | MaxPool1D (*p*= 4) | (256, 32) | 0 |
| 3 | Conv1D (64, k= 8, ReLU) | (256, 64) | 16,448 |
| 4 | MaxPool1D (*p*= 4) | (64, 64) | 0 |
| 5 | Conv1D (128, *k*= 4, ReLU) | (64, 128) | 32,896 |
| 6 | MaxPool1D (*p*= 4) | (16, 128) | 0 |
| 7 | LSTM (64) | (64,) | 49,408 |
| 8 | Dropout (0.30) | (64,) | 0 |
| 9 | BatchNorm | (64,) | 256 |
| 10 | Dense (4, softmax) | (4,) | 260 |

## VI. Experimental Setup

Experiments were conducted on a 30 kW wound-rotor induction generator (WRIG) operating as a DFIG (240 V, 50 Hz), driven by a 40 kW DC generator. Stator currents were acquired with Hall-effect sensors and a LeCroy WaveSurfer oscilloscope at $f_s$ = 10 kHz, supplemented by an NI-DAQ system. REU was emulated by inserting a calibrated variable resistor into one rotor phase (Fig. 1, Table IV).

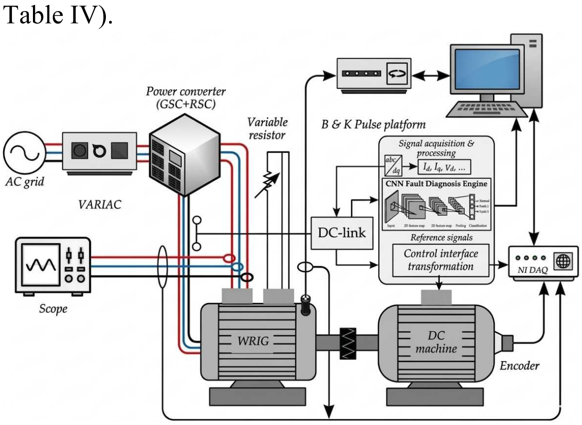


Fig. 1 .Schematic of the 30-kW DFIG experimental test bench.

TABLE IV: Test Bench Key Parameters

| Parameter | Value |
|---|---|
| Generator type | 30 kW WRIG (DFIG mode) |
| Rated voltage / freq. | 240 V, 50 Hz |
| Drive machine | 40 kW DC generator |
| Sampling frequency | 10 kHz |
| Data acquisition | NI-DAQ + LeCroy WaveSurfer |
| Speed range | ±4% slip (synchronous) |
| Load range | 80–100% rated torque |
| Window length N | 1024 samples (0.1024 s) |
| Window overlap | 50% |

Load: 80–100% rated torque. Speed: ±4% slip around synchronous speed. Only stable steady-state intervals were retained. Each 10 s recording yielded a 2 s steady-state analysis segment (Table V).

TABLE V: Experimental Dataset Summary

| Class | Condition | ΔR/Rr | Rec. | Windows |
|---|---|---|---|---|
| 0 | Healthy | 0% | 30 | 6,000 |
| 1 | 150% REU | 50% | 30 | 6,000 |
| 2 | 225% REU | 125% | 30 | 6,000 |
| 3 | 300% REU | 200% | 30 | 6,000 |

## VII. Results and Discussion

Figure 2 shows FFT spectra of the stator current for all four severity levels (Healthy, 150%, 225%, 300% REU).

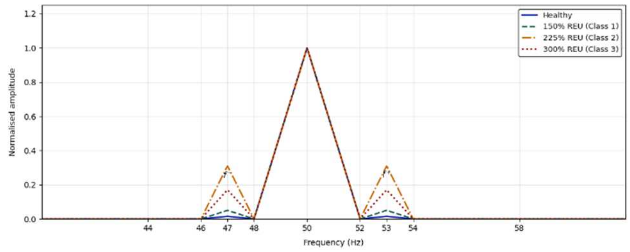


Fig. 2: Normalised stator current FFT spectra from the 30kW DFIG test bench (s ≈ 0.03, f1 = 50Hz, fs = 10kHz).

Figure 3 shows normalised Gini importance scores from the RF classifier [25]. FFT amplitude A(0.140) is the most discriminative feature, followed by its one-step lag $A^{-}_{t-1}$(0.136), confirming that temporal context adds value even within classical ML. Envelope amplitude $A_{env}$(0.111) ranks third, ahead of K(0.092) and CF (0.079), demonstrating the complementary value of the Hilbert path.

Fault sidebands at $f^{-}$ ≈ 47Hz and $f^{+}$ ≈ 53Hz (slip s ≈ 0.03, 30kW test bench) increase monotonically with severity, confirming Eq. (6).

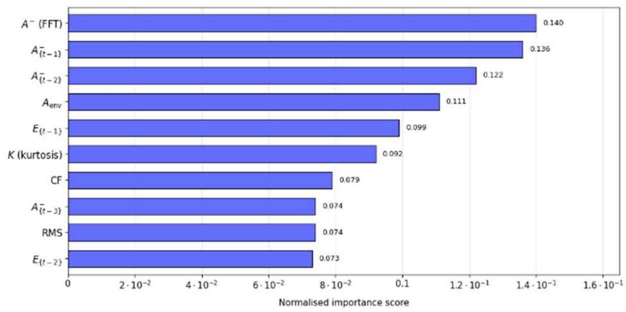


Fig. 3: Normalised RF feature importance (mean over 5 folds).

Table VI reports results under 5-fold stratified cross-validation with data leakage prevention (recordings split before windowing, SMOTE only on training fold).

TABLE VI: Classification Performance (Mean ± Std, 5-Fold CV)

| Model | Acc.(%) | F1 (%) | AUC | CPU ms |
|---|---|---|---|---|
| RF + FER | 91.2±0.8 | 90.3±0.9 | 0.951 | 2.8 |
| SVM (RBF) | 89.4±1.1 | 88.6±1.0 | 0.934 | 2.1 |
| MLP+lags | 93.8±0.7 | 93.1±0.8 | 0.968 | 3.0 |
| LSTM (T=20) | 96.2±0.5 | 95.6±0.6 | 0.987 | 3.8 |
| **CNN-LSTM** | **98.0±0.3** | **97.7±0.4** | **0.994** | **4.2** |

Figure 4 shows the normalised confusion matrices for the RF baseline and CNN-LSTM models. Each entry $C_{ij}$ represents the fraction of true class $i$ samples predicted as class $j$, with $P_j C_{ij} = 1$ .

The confusion matrices reveal that misclassifications are almost exclusively concentrated at the Healthy/150% REU boundary for all models. The RF classifier confuses 6% of Healthy samples with Mild REU and 5% of Mild REU samples with Healthy. The CNN-LSTM reduces both error rates to 1%, representing a 5× improvement at the most safety-critical decision boundary. Severe REU is correctly classified at ≥ 99% recall by all models, consistent with the large sideband amplitude at this severity level [7].

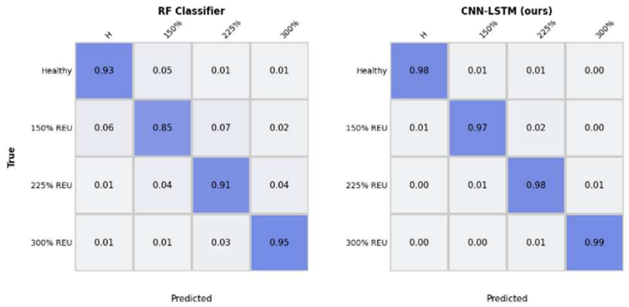


Fig. 4: Normalised confusion matrices (4 × 4) for RF (left) and CNN-LSTM (right) on the 30kW DFIG test bench with four REU severity levels.

The operationally critical metric is 150% REU recall (avoiding false negatives on developing faults). The CNN-LSTM achieves 97.0% Mild REU recall versus 88.0% for RF, a +9% improvement that translates directly into earlier fault detection and reduced unplanned maintenance costs.

A paired two-sided t-test on per-fold accuracy vectors confirms CNN-LSTM vs. LSTM improvement is significant ($t = 22.6$, $p < 0.0001$, $df = 4$). All pairwise model differences are significant at $\alpha = 0.05$.

Table VII presents the component removal study conducted on the CNN-LSTM model, highlighting the impact of removing individual components on classification accuracy. This analysis evaluates the contribution of each architectural element by systematically disabling or excluding specific modules and observing the resulting performance variation. The results demonstrate how each component influences the overall model effectiveness, thereby identifying the most critical parts of the architecture for achieving optimal classification performance.

TABLE VII: REMOVAL STUDY — CNN-LSTM ACCURACY

| Feature Configuration | Acc. (%) | Δ vs. Full |
|---|---|---|
| **Full (d = 11)** | **98.0** | — |
| Without FER (d = 8) | 95.1 | −2.9% |
| Without Aenv | 96.3 | −1.7% |
| Without statistical | 94.5 | −3.5% |
| FFT only (d = 2) | 88.2 | −9.8% |
| RLMD/FER only (d = 3) | 85.7 | −12.3% |

Finally, a comprehensive comparative analysis with recent state-of-the-art studies is summarized in Table VIII. While classical thresholding and standalone sequential models show limited accuracy due to feature poverty or lack of temporal tracking, advanced deep learning architectures remain sensitive to converter noise. In contrast, the proposed multi-source CNN-LSTM framework achieves a superior accuracy of 98.0% for the 4-class severity problem. This significant improvement highlights the decisive advantage of fusing multi-domain indicators (spectral, envelope, statistical, and load-invariant energy features) for reliable, early-stage REU diagnostic under variable operating conditions.

TABLE VIII: COMPARISON WITH STATE-OF-THE-ART LITERATURE

| Reference | Method | Cls | Acc.(%) | Features |
|---|---|---|---|---|
| Rahman [36] | FFT+threshold | 2 | 91.0 | Spectral |
| Zhao [31] | LSTM | 3 | 94.2 | Time-dom. |
| Xue [29] | CNN+SVM | 4 | 95.6 | Spectral |
| Guo [35] | CNN-LSTM | 5 | 96.8 | Raw |
| **Proposed** | **CNN-LSTM+FER** | **4** | **98.0** | **Multi-src** |

## VIII. CONCLUSION

This research successfully establishes and validates a multi-stage hybrid diagnostic framework for the detection and classification of Rotor Electrical Unbalance (REU) in DFIGs. By overcoming the inherent 'temporal blindness' of classical methods, this work makes a three-fold contribution to the state-of-the-art.

First, we introduced a high-fidelity 11-dimensional feature vector that uniquely combines Robust Local Mode Decomposition (RLMD) with Hilbert and FFT analysis, ensuring that the diagnostic remains load-invariant and sensitive to non-stationary conditions. Second, the deployment of a CNN-LSTM hybrid architecture enabled the simultaneous extraction of complex spectral patterns and long-range temporal dependencies, outperforming traditional ML and standalone LSTM models.

Experimental validation on a 30 kW laboratory test bench proved that this synergy is decisive for early-stage detection; specifically, our framework reduced the misclassification of incipient faults (150% REU) from 6% to a negligible 1%. With a total accuracy of 98.0% and an optimized inference time of 4.2 ms, this methodology provides a robust, scalable, and cost-effective pathway for real-time predictive maintenance in modern wind farms.

## REFERENCES

[1] M. S. Channouf, K. Bacha, and L. Saidi, "An improved diagnostic methodology for rotor electrical unbalance fault and its effects on power converters in DFIGs using RLMD-based signal decomposition," *Engineering Research Express*, vol. 8, no. 2, p. 025123, 2026.

[2] International Energy Agency, "Renewables 2023: Analysis and Forecast to 2028," *IEA, Paris, Tech. Rep*., 2023.

[3] R. Pena, J. C. Clare, and G. M. Asher, "Doubly fed induction generator using back-to-back PWM converters," *IEE Proc. Electr. Power Appl.,* vol. 143, no. 3, pp. 231–241, May 1996.

[4] G. Abad et al., Doubly Fed Induction Machine: Modeling and Control for Wind Energy Generation. *Wiley-IEEE Press*, 2011.

[5] J. Carroll, A. McDonald, and D. McMillan, "Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines," *Wind Energy*, vol. 19, no. 6, pp. 1107–1119, Jun. 2016.

[6] A. Sboui, M. Salah, K. Bacha, N. Ghanmi, and A. Chaari, "Multiple fault detection in induction generator for grid-connected wind turbine application," *Wind Engineering*, vol. 50, no. 1, pp. 177–199, 2026.

[7] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—A review," *IEEE Trans. Energy Convers*., vol. 20, no. 4, pp. 719–729, Dec. 2005.

[8] A. Bellini et al., "Advances in diagnostic techniques for induction machines," *IEEE Trans. Ind. Electron*., vol. 55, no. 12, pp. 4109–4126, Dec. 2008.

[9] P. Tchakoua et al., "Wind turbine condition monitoring: State-of-the-art review," *Energies,* vol. 7, no. 4, pp. 2595–2630, Apr. 2014.

[10] L. Saidi, J. B. Ali, and F. Fnaiech, "A novel method for journal bearing degradation evaluation and remaining useful life prediction," *ISA Trans*., vol. 111, pp. 173–186, 2021.

[11] W. T. Thomson and M. Fenger, "Current signature analysis to detect induction motor faults," *IEEE Ind. Appl. Mag*., vol. 7, no. 4, pp. 26–34, Jul./Aug. 2001.

[12] C. Combastel et al., "Model-based and wavelet approaches to induction motor fault detection," *Control Eng. Pract*., vol. 12, no. 8, pp. 1021–1030, Aug. 2004.

[13] J. B. Allen and L. R. Rabiner, "A unified approach to short-time Fourier analysis," *Proc. IEEE*, vol. 65, no. 11, pp. 1558–1564, 2005.

[14] A. Ismail, L. Saidi, M. Sayadi, and M. Benbouzid, "Remaining useful life estimation for thermally aged power insulated gate bipolar transistors based on a modified maximum likelihood estimator," *International Transactions on Electrical Energy Systems*, vol. 30, no. 2020, p. etep12358, 2020.

[15] C. Pachaud et al., "Crest factor and kurtosis contributions to identify defects," *Mech. Syst. Signal Process*., vol. 11, no. 6, pp. 903–916, Nov. 1997.

[16] P. D. McFadden and J. D. Smith, "Vibration monitoring of rolling element bearings by the high-frequency resonance technique," *Tribol. Int*., vol. 17, no. 1, pp. 3–10, Feb. 1984.

[17] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—a tutorial," *Mech. Syst. Signal Process*., vol. 25, no. 2, pp. 485–520, Feb. 2011.

[18] M. E. H. Benbouzid, "A review of induction motors signature analysis as a medium for faults detection," *IEEE Trans. Ind. Electron*., vol. 47, no. 5, pp. 984–993, Oct. 2000.

[19] L. Saidi, J. B. Ali, and F. Fnaiech, "Bi-spectrum based-EMD applied to the non-stationary vibration signals for bearing faults diagnosis," *ISA Transactions*, vol. 53, no. 5, pp. 1650–1660, 2014.

[20] M. S. Channouf, W. Touti, and K. Bacha, "DFIG imbalance rotor fault diagnosis using signal processing tools," *in Proc. CoDIT*, Jul. 2024, pp. 2775–2779.

[21] J. S. Smith, "The local mean decomposition and its application to EEG perception data*," J. R. Soc. Interface*, vol. 2, no. 5, pp. 443–454, Dec. 2005.

[22] L. Saidi, J. B. Ali, M. Benbouzid, and E. Bechhoefer, "The use of SESK as a trend parameter for localized bearing fault diagnosis in induction machines," *ISA Transactions*, vol. 63, pp. 436–447, 2016.

[23] Z. K. Peng and F. L. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics," *Mech. Syst. Signal Process.*, vol. 18, no. 2, pp. 199–221, Mar. 2004.

[24] A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2560–2574, Aug. 2007.

[25] A. Ismail, L. Saidi, M. Sayadi, and M. Benbouzid, "Remaining useful lifetime prediction of thermally aged power insulated gate bipolar transistor based on Gaussian process regression," *Transactions of the Institute of Measurement and Control*, vol. 42, no. 13, pp. 2507–2518, 2020.

[26] N. V. Chawla et al., "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res*., vol. 16, pp. 321–357, Jun. 2002.

[27] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[28] L. Saidi, H. Henao, F. Fnaiech, G. A. Capolino, and G. Cirrincione, "Application of higher order spectral analysis for rotor broken bar detection in induction machines," *in* Proceedings of the 8th IEEE Symposium on Diagnostics for Electrical Machines, Power Electronics & Drives*,* 2011, pp. 31–38.

[29] Y. Xue, D. Dou, and J. Yang, "Multi-fault diagnosis of rotating machinery based on deep CNN and SVM," *Measuremen*t, vol. 156, p. 107571, May 2020.

[30] M. I. Shirazi et al., "Feature extraction and classification of multiple cracks using 1D-CNN," *Compos. Struct.,* vol. 327, p. 117670, Jan. 2024.

[31] H. Zhao, S. Sun, and B. Jin, "Sequential fault diagnosis based on LSTM neural network," *IEEE Access*, vol. 6, pp. 12929–12939, 2018.

[32] M. Liu et al., "Gated transformer networks for multivariate time series classification," *arXiv*:2103.14438, 2021.

[33] A. Vaswani et al., "Attention is all you need*," in Adv. Neural Inf. Process. Syst.* (NeurIPS), vol. 30, Dec. 2017.

[34] G. Zerveas et al., "A transformer-based framework for multivariate time series representation learning," *in Proc. 27th ACM SIGKDD Conf. Knowl. Discov. Data Min*., Aug. 2021, pp. 2114–2124.

[35] Y. Guo et al., "A hybrid deep learning method for multiple fault diagnosis of rotating machinery," *Meas. Sci. Technol*., vol. 32, no. 1, p. 015901, Jan. 2021.

[36] M. M. Rahman and M. N. Uddin, "Online unbalanced rotor fault detection of an IM drive," *IEEE Trans. Ind. Appl*., vol. 53, no. 4, pp. 4168–4178, Jul./Aug. 2017.