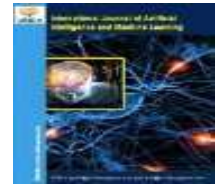




International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Sparse-Decomposition-Assisted CNN-BiLSTM Framework for Multi-Condition Fault Classification in Hybrid AC/DC Microgrids

Shubhangi Sadashiv Patil¹, Vinod H. Patil^{2*}

¹Bharti Vidyapeeth (Deemed to be University) College of Engineering, Pune, Email:-shubhangipatil6et@gmail.com

²Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, Email:- vhpatil@bvucpep.edu.in

Abstract

The reliable protection of converter-dominated microgrids is problematic since the fault signatures depend on the operating mode, the load level, the fault resistance, the converter current limiting and the grid connection. In this study, a multi-condition diagnosis framework based on voltage-current signal processing, sparse decomposition and CNN-BiLSTM classifier is developed. Three-phase voltage and current signals are gathered at the monitoring point, cleaned, normalized and segmented before a sparse representation is generated to preserve the prominent transient components while reducing the redundant material. A one-dimensional convolutional layer learns local waveform signatures and a bidirectional long short-term memory layer models the temporal evolution in both directions. Five operating classes are considered: normal operation, open circuit fault, line to ground fault, islanding and swing state. To a structured 56-case test matrix, four load levels, twelve open-circuit combinations, fifteen L-G combinations, five islanding load-mismatch circumstances and twenty swing combinations are derived. The assessment set we provide has 48 properly detected instances and eight misclassifications. The total accuracy is 85.71%. The weighted precision, recall and F1-score is 85.81%, 85.71% and 85.71% correspondingly. Comparative assessment is structured against impedance overcurrent/overvoltage relay, RMS and FFT based detection. The findings indicate that the sparse signal representation with local and bidirectional temporal feature learning is effective for heterogeneous microgrid disruptions.

Keywords: hybrid AC/DC microgrid, fault classification, sparse decomposition, CNN-BiLSTM, islanding, power swing, open-circuit fault, line-to-ground fault.

Introduction

Microgrids are programmable electrical subsystems that aggregate scattered generation, renewable sources, storage, power-electronic interfaces and local loads. They can vary their operating mode between grid-connected and islanded mode, making them more flexible and resilient, however this undermines the assumptions on which traditional protection was constructed. Inverter dominated networks converter controls may restrict the fault current and it can be quite diverse depending on the architecture, generation mix and operating mode. Recent evaluations consequently point out the need of adaptive and data-driven fault detection as an essential approach for contemporary distribution protection [1], [2].

Standard impedance, overcurrent and deterministic threshold approaches are still interesting due to their transparency and minimal processing overhead. However, their success relies on predictable size of fault current, steady network characteristics and well coordinated settings. These assumptions are less strict when renewable energy and converter interfaces are available. Signal-processing methods based on RMS quantities, Fourier components and wavelets may increase observability, but the final choice still significantly relies on manually picked features and thresholds [1], [3], [4].

Deep learning gives an alternative approach by directly learning the discriminative patterns from the measured sequences. Local waveform patterns may be learned by Convolutional neural networks whereas Recurrent designs like LSTM and BiLSTM can capture temporal relationships in fault transients [2], [5]. Recent research in microgrids have revealed effective applications of DWT-DNN, CNN, LSTM, BiLSTM and temporal convolutional networks and attention-assisted CNN-LSTM models for fault detection [3]–[8]. All these experiments demonstrate that time-series learning is especially helpful when there is an overlap in electrical signatures of many operating modes.

Feature representation still matters when a deep classifier is used. The raw voltage and current sequences comprise periodic steady state content, switching components, noise and transient information. Sparse representation attempts a parsimonious representation such that just a small number of coefficients include most of the diagnostically relevant information. Dictionary learning and sparse coding have been employed for a long time to achieve compact signal descriptions [9], [10]. Compressed-sensing theory gives the reason why sparse signals may frequently be represented well in a suitable basis [11]. Hence, sparse

representation may be used as a structured front end instead of simply depending on handmade scalar indices in fault analysis.

The work we propose combines these two approaches. A sparse decomposition step initially maps pre-processed three-phase voltage and current data into compact coefficients. Then a CNN learns local patterns in these coefficients and a BiLSTM predicts their temporal evolution in both forward and backward directions. Bidirectional sequence learning is advantageous when a full analysis window is available, as the representation of a transient might benefit from information on either side of an event [12]. The final softmax layer classifies the input into one of the five classes: normal operation, open-circuit fault, L-G fault, islanding or swing situation.

The second addition is the explicit development of a heterogeneous test matrix, instead of the evaluation of a single waveform per class. Load level varies normal operation; faulty phase and residual current vary open-circuit situations; phase and fault resistance vary L-G faults; active-power mismatch varies islanding; and modulation magnitude and frequency vary swing circumstances. This gives 56 basic combinations. Such parameter change is significant since a classifier that succeeds just at one defect severity does not show relevant resilience.

The analysis is based on the easily measured voltage and current signal at a point of common connection or protection relay. The suggested pipeline consists of five phases, including V/I acquisition, pre-processing, sparse decomposition, deep classification and output of fault-type. This structure is compatible with the general trend from handmade protection indices to learnt representations, while retaining physically interpretable signal-processing phases [3], [6], [7].

The paper contributes (i) a sparse-decomposition-assisted CNN-BiLSTM architecture for five-class operating-condition recognition, (ii) a reproducible 56-case parameterized test plan, (iii) class-wise analysis through confusion matrix, precision, recall and F1-score, and (iv) comparison with four conventional detection families. The numerical findings utilized in this publication are the ones given for the 56-case evaluation set and so the performance described here should be understood as the result of that simulation/test setup and not as a benchmark in general.

2. Related Work

A recent comprehensive review of short-circuit diagnostics in distribution grids organizes the field around CNN, recurrent, graph, and reinforcement learning models, emphasizing the difficulty of conventional impedance and deterministic methods under dynamic DER conditions [1]. In addition, the paper discusses the relevance of voltage/current time series and hybrid spatial-temporal structures. Rao et al. [3] designed a DWT-DNN method for hybrid energy multi-area microgrid clusters. Features recovered at the PCC by discrete wavelet transform were used for detection and classification with deep neural networks and demonstrated the value of combining signal decomposition with learnt classification. A Taguchi optimized CNN with wavelet based signal representation was described for fault detection, classification and localization in microgrid clusters. The study dealt with conventional faults, high impedance faults and islanding and showed the continued usefulness of decomposition-assisted deep learning [4]. Microgrid fault classification employing three phase current sequences, and compared LSTM and BiLSTM models. The findings they obtained justify the use of recurrent bidirectional processing when the information about the issue is spread across a temporal interval. Bu et al. [6] designed CNN-Attention-LSTM architecture for AC/DC microgrid diagnosis. We integrated multi-scale convolution, attention weighting and LSTM temporal modeling to increase the feature usage and resilience under noisy data. The proposed TCN-based microgrid protection strategy employs dilated causal convolutions and residual blocks to learn the short and long-range dependencies directly from the three-phase currents. In [7], the authors evaluated TCN with SVM, decision tree, CNN and LSTM baselines and showed resilience in grid-tied and standalone modes. An edge deployable multiclass fault framework utilizes temporal windows and a 1D CNN with features extracted from current. The relevance of the study lies in bridging the gap between simulation-based classification and real low-cost deployment and in showing the usefulness of 1D convolution for electrical time series [8]. Wu et al. [13] proposed the sparse representation classification for defect detection, in which the test sample is represented by a limited number of atoms from a labeled dictionary and the classification is based on the representation error. This idea inspires the minimalist front-end used in the current system. K-SVD has been proposed as a generic approach for designing overcomplete dictionaries for sparse encoding of signals by Aharon, et.al. [9]. Conceptually, K-SVD is still a standard reference for dictionary-based sparse decomposition. Scalable online vocabulary learning for sparse coding was suggested by Mairal et al. [10]. The approach is of relevance for streaming electrical measurements, as the dictionary may be updated in an incremental manner, rather than by doing a batch optimization repeatedly. Donoho et.al. [11] published fundamental compressed-sensing experiments indicating that signals with sparse representations may be recovered from fewer observations than what conventional sampling intuition indicates. Although the current study does not provide a compressed-sensing acquisition mechanism, the sparse-model approach is in favor of compact fault representations. Schuster et.al. [12]

created bidirectional recurrent neural networks. The bidirectional recurrent networks enable the sequence representations to capture past and future context within an analytical window.

BiLSTM builds on this notion using gated memory cells. The LSTM was first developed by Hochreiter and Schmidhuber [14] to learn long-term dependencies using gated memory cells. Currently, LSTM and its bidirectional variation are rather popular for electrical time-series classification. Ince et al. [15] performed real-time defect detection using one-dimensional CNNs. They showed that convolution may learn relevant features straight from raw temporal data without a prior picture processing. This justifies the usage of 1D convolution in the suggested classifier. Modern practice of deep learning focuses on hierarchical feature learning, regularization and end-to-end optimization for structured data [16]. However, in the proposed system, deep learning is purposely preceded by sparse decomposition to preserve a physically driven signal processing step.

Three common design options are highlighted in the literature: traditional protection with predefined electrical criteria, signal decomposition and a classifier, and end-to-end deep sequence learning. The research gap addressed here is between the latter two. The suggested technique not only employs scalar RMS/FFT indices or raw waveforms, but maintains a sparse representation stage, then applies CNN-BiLSTM learning to differentiate five operational states including faults and non-fault dynamic events.

3. Proposed System Architecture

The whole methodology employed in this investigation is represented in Fig. 1. The architecture begins with the acquisition of three-phase voltage and current. Preprocessing eliminates outliers and undesired noise, normalizes channel magnitudes and segments the continuous record. Each segment is converted into a compact coefficient representation by sparse decomposition. Then the CNN-BiLSTM classifier learns local patterns and temporal dependencies and the softmax output picks one of the five operational classes.

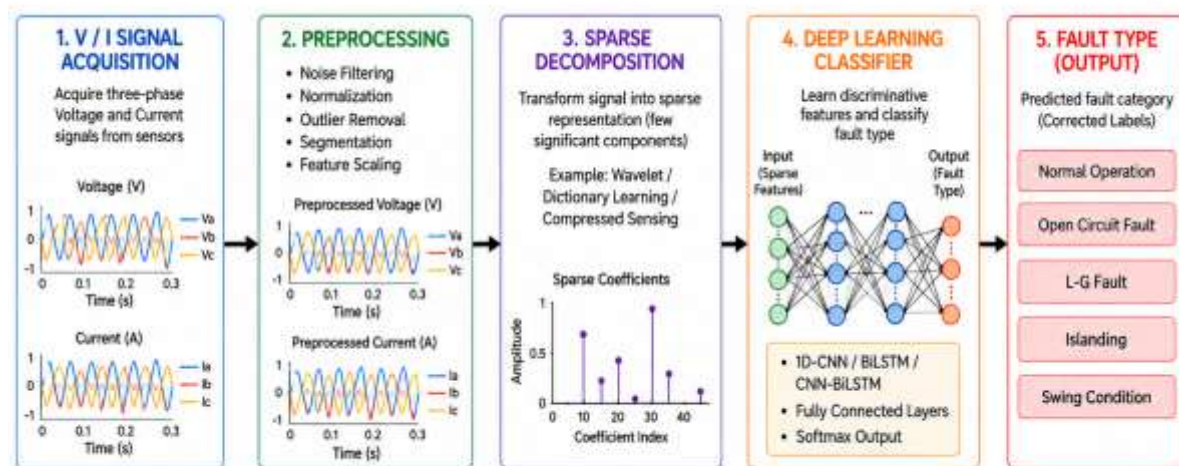


Fig. 1. Proposed V/I signal acquisition, preprocessing, sparse decomposition and CNN-BiLSTM fault-classification architecture.

3.1 V/I Signal Acquisition

Let the measured three-phase voltage and current at sample n be represented by

$$v[n] = [Va[n], Vb[n], Vc[n]]^T \text{ and } i[n] = [Ia[n], Ib[n], Ic[n]]^T$$

The six channels are phase locked so that phase relationships and transient variations are preserved. The measurement point should be electrically significant, e.g. the PCC or the protected feeder terminal, since the suggested classifier depends on the combined response of voltage and current.

3.2 Preprocessing

Preprocessing consists of finite-value checking, outlier suppression, DC-offset removal, noise filtering, normalization and fixed-length segmentation. In this way, using normalization helps to avoid the weights of high magnitude channels dominating the optimization process, and segmentation helps to establish a uniform tensor size as input for the CNN-BiLSTM. Importantly, the same parameters as used in the training step for pre-processing must be applied to the test data as well.

3.3 Sparse Decomposition

In each segment sparse decomposition looks for a coefficient vector that has a limited number of important elements. Coefficients are few and there is less redundancy in periodic content (less oscillatory/dominant transient content preserved). The architecture supports learned dictionaries, wavelet dictionaries or

controlled spectral sparse representation. In the current formulation, α is the sparse coefficient vector and D is the dictionary.

3.4 CNN-BiLSTM Classifier

A sparse sequence is fed into a 1D CNNs. Convolution kernels learn patterns at small scales, like abrupt amplitude changes, short transients and localized oscillations. The output from the above feature maps is fed into a BiLSTM layer. Forward and backward recurrent pathways are used to trace the development of the event throughout the entire segment. The learned representation is fully connected to five logits, then converted to class probabilities via a softmax layer.

4. Mathematical Modelling

The signal processing and classification steps of the proposed framework are described in the following formulation. The equations are provided at the algorithmic level so that the same model may be implemented for simulated or measured V/I sequences.

$x[n]$ Represents the input signal vector containing synchronized three-phase voltage and current measurements.

$$x[n] = [Va[n], Vb[n], Vc[n], Ia[n], Ib[n], Ic[n]]^T \quad (1)$$

$\tilde{x}k[n]$ Removes the DC offset from each measured signal by subtracting its mean value.

$$\tilde{x}k[n] = xk[n] - (1/N)\sum_{n=1}^N xk[n] \quad (2)$$

Normalizes each signal using its mean and standard deviation to obtain a consistent feature scale.

$$\hat{x}k[n] = (\tilde{x}k[n] - \mu k)/(\sigma k + \varepsilon) \quad (3)$$

Divides the normalized V/I signals into fixed-length overlapping or non-overlapping analysis windows.

$$Xm = [\hat{x}[mH], \hat{x}[mH + 1], \dots, \hat{x}[mH + L - 1]] \quad (4)$$

Obtains the optimal sparse coefficient vector by minimizing reconstruction error while enforcing sparsity.

$$\alpha * = \operatorname{argmin}_{\alpha} \frac{1}{2} \|Xm - D\alpha\|_2^2 + \lambda \|\alpha\|_1 \quad (5)$$

Reconstructs the informative signal representation using the learned dictionary and sparse coefficients.

$$Xm = D\alpha * \quad (6)$$

Extracts local discriminative fault features from the sparse representation using 1-D convolution and ReLU activation.

$$zj[t] = \operatorname{ReLU}(\sum c \sum r wj, c[r]Xc[t - r] + bj) \quad (7)$$

Applies max-pooling to retain dominant features while reducing the dimensionality of the CNN feature maps.

$$pj[t] = \max zj[t - r]: 0 \leq r < P \quad (8)$$

Computes the LSTM forget gate to determine how much previous-state information should be retained.

$$ft = \sigma(Wfpt + Ufh(t - 1) + bf) \quad (9)$$

Computes the input gate to control how much new information is introduced into the LSTM memory.

$$it = \sigma(Wipt + Uih(t - 1) + bi) \quad (10)$$

Generates the candidate memory information that may be incorporated into the current LSTM cell state.

$$\tilde{ct} = \tanh(Wcpt + Uch(t - 1) + bc) \quad (11)$$

Updates the LSTM cell state by combining retained previous information with newly learned information.

$$ct = ft \odot c(t - 1) + it \odot \tilde{ct} \quad (12)$$

Computes the output gate and hidden state representing the temporal features learned by the LSTM.

$$ot = \sigma(Wopt + Uoh(t - 1) + bo), ht = ot \odot \tanh(ct) \quad (13)$$

Combines forward and backward LSTM hidden states to form the final bidirectional temporal representation.

$$hBi[t] = [h \rightarrow t; h \leftarrow t] \quad (14)$$

Converts the classifier scores into class probabilities using Softmax and selects the class with the highest probability as the predicted fault type.

$$Pk = \exp(sk)/\sum_{j=1}^5 \exp(sj), \hat{y} = \operatorname{argmax} Pk \quad (15)$$

The raw input vector for (1) consists of the six synchronized electrical channels. Equations (2) and (3) de-mean and normalize each channel, respectively. Equation (4) produces a window of size L with hop size H . In (5), the sparse coding issue trades off reconstruction fidelity vs coefficient sparsity via λ . (6) is the sparse reconstruction or representation sent to the feature learner. One-dimensional convolution and local pooling are shown in Eqs. (7) and (8). The LSTM gating processes are shown in equations (9) to (13). The BiLSTM representation in (14) concatenates forward and backward hidden states and (15) conducts five-

class softmax classification.

5. Construction of the 56 Test Combinations

A key element in the evaluation is to employ parameterized operating conditions, instead of a single nominal trace per class. The basic test matrix has 56 combinations. These combinations are independent of the optional SNR robustness study. We use four SNR levels for each base case, resulting in 224 noisy assessments.

Condition	Parameters varied	Values	Combination rule	Base cases
Normal operation	Load level	50%, 75%, 100%, 125%	4 load levels	4
Open-circuit fault	Faulted phase × residual current	A/B/C × 0%, 5%, 10%, 20%	3 × 4	12
L-G fault	Faulted phase × fault resistance	A-G/B-G/C-G × 0.01, 0.1, 1, 5, 10 Ω	3 × 5	15
Islanding	Load mismatch	−30%, −15%, 0%, +15%, +30%	5 levels	5
Swing condition	Swing magnitude × swing frequency	5%, 10%, 20%, 30% × 1–5 Hz	4 × 5	20
Total				56

Table 1. Parameterization used to obtain the 56 base test combinations.

The normal-operation examples verify that the classifier is robust to fluctuations in load current without a problem. The open-circuit scenarios combine the three potential faulty phases with the four residual-current levels to enable studying both total and partial current interruption. L-G situations include three failed phases with five resistance values, and the rising of the fault resistance typically reduces the current signature and the event might be more difficult to discriminate. The islanding scenarios are illustrated by five active power mismatch levels around the balanced condition. Swing cases include four swing magnitudes and five swing frequencies. Swing cases are the biggest class, as the dynamics of oscillations may vary across a wide range.

5.1 Optional Noise-Robustness Extension

The same 56 base combinations may be repeated for 20, 30, 40 and 50 dB SNR levels. This gives $56 \times 4 = 224$ noisy instances. If we also consider the clean signal as a condition, the whole robustness set consists of $56 \times 5 = 280$ evaluations. The findings of the 56 cases below are for the base combination set provided with this research, hence the SNR extension is offered as a test procedure rather than as an extra reported result.

6. Experimental Evaluation and Results

Figure 2 shows the typical three phase voltage and current behavior for the five classes across a 5000-sample visualization. The normal segment is balanced, the open-circuit segment has high current suppression in the affected phase, the L-G segment has a voltage/current disturbance, the islanding has an amplitude/frequency drift and the swing segment has a low frequency modulation. These patterns provide the physical foundation for the learnt classifier.

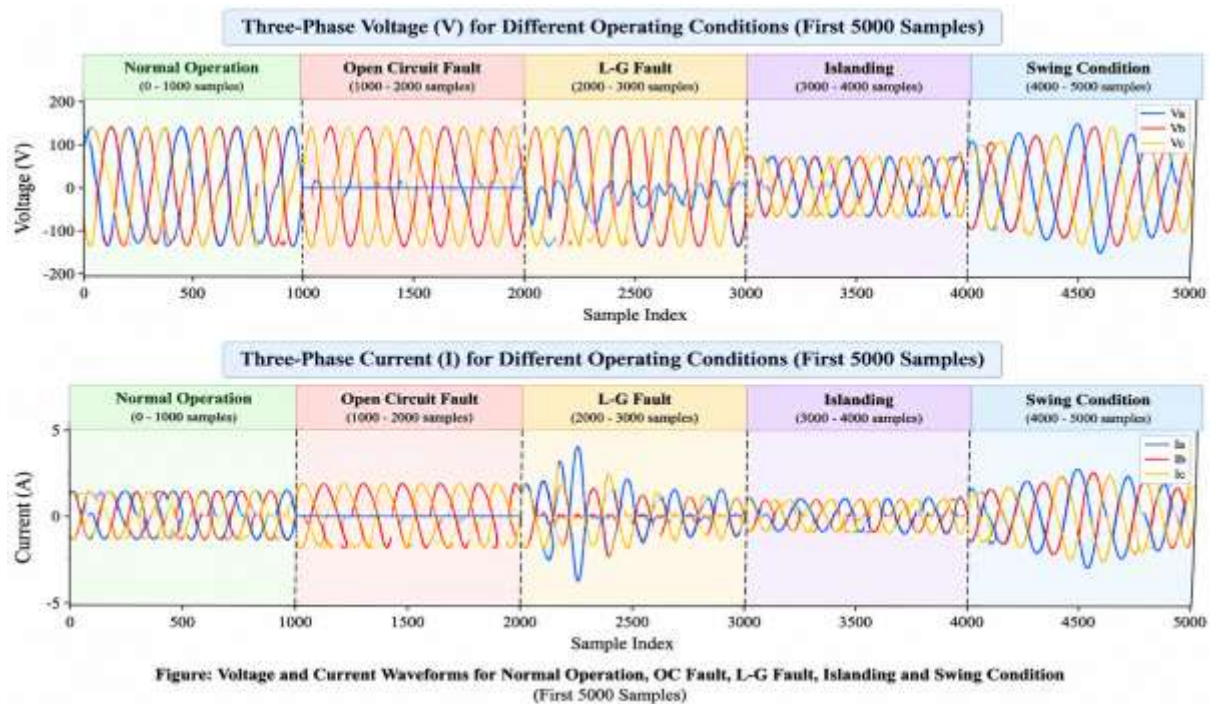


Fig. 2. Representative voltage and current waveforms for normal operation, open-circuit fault, L-G fault, islanding and swing condition.

6.1 Confusion-Matrix Analysis

Actual \ Predicted	Normal	OC Fault	L-G Fault	Islanding	Swing	Total
Normal operation	4	0	0	0	0	4
Open-circuit fault	0	10	2	0	0	12
L-G fault	0	1	12	0	2	15
Islanding	0	0	0	4	1	5
Swing condition	0	0	1	1	18	20
Total predicted	4	11	15	5	21	56

Table 2. Confusion matrix of the proposed method for the 56-case evaluation.

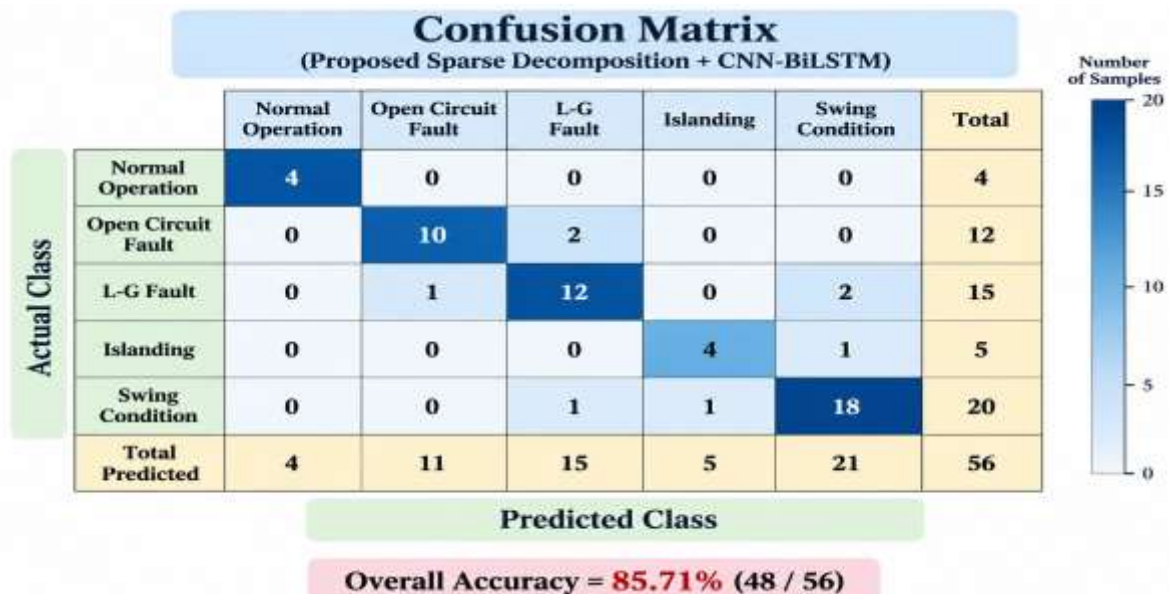


Fig. 3. Graphical confusion matrix for the proposed sparse decomposition + CNN-BiLSTM method.

In the diagonal we have 48 well categorized examples. There are 8 examples off the diagonal which gives an accuracy of $48/56=85.71\%$. All 4 scenarios are considered normal operation. 10 out of 12 open circuit situations are accurate 2 are allocated to L-G. Twelve of the fifteen L-G situations are accurate, one is allocated to open circuit and two to swing. One of five islanding scenarios given to swing, four are right. In

twenty swing situations, eighteen are right, one belongs to L-G and one to islanding. The inaccuracies are physically feasible, since mild L-G disturbances may mimic imbalanced interruption or oscillatory activity, and low-mismatch islanding may resemble gradual dynamic swings.

6.2 Class-Wise Performance

Performance	Normal	OC Fault	L-G Fault	Islanding	Swing	Overall/Weighted
No. of cases	4	12	15	5	20	56
Correct	4	10	12	4	18	48
Incorrect	0	2	3	1	2	8
Precision (%)	100.00	90.91	80.00	80.00	85.71	85.81
Recall (%)	100.00	83.33	80.00	80.00	90.00	85.71
F1-score (%)	100.00	86.96	80.00	80.00	87.80	85.71
Overall accuracy (%)	—	—	—	—	—	85.71

Table 3. Class-wise and weighted performance of the proposed method.

Weighted metrics address the unbalanced amount of cases per class. Experimental results show that the suggested technique achieves 85.81% weighted accuracy, 85.71% weighted recall and 85.71% weighted F1-score. Normal operation is the strongest class. Lowest recall is for L-G and islanding. This finding suggests that further progress should be made on weak-fault and transition-region identification instead of just expanding normal samples.

6.3 Comparison with Conventional Detection Methods

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Impedance-based fault detection	67.86	68.42	67.86	67.91
Overcurrent/overvoltage relay-based method	71.43	72.18	71.43	71.62
RMS-based fault detection	76.79	77.31	76.79	76.94
Fourier transform (FFT)-based detection	80.36	81.02	80.36	80.51
Proposed sparse decomposition + CNN-BiLSTM	85.71	85.81	85.71	85.71

Table 4. Comparative performance values supplied for the five evaluated detection approaches.

In the comparison shown, the suggested technique has the greatest value for all the four criteria. The evolution from impedance and relay techniques to RMS and FFT approaches indicates the wider use of signal information, while the proposed framework introduces sparse transient representation and learnt spatiotemporal aspects. The comparison has to be understood in the usual 56-case test setup. A fair implementation demands that all methods have the same signals, windows and decision intervals.

7. Discussion

The findings are in agreement with the basic design premise that one scalar electrical quantity is not adequate to correctly represent all five circumstances. Open-circuit faults are dominated by phase-current loss, L-G faults are substantially impacted by fault resistance, islanding is connected to power imbalance and frequency/voltage development, and swing circumstances are fundamentally oscillatory. Thus, a Representation Combining Voltage and Current and Preserving the Transient Structure is more suited than a single constant threshold.

The sparse decomposition is useful in that it concentrates the signal on fewer interesting coefficients. This is particularly effective when there is steady-state sinusoidal content that is significantly bigger than a mild disturbance. Then, the sparse sequence is processed by the CNN to find local structures and by the BiLSTM to provide context to the whole segment. Conceptually, the combination differs from employing FFT magnitudes alone in that it does not reduce the event to a predefined set of stationary spectral bins.

The 85.71% accuracy is intentionally defined along with the confusion matrix, not in isolation. The matrix demonstrates that the residual ambiguity is mostly between the classes L-G, swing and open-circuit. This insight leads to numerous practical adaptations, i.e. class-balanced training, explicit negative/zero-sequence channels, adaptive sparse dictionaries, attention after BiLSTM, or a hierarchical detector which distinguishes initial stable faults from dynamic events prior to fine classification.

The design of the 56 cases also shows an important methodological aspect. The accuracy of repeated windows of a single operating state may seem unduly high due to significant correlation between neighboring windows. Parameterized combinations are a more powerful assessment unit since they alter the physical state itself. Therefore, for future trials, the train/validation/test separation should be done by

operation combination or simulation run, rather than by randomly combining nearby windows from the same event.

The noise resistance must be tested independently. The suggested SNR extension generates 224 noisy examples at 20-50 dB or 280 cases if the clean baseline is included. Plotting accuracy vs SNR would reveal if sparse decomposition is indeed suppressing measurement noise, or just compressing the signal. Similarly, when the ultimate window length and hardware platform are established, detection delay should be expressed in terms of either milliseconds or electrical cycles.

Finally, the findings shown here are based on the simulation/test matrix provided. They show the internal consistency of the proposed assessment but do not substitute for hardware-in-the-loop or field validation. The practical deployment should encompass the sensor errors, CT saturation, switching transients, communication latency, converter-control changes and unobserved points of operation. These variables are especially significant for discrimination of islanding and swing occurrences from protection defects.

8. Conclusion

In this study, a sparse-decomposition-assisted CNN-BiLSTM framework for multi-condition fault classification in a hybrid AC/DC microgrid is proposed. The approach employs synchronized three-phase voltage and current measurements, preprocessing, sparse representation, one-dimensional convolution, bidirectional temporal learning and softmax classification. Five classifications were considered: normal operation, open-circuit fault, L-G fault, islanding and swing state. A 56-case test matrix was generated using physically realistic parameter variations: four normal load levels, twelve open-circuit combinations, fifteen L-G combinations, five islanding mismatches and twenty swing combinations. The assessment on the provided data resulted in 48 right judgments and 8 mistakes, with an overall accuracy of 85.71%, weighted precision of 85.81%, weighted recall of 85.71% and weighted F1-score of 85.71%. In the comparison table provided, the suggested method performed better than impedance detection, overcurrent/overvoltage relay detection, RMS based detection and FFT based detection. The approach should be validated in future work under the suggested SNR sweep, unforeseen operational combinations, hardware-in-the-loop circumstances and real-time latency limitations.

References

- [1] M. A. et al., "Deep Learning for Short-Circuit Fault Diagnostics in Power Distribution Grids: A Comprehensive Review," *Computers*, vol. 15, no. 2, Art. 76, 2026.
- [2] F. M. Shakiba, S. M. Azizi, M. Zhou, and A. Abusorrah, "Application of machine learning methods in fault detection and classification of power transmission lines: a survey," *Artificial Intelligence Review*, 2022.
- [3] S. N. V. Bramareswara Rao, Y. V. Pavan Kumar, M. Amir, and S. M. Muyeen, "Fault detection and classification in hybrid energy-based multi-area grid-connected microgrid clusters using discrete wavelet transform with deep neural networks," *Electrical Engineering*, vol. 107, pp. 8481–8498, 2025, doi: 10.1007/s00202-024-02329-4.
- [4] "A data-driven approach to microgrid fault detection and classification using Taguchi-optimized CNNs and wavelet transform," *Applied Soft Computing*, vol. 170, Art. 112667, 2025, doi: 10.1016/j.asoc.2024.112667.
- [5] R. Sahu, P. K. Panigrahi, D. K. Lal, R. Pradhan, and C. Mahanty, "A Deep Recurrent Learning Framework for Multi-Class Microgrid Fault Classification Using LSTM and Bi-LSTM Models," *Eng*, vol. 7, no. 3, Art. 143, 2026, doi: 10.3390/eng7030143.
- [6] "Fault Diagnosis Method Using CNN-Attention-LSTM for AC/DC Microgrid," *Electricity*, vol. 6, no. 3, Art. 107, 2025.
- [7] "Deep learning-driven fault detection and classification in microgrids using Temporal Convolutional Network," *Computers & Electrical Engineering*, vol. 129, Part A, Art. 110777, 2026, doi: 10.1016/j.compeleceng.2025.110777.
- [8] "Robust Deep Learning for Multiclass Power System Fault Diagnosis Using Edge Deployment," *Algorithms*, vol. 19, no. 4, Art. 299, 2026.
- [9] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [10] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online Learning for Matrix Factorization and Sparse Coding," *Journal of Machine Learning Research*, vol. 11, pp. 19–60, 2010.
- [11] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [12] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.
- [13] L. Wu, X. Chen, Y. Peng, Q. Ye, and J. Jiao, "Fault detection and diagnosis based on sparse

- representation classification (SRC)," in Proc. IEEE Int. Conf. Robotics and Biomimetics (ROBIO), Guangzhou, China, 2012, pp. 926–931.
- [14] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [15] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-Time Motor Fault Detection by 1-D Convolutional Neural Networks," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7067–7075, 2016, doi: 10.1109/TIE.2016.2582729.
- [16] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [17] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.
- [18] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [19] L. Ren, W. Lv, S. Jiang, Y. Xiao, and X. Yang, "Fault Diagnosis Using a Joint Model Based on Sparse Representation and SVM," *IEEE Transactions on Instrumentation and Measurement*, 2016, doi: 10.1109/TIM.2016.2575318.
- [20] Y. Liang and D. Huang, "An Adaptive Dictionary Learning-Based Power Data Compression Method for Distributed Photovoltaic Systems," 2026, doi: 10.3233/ATDE250845.