Original scientific paper
https://doi.org/10.33180/InfMIDEM2024.105

Informacije MIDEM
Journal of Microelectronics,
Electronic Components and Materials
Vol. 54, No. 1(2024), 51 – 63

Fault Diagnosis of Asymmetric Cascaded Multilevel Inverter using Ensemble Machine Learning

Kavitha Rangasamy

Department of Electrical and Electronics Engineering, Kumaraguru College of Technology, Coimbatore, Tamilnadu, India

Abstract: Cascaded Multi-Level Inverters (CMLI) are used in a wide range of high-power industrial drives and for integrating solar PV system. Asymmetric Cascaded Multilevel Inverter (ACMLI) produces an output voltage with reduced Total Harmonic Distortion (THD) when compared to Symmetric Cascaded Multilevel Inverter (SCMLI). ACMLI comprises of more semiconductor devices and thus reliability is a major concern. Efficient, high speed and precise fault detection is required for ACMLI to reduce failure rates and avoid unplanned shutdown. RMS voltage, mean voltage and THD under various single and double switch fault conditions are used as features for fault diagnosis. Fault diagnosis method for ACMLI based on probabilistic principal component analysis (PPCA) and Ensemble Machine Learning (EML) is presented. PPCA is used to optimize data and reduce the size of fault features. Finally, an EML classifier combining Support Vector Machine (SVM), K-Nearest Neighborhood (KNN) and Decision Tree (DT) is employed to diagnose the various open circuit faults. The proposed fault diagnosis method is validated using an experimental setup. The simulation and experimental result shows that EML technique diagnosis the fault with 99.32% accuracy.

Keywords: Multilevel Inverter; fault; Principal Component Analysis (PCA); SVM; Ensemble Machine Learning

Diagnostika napak asimetričnega kaskadnega večnivojskega pretvornika z uporabo skupinskega strojnega učenja

Izvleček: Kaskadni večnivojski pretvorniki (CMLI) se uporabljajo v številnih visokozmogljivih industrijskih pogonih in za integracijo fotonapetostnega sistema. Asimetrični kaskadni večnivojski pretvornik (ACMLI) proizvaja izhodno napetost z manjšim skupnim harmonskim popačenjem (THD) v primerjavi s simetričnim kaskadnim večnivojskim pretvornikom (SCMLI). ACMLI vsebuje več polprevodniških naprav. Za ACMLI je potrebno učinkovito, hitro in natančno odkrivanje napak, da se zmanjša število okvar in prepreči nenačrtovana zaustavitev. Napetost RMS, srednja napetost in THD pri različnih okvarah z enim in dvema stikaloma se uporabljajo kot značilnosti za diagnosticiranje okvar. Predstavljena je metoda diagnosticiranja napak za ACMLI, ki temelji na verjetnostni analizi glavnih komponent (PPCA) in skupinskem strojnem učenju (EML). PPCA se uporablja za optimizacijo podatkov in zmanjšanje velikosti značilnosti napak. Za diagnosticiranje različnih napak odprtega tokokroga je uporabljen klasifikator EML, ki združuje podporni vektorski stroj (SVM), K-najbližje sosedstvo (KNN) in odločitveno drevo (DT). Predlagana metoda za diagnosticiranje napak je potrjena z eksperimentalno postavitvijo. Simulacija in eksperimentalni rezultati kažejo, da tehnika EML diagnosticira okvare z 99,32-odstotno natančnostjo.

Ključne besede: večnivojski pretvornik; napaka; analiza glavnih komponent (PCA); SVM; skupinsko strojno učenje

** Corresponding Author's e-mail: kavithain2@gmail.com*

How to cite:
K. Rangasamy, "Fault Diagnosis of Asymmetric Cascaded Multilevel Inverter using Ensemble Machine Learning", Inf. Midem-J. Microelectron. Electron. Compon. Mater., Vol. 54, No. 1(2024), pp. 51–63

1 Introduction

CMLI is one of the most prominent converter topologies for renewable energy-based distribution system and high-power industrial drive applications because of its excellent scalability, low harmonic distortions, modular topology, and high efficiency [1]. As the number of voltage steps at output terminals increases, harmonic distortion reduces. However, this inherently leads to an increase in the number of power semiconductor devices [2]. ACMLI topology in natural sequence ratio (1:2:3) provides a higher number of voltage levels with the symmetrical arrangement of power semiconductor switching devices. Reduced number of dc sources, high-speed capabilities, minimal switching loss, and high conversion efficiency are the prominent advantages of the ACMLI topology. The circuit topology becomes more complicated when the number of H-bridge cells grows exponentially, and thus the possibility of a power semiconductor device failure increases, resulting in abnormal working circumstances [3].

Short Circuit Fault (SCF) and Open Circuit Fault (OCF) in semiconductor devices are the general types of faults in power devices and account for 38% of errors. Short circuit faults happen in short duration and are extremely destructive, which causes serious impacts. SCF's are transformed into OCF by a fast-acting fuse. OCF degrades inverter performance, distorts the output currents, and causes considerable risk of secondary faults in the load, grid, and converter systems. While applying CMLI to PV systems, if CMLI faults remain unidentified on time, it adversely impacts reliability, and even triggers fires. Hence, it becomes essential to design and develop intelligent fault diagnosis system to provide reliable operation of ACMLI systems [4-8].

Fault diagnostic approaches are based on current and voltage-based methods. Current based method detects OC faults through the phase current, and current residual [9]. The voltage-based fault diagnosis approach locates the fault by combining the output voltage and the diagnosis model. Line voltage error-based method is proposed for open circuit faults; however, it requires higher sampling time. The open switch faults are detected using a Fourier series and a histogram of the trajectory of voltage [10].

The data-driven fault detection method extracts voltage signals and applies signal processing methods and classification algorithms. This method provides attractive solutions due to the progress in Machine Learning (ML) algorithms and computation capability [11-12]. Fast Fourier Transform (FFT) with a hidden Markov model provides slower response in defect detection due to the sophisticated matrix operations [13].

The defect diagnostic approach for a three-level inverter utilizing a Back Propagation Neural network (BPN) and genetic algorithm is presented [14]. FFT analysis is employed to extract fault information, relative PCA and support vector machine (SVM) are employed to diagnose OCFs for SCMLI [15]. PCA and Multiclass-Relevance Vector Machine (MRVM) are applied to lower the facet of the OCF in symmetric CMLI [16].

The approach incorporates the grey wolf algorithm linked with enhanced SVM technology. The methodology uses the grey wolf algorithm to optimize the SVM parameters [17]. SVM and KNN with PPCA are used to detect the OCF under different fault circumstances in a CMLI. The SVM based fault analysis technique is effective and presents 97.6% accuracy with 70% training dataset [18].

Sparse representation with deep convolutional neural networks (DCNNs) is applied to provide an intelligent fault detection method for CMLI. By harnessing the beneficial characteristics of both approaches, the technique aims to improve defect identification efficiency and accuracy [19].

Different ML techniques, like KNN, SVM and Naive Bayesian classifier are applied. THD, RMS and mean voltage, and harmonics up-to 12^{th} order are considered as characteristic features. The classification accuracy achieved using CMLI is 95.56%, and that of Packed U-cell (PUC) inverters is 94.28% [20]. Affine-Invariant Riemannian Metric Autoencoder Random Forest (AIRMAR) is proposed to recognize OCF in MLI [21].

To enhance the features for classification, this model employs a short-time Fourier transform (STFT), that generates a time-frequency image. Multiscale Kernel Convolution Neural Network (MKCNN) is employed that utilizes the benefits of 2D images to capture the spatial relations among diverse features. MKCNN is compared with the Back Propagation Neural Network (BPN) and CNN [22].

The shift in pole voltage when a fault occurs is investigated and Rolling average of pole voltage is extracted. Pole voltage categories and reference voltage are the input characteristics for a decision tree classifier to identify single and double switch OCF [23].

An Adaptive Hilbert-Huang filter using Scale-Invariant Feature Transform (SIFT) with a Convolution Neural Network (CNN) classifier is employed [24]. Artificial neural network (ANN) is used to detect and diagnose faults in both binary and trinary configurations in an ACMLI [25]. Most of the literature has addressed fault diagnosis in SCMLI topology, and only very few have focused on fault diagnosis in ACMLI topology.

The following are the primary contributions of the paper: (i) Developing an ML based fault diagnosis technique with greater classification accuracy that detects both single and double switch fault scenarios (ii). The proposed technique can be extended for fault diagnosis of any CMLI configuration (iii) Simulation of the ACMLI has been performed in both normal and faulty conditions. (iv) Experimental implementation of the proposed technique is validated for single and double switch fault classification in ACMLI.

2 Fault diagnosis system

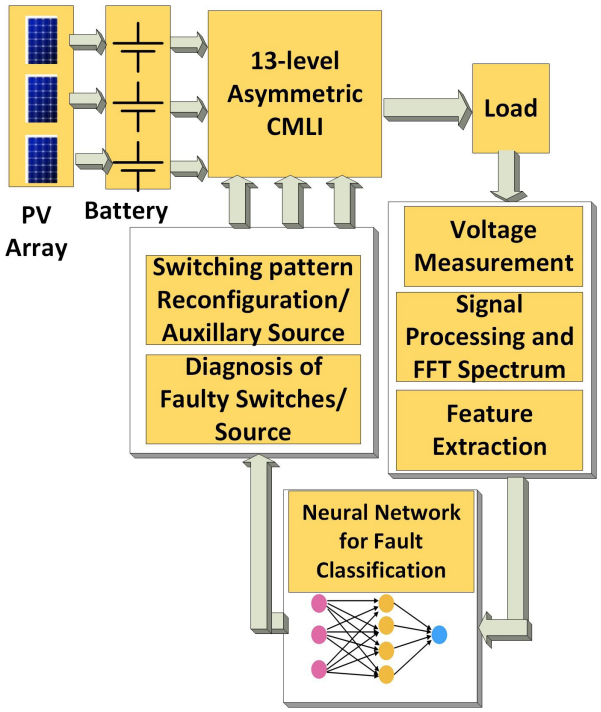


Figure 1: Structure of fault diagnosis system

Fig. 1 depicts the structure of a fault diagnosis system. The following four major states build up the system: 1) Feature extraction and dimension reduction; 2) Neural Network based fault classification; 3) Fault diagnosis; 4) Switching pattern reconfiguration. The voltage input signal transformation is performed to extract essential characteristics, and the output is directed to neural network classification. The network is trained with normal scenario and fault scenario, and each matching output is assigned a binary label.

FFT is a digital signal processing microchip implementation approach. FFT approach has a strong identity feature that allows it to distinguish between normal and irregular fault signals. The precisely assembled information in fig. 2. lays the foundation for developing a reliable fault identification and diagnosis system.

The features used for fault classification are: FFT, RMS voltage (Vrms), mean voltage (Vmean), THD and harmonics (up-to 19th order). These values are obtained under normal, single switch fault, double switch fault conditions. Mean voltage indicates a shift and change in symmetrical pattern during faulty condition. RMS voltage provides information on the reduction in peak value due to fault. THD and lower order dominant harmonics (up-to 19^{th}) are considered as they reveal the distortion in waveform shape.

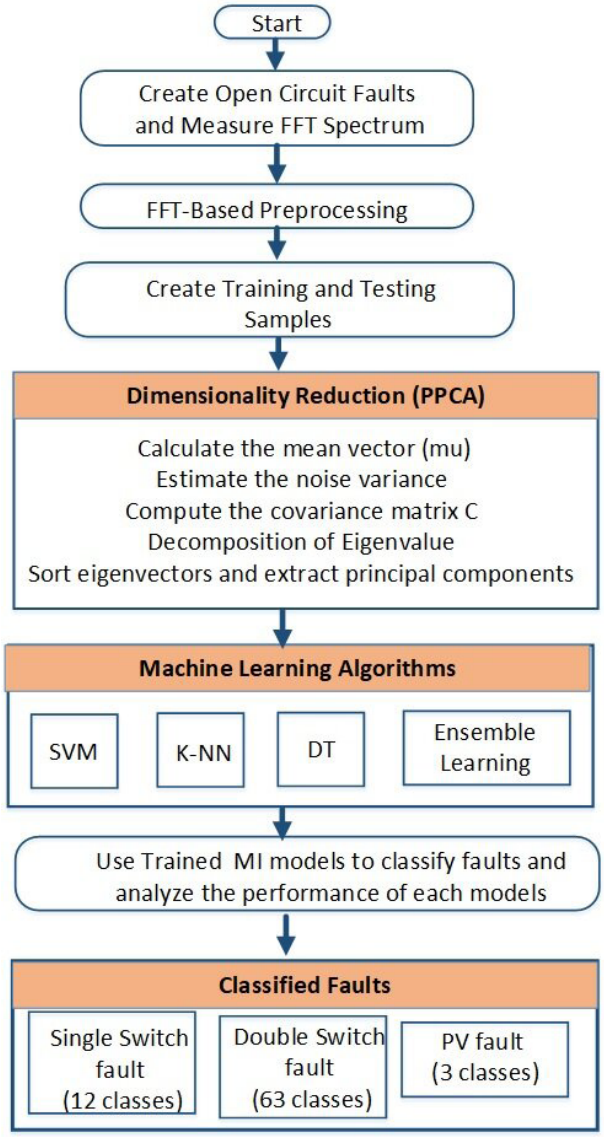


Figure 2: Flow chart of fault diagnosis system

3 ACMLI

The single-phase thirteen-level ACMLI is shown in Fig. 3. The ACMLI has three H-Bridge modules and 12 power switching semiconductors, and three dc voltage sources supplied with values of Vdc: 2Vdc: 3Vdc. Vdc of 50V,

100V, and 150V are provided to produce a peak voltage of 300V. Each MOSFET switch is identified based on the bridge location (A, B, or C), as SA1, SA2, SB1, SB2, and so on. In natural sequence ACMLI, N-1 number of MOSFET's are required for an L-level, and thus thirteen-level inverter requires 12 MOSFETs. The output voltage progresses in the levels of 1Vdc, 2Vdc, 3Vdc, 4Vdc... 13 Vdc. The switching sequence of ACMLI is presented in Table I. ACMLI is more susceptible to switch malfunctions as it has more semiconductor switches.

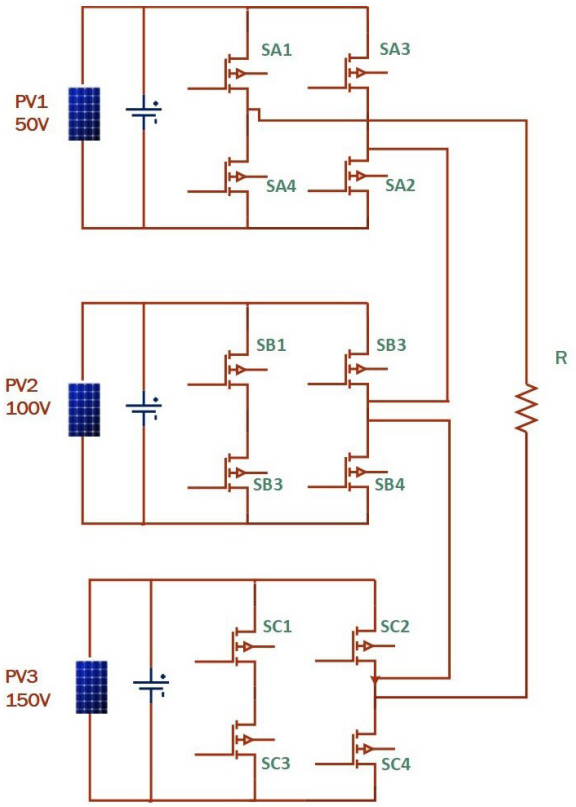


Figure 3: Circuit diagram of ACMLI

Table 1: Switching sequence of ACMLI.

	0	Vdc						-Vdc					
	0	1	2	3	4	5	6	1	2	3	4	5	6
SA1	1	1	1	1	1	1	1	0	0	0	0	0	0
SA2	0	1	0	1	0	1	1	0	1	1	0	1	0
SA3	1	0	1	0	1	0	0	1	0	0	1	0	1
SA4	0	0	0	0	0	0	0	1	1	1	1	1	1
SB1	1	1	1	1	1	1	1	0	0	0	0	0	0
SB2	0	0	1	0	0	1	1	1	0	1	1	0	0
SB3	1	1	0	1	1	0	0	0	1	0	0	1	1
SB4	0	0	0	0	0	0	0	1	1	1	1	1	1
SC1	1	1	1	1	1	1	1	0	0	0	0	0	0
SC2	0	0	0	1	1	1	1	1	1	0	0	0	0
SC3	1	1	1	0	0	0	0	0	0	1	1	1	1
SC4	0	0	0	0	0	0	0	1	1	1	1	1	1

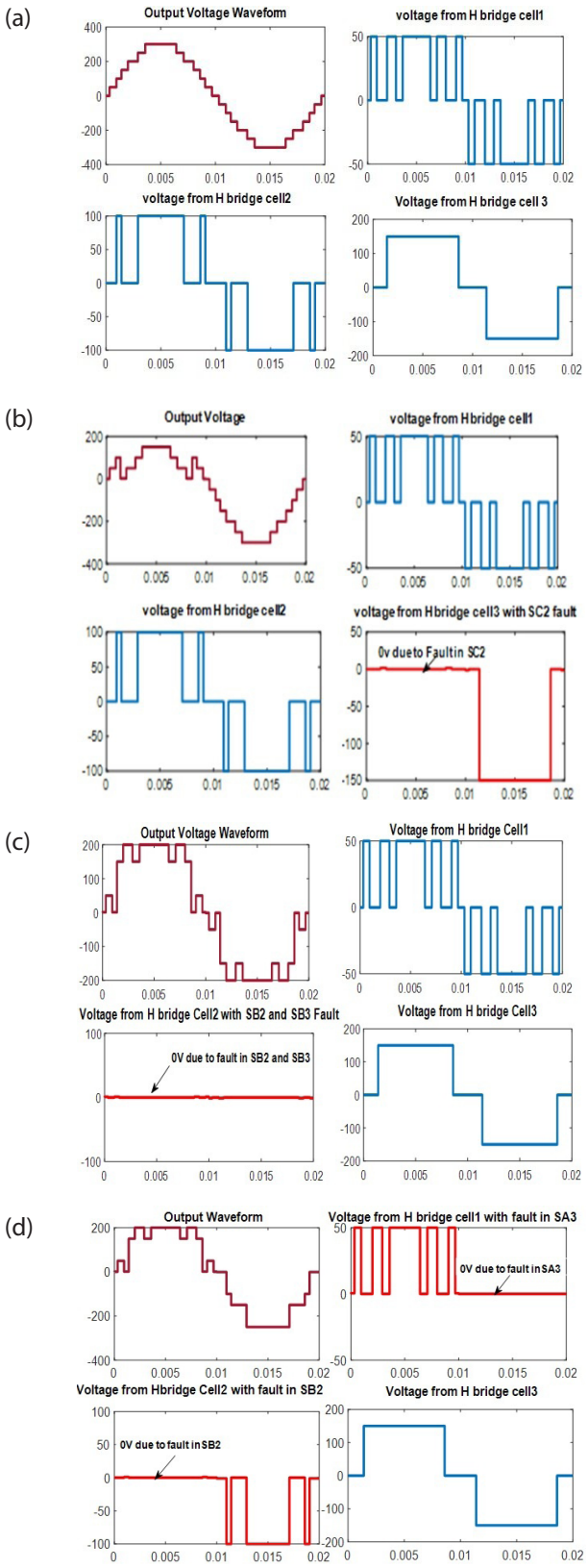


Figure 4: Output and individual H-Bridge voltage waveform during fault conditions in ACMLI: (a) Normal condition (b) Single switch fault (c) Two switch fault in same H-Bridge cell (d) Two switch fault in different H-Bridge cell

4 Fault in ACMLI

Simple fault in ACMLI is single MOSFET malfunctioning at an instant of time, and intricate fault is two or more MOSFET malfunctioning concurrently. This paper deals with simple as well as intricate fault diagnosis. The most difficult aspect of fault detection is obtaining additional information to distinguish between comparable defects in distinct switches.

Twelve single switch faults, sixty-three double switch faults, and three input source faults are examined for validating the proposed method as shown in Table 2.

Fig .4(a) illustrates the output waveform of the ACMLI without any flaws, and distinct waveforms obtained in each H-Bridge cell. In the positive cycle, conduction takes place through switches Sn1 and Sn2 whereas in negative cycle, conduction takes place through Sn3 and Sn4 with n generalized as n=a, b, c.

Case (i) Single switch fault: The fault is located in switch SC2, thus voltage in the positive cycle of H-Bridge cell 3 is zero as shown in fig.4(b). This fault does not permit the source Vdc3 to pervade through it. Hence in H-bridge cell 3, level 3Vdc is lost which causes variation of voltage profile in positive cycle whereas the negative cycle remains unaffected.

Case (ii) Two switch faults in same H-Bridge cell: The fault is located in switches SB2 and SB3 in the same H-Bridge cell 2, thus it affects the flow of current in both positive and negative cycle. As observed in fig. 4(c), this scenario does not allow both half cycles and yields symmetrical variation of output voltage profile in both the cycles.

Case (iii) Two switch faults in different H-Bridge cells: The fault is located in switches of different H-Bridge cells, namely switches SA3 and SB2. This event triggers an asymmetrical variation in output voltage profile as shown in fig. 4(d). The condition (SA3 fault) does not permit voltage Vdc1 to pass through in negative cycle and voltage (SB2 fault) does not permit Vdc2 in positive half cycle.

Table 2: Switching sequence of ACMLI

Type of faults	Labels
Healthy	1
single switch fault	2 to 13
Double switch fault	14 to 76
Input source fault	77 to 79
Total Number of samples	1183

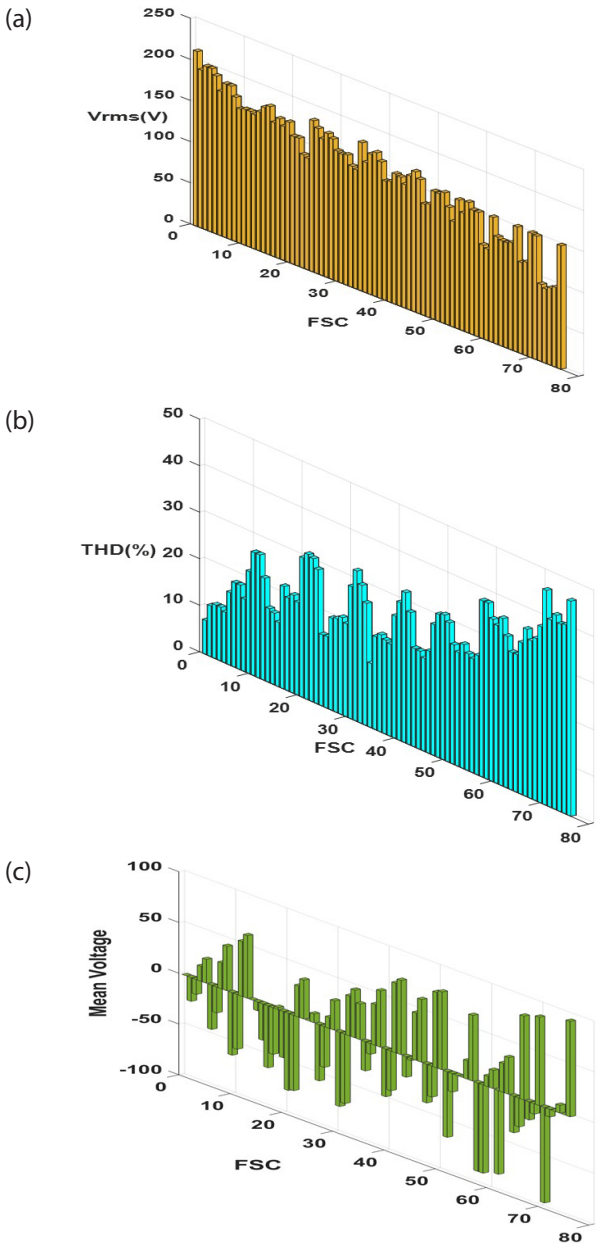
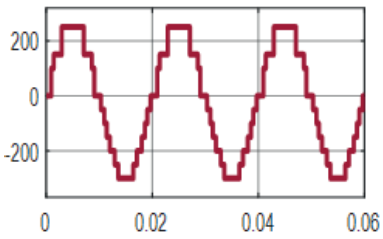


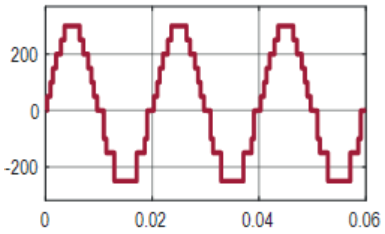
Figure 5: RMS voltage and THD for various Fault switch condition (FSC): (a) RMS voltage (b) THD (%) (c) mean voltage.

The fault case scenarios are labelled as follows: Healthy (1), single switch fault (2 to 13), and double switch fault (14 to 76), and input source fault (77 to 79) as shown in Table 2. The values are decoded into binary form and provided as the label for fault classification. Thus, a total of 1183 samples are provided with 948 for training and 235 for testing data set.

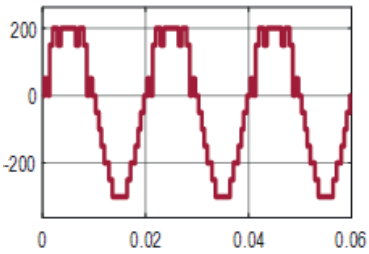
RMS voltage, mean voltage, lower order harmonics and THD are recorded which serve as data set for various single and double OCF fault conditions. From fig.5(a) and fig.5(b) it is observed that certain faults result in similar values of RMS voltage and THD. Thus, it is nec-



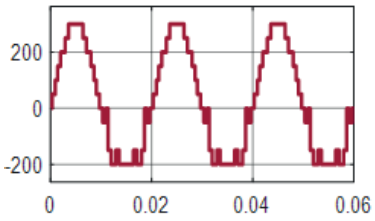
a) SA2



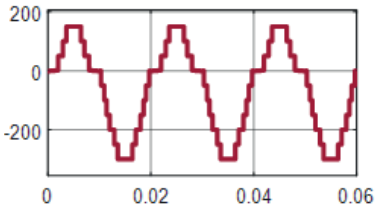
b) SA3



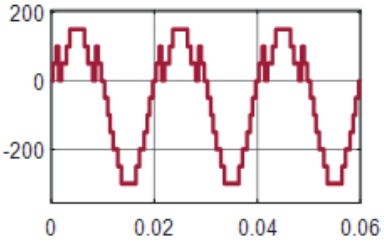
c) SB2



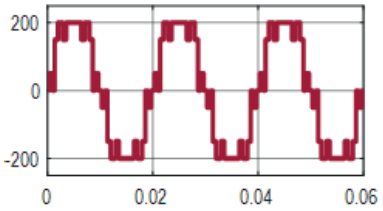
d) SB3



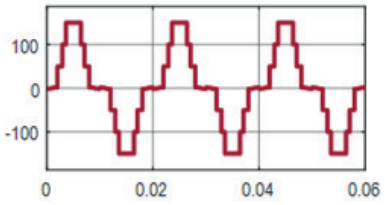
e) SC1



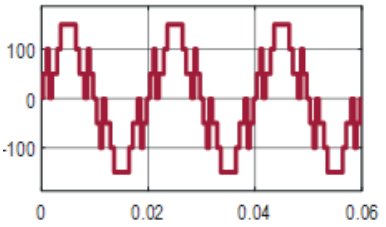
f) SC2



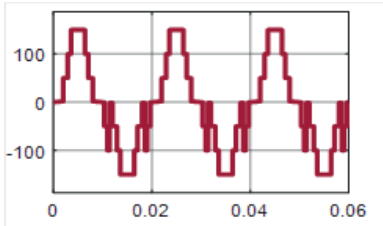
g) SB2 and SB3



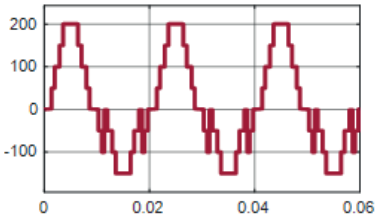
h) SC1 and SC4



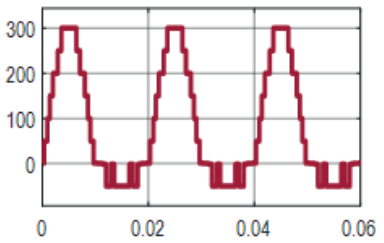
i) SC3 and SC2



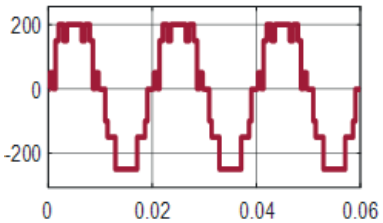
j) SC1 and SC3



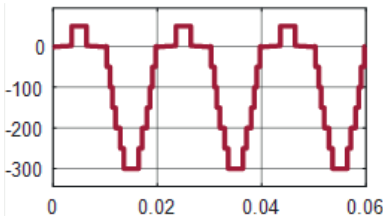
k) SB1 and SB3



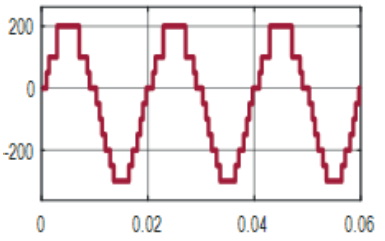
l)SB3 and SC4



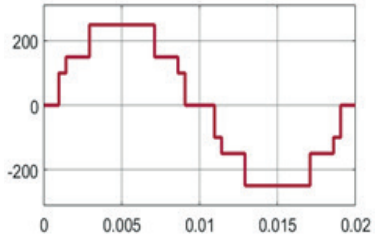
m) SA3 and SB2



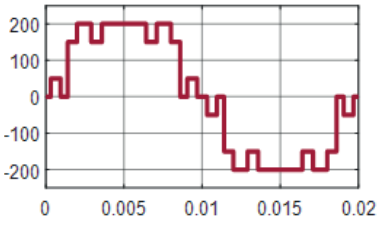
n) SB1 and SC1



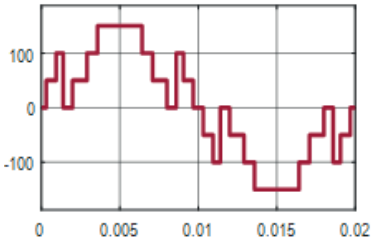
o) SA1 and SA2



p) input source fault 1



q) input source fault2



r) input source fault3

Figure 6: Simulation results of output voltage under various fault switch condition

essary to consider the mean voltage shown in fig.5(c), which shows variation in polarity and amplitude under fault conditions. The mean voltage serves as a rich indicator of type and location of faults.

Simulation results of output voltage waveform for various types of single OCF are presented in fig.6(a) to (f), double switch OCF is given in fig.6(g) to (o), and input source fault is shown in fig. 6(p) to (r). Both even and odd harmonics are considered as the waveform is asymmetrical in most fault cases.

As observed in fig 6(c) and 6(d), the voltage waveforms during SB3 switch fault and SB2 switch fault are complement to each other with similar value of RMS voltage and THD. SB2 switch failure shown in fig. 6(c) causes a faulty voltage waveform with reduction in peak voltage in the positive cycle, with a negative mean voltage of 25.48V. SB3 switch fault causes a similar effect with reduction in voltage in the negative cycle as shown in fig. 6(d) and results in positive mean voltage (+25.48 V.) Hence, the polarity of mean voltage alone varies for SB2 and SB3 fault.

Single switch fault (SA1, SA4), (SA2, SA3), (SB1, SB4), (SB2, SB3), (SC1, SC4), (SC2, SC3) result in similar RMS voltage and THD as shown in fig.5 (a) and (b) with variation in mean voltage as shown in Fig.5(c).

As indicated in fig.5(c) double switch faults in certain scenario causes variation of mean voltage alone. For instance, the double switch (SC1 and SC3), (SC2 and SC4) results in similar value of RMS voltage and THD. However, the mean voltage varies in polarity as (-7.62, +7.62) respectively.

Thus, in some fault scenarios, though Vrms and THD are similar, the mean voltage exhibits a change in voltage polarity or magnitude, which serves as an essential feature for fault identification.

The system identifies the label or switches in which fault had happened. Only from the labels, label 1 is classified as no fault, label 2-13 is classified as single switch fault and labels 14 to 76 is classified as double switch fault. Fault classes 77 to 79 are identified as input source faults.

5 Data Preprocessing

FFT profile (up to 19th harmonics) are taken and PPCA is applied to extract essential information and reduce the dimensionality of data. The PPCA method is used to transform a group of correlated data into a lower-dimensional set. PPCA iteratively refines the model parameters to maximize the likelihood of the actual data. By mapping data onto a lower-dimensional substructure, it retains most of the original data's variation. The aim of PCA is to find a set of principal components that captures the most variance information.

The flow chart of fault diagnosis is represented in fig. 2. The PPCA relates the actual variable with reduced dimension latent variable which induces a corresponding distribution in the data space and given by Eq. (1)

$$h = U_{LM} H + \mu + \varepsilon \tag{1}$$

H is the latent variable principal score of reduced dimension harmonics and ε is multivariate Gaussian error rate, μ the mean vector, ULM is the linear transformation matrix determined by maximum likelihood techniques.

The procedure is described as follows:

i) Input harmonics h_{d*q} and error(ε) where P and R are orthogonal matrices.

ii) Determine the empirical mean of harmonics using Eq. (2) and co-variance using Eq. (3)

$$h_m = \frac{1}{q}\sum_{i=1}^{q} h_i \tag{2}$$

$$A = \frac{1}{\sqrt{q}}\left[\left(h_1 - h_m\right), \left(h_2 - h_m\right) ..... \left(h_q - h_m\right)\right] \tag{3}$$

iii) Decompose A as $A = PDR^T$ where D is the Diagonal matrix

iv) Find Eigen values $\lambda_1, \lambda_{2 \ldots} \lambda_r$ using a diagonal matrix D = diag ($d_1, d_2, \ldots d_r$) r = min (d, b) with d1 ≥ d2 … ≥ dr

v) Arrange eigen values in descending order; λ j ≥ λ j+1
where (1 ≤ I … ≤ d – 1),
$\lambda_i^2 = d_i^2 \left(1 \leq \ i \ldots \leq \min\left(\mathrm{d}, \mathrm{q}\right)\right)$ for j=1 to d-1 do
Pj = Rj,1

vi) Calculate Reconstruction error based on Eq. (4)

$$Q = \frac{1}{q}\sum_{i=1}^{q} \| \ h_i - \ P_j P_i^T \ \left(h_i - h_m\right) - h_m \ \|^2 \tag{4}$$

If Q < ε then b=j; j=d-1

vii) Find variance and orthogonal matrix of maximum likelihood by Eq. (5) and (6) respectively

$$\sigma_{ML}^2 = \frac{1}{d-q}\sum_{i=q+1}^{d} \lambda_{\mathrm{i}} \tag{5}$$

$$P_{ML} = P_q \sqrt{\left(\Lambda_q - \sigma^2 PI_q\right)} \tag{6}$$

The reduced dimensional output matrix of the PPCA is the first five eigen values represented by H = [$\lambda_1, \lambda_{2 \ldots}$

λ_5], ULM is the linear transformation matrix decided by maximum likelihood techniques, μ the mean vector.

6 Machine learning algorithms

6.1 Support vector machine (SVM)

Support Vector Machine (SVM) is primarily employed for regression and classification problems. It functions exceptionally well in circumstances where exacerbated boundaries for decision-making are to be established. The goal of SVM is to split the data points into separate fault categories and optimize the distance between fault classes.

The reduced dimension matrix x_i is x_i = [Vmean, Vrms, THD, H] where H= [λ_1, $\lambda_{2 \ldots}$ λ_5] is obtained by PPCA and yi is the output fault classification [1, 2,…..79] associated with each xi. x_i represents data point variables necessary for fault classification of Fi. Arbitrary hyperplane is characterized as a series of points meeting hyperplane as given in Eq. (7)

$$W^{\circ}x_i - b = 0 \tag{7}$$

where ° is scalar product; W is weight/normal vector perpendicular to plane; b is the bias.

W= (w1, w2 … wp) and x_i = (xi1, xi2 …, xip), the scalar product is represented in Eq. (8)

$$w_{\circ}xi = \sum_{i=1}^{p} w_j x_{ij} \tag{8}$$

The most important part of SVM method is to obtain weight vector (w) and bias value b such that hyperplane equation specified by Eq. (1) maximizes the margin between fault classes. The offset of maximum-margin hyperplane is b / ||W|| and ||W|| denotes the length of vector W. The goal is to find the hyperplane that separates fault classes with minimum error. For a maximum margin hyperplane to compute the support vector on either side the optimization problem is formulated as Eq. (9)

$$\min \frac{1}{2}\left\|\mathrm{W}\right\|^2 + \mathrm{C}\sum_{i=1}^{p}\xi_i y_i (w.\Phi(x_i) + b \geq 1 - \xi_i \tag{9}$$

$$\xi_i \geq 0$$

C is the regularization parameter to avoid overfitting and ξ is the slack variable which measures error made at point (xi, yi,)), ϕ is the function that transforms x_i from p to q dimensions. The harmonic measurement is too close, so Kernel function K is used to denote the proximity between samples xi and xj. The Lagrange multiplier is used to determine the weight vector w and bias b from the solution αi as shown in Eq. (10)

$$max\sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N} y_i\, y_j \alpha_i\, \alpha_j K\left(x_i\ , x_j\right) \tag{10}$$

RBF kernel (where σ > 0 is a user defined parameter) is given by Eq. (11)

$$K\left(x_i, x_j\right) = \exp\ \frac{\| x_i - x_j \|^2}{2\sigma^2} \tag{11}$$

6.2 KNN (K nearest neighborhood)

KNN is an instance-based technique that classifies new data points through groups based on the group which has the greatest number of members among its nearest neighborhoods. KNN establishes prediction based on correlations with adjacent points instead of specifically learning a model. The KNN is based on the Euclidean distance between xi and yi. SVM performance is vulnerable to the regularization parameter (C) and kernel parameters whereas KNN effectiveness depends on the selection of K and the distance metric. The Euclidean distance with n number of training samples is computed by Eq. (12)

$$\sum_{i=1}^{N}\left| x_i - y_i\right|^2 \tag{12}$$

The KNN algorithm is implemented as follows:

- Calculate the distance between x_new and x using the distance metric. Store the distances and their corresponding indices.
- Choose K Neighborhoods based on the calculated distances. Count the occurrences of each class among the K neighbors. Determine the fault class with the highest occurrence rate as the predicted fault class for X_new.

6.3 Decision tree (DT)

Decision Tree continuously splits the feature set into sections to generate a hierarchical tree-like decision-making framework. Every internal node point in a decision tree is based on the characteristics [Vmean, Vrms, THD, H]. To reduce the variance, the most effective characteristic is selected, which divides the data at every node. The process persists until a stopping criterion, namely tree depth is met. The designing parameters of DT are entropy and Gini Index (GI), samples in the leaf (minimum), depth of the tree, and the splitter. En-

tropy and GI are more responsive to node probability changes.

For each feature:

(i) Calculate the impurity of the current node using an entropy using Eq. (13)

$$Entrophy(S) = \sum_{i=1}^{n} P_i log_2 P_i \quad (13)$$

where Pi is the probability of fault classes, S is the case set and n is the number of fault classes.

(ii) Calculate information gain that indicates the reduction in impurity using Eq. (14)
Gain = Impurity Before Split - Weighted Average Impurity after Split

$$Gain = Entrophy(S) - \sum_{i=1}^{n} \frac{S_i}{S} Entrophy(S_i) \quad (14)$$

S-Sample of entire fault input
Si-Sample input of i^{th} fault class

(iii) Select the feature with the highest information gain as the splitting attribute.

(iv) Create a root node with the selected feature.
For each possible value of the selected feature:
Create a child node. Recursively apply the algorithm to the subset of training data corresponding to that value and the remaining features.
Attach the child node to the root node.
Return to the constructed Decision Tree.

(v) Compute information gain for each feature and choose the feature with the highest gain.

DT follows numerous paths if their input information is slightly modified. Reliability and stability are essential in fault identification. During the validation of the proposed fault analysis technique, single and double switch OCF classes are exactly diagnosed, and the faulty MOSFETs are located with an accuracy of 99.32 %.

6.4 Extreme machine learning

To improve the performance of fault diagnosis, a bagged ensemble learning algorithm is employed. Bagging refers to an ensemble technique that generates multiple base models by training each on a bootstrapped sample of the data (randomly selected dataset with replacement). Ensemble Learning with three predictive algorithms namely SVM, KNN and decision tree are employed in this paper. The prediction from each base model is then selected by majority voting techniques, where classification is done based on the fault class with maximum number of votes as shown in fig.7.

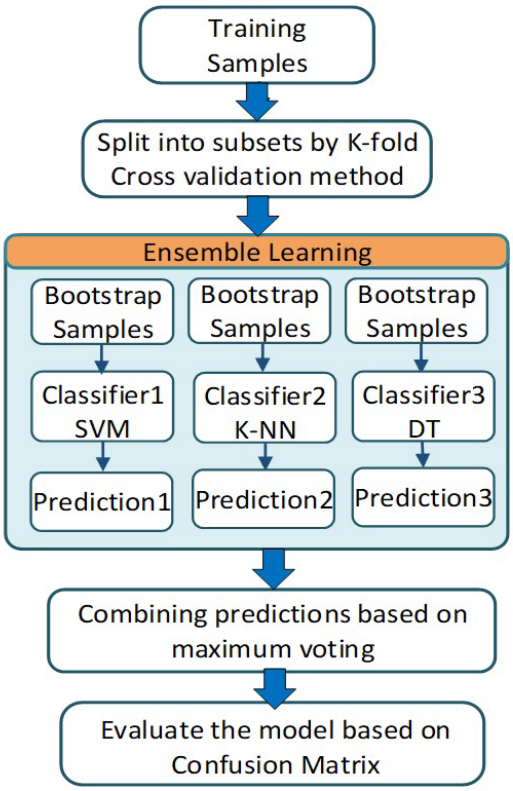


Figure 7: Flow chart of Ensemble Learning

7 Experimental results

The experimental platform of a thirteen-level inverter shown in fig.8. is established with MOSFET. In addition, the FFT spectrum of the output voltage of the ACMLI was computed by DSP TMS320F28034 controller. An open-circuit fault was created by removing the gate signal, and the voltage under faulty condition is measured.

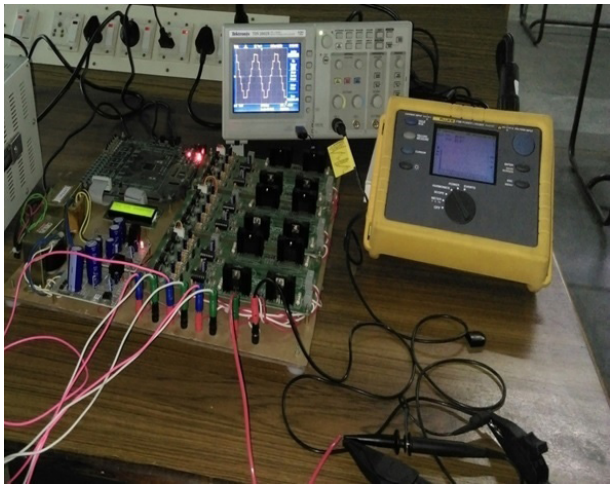


Figure 8: Experimental Model

The voltage signals are then processed using DSP, harmonics and THD are computed by employing FFT. The mean average voltage of ac waveform is computed by sampling the ac voltage over a period and computing the mean value digitally.

The experimental data set is obtained for 1183 data samples, of which 948 are used as training data, and the remaining 235 are taken as testing data.

From Fig. 9(a) –(d), it is evident that noise impact in experimental data is high when fault condition is performed in real-time hardware. Ensemble learning is suitable for this category as it lowers the variance within a noisy dataset and yields higher accuracy, when compared to other ML techniques. In the experimental setup, healthy, OCF, input source faults are examined. The OCF is established by removing the gate signal from the related switch. Many samples of the output voltage are taken for the same fault condition to provide complete insight into the fault classes.

Fig. 9(a) indicates the output voltage waveform for the healthy class without fault. Fig. 9(b) and 9(c) indicate single and double switch faulty instance taken from the experiment. During fault in switches SC3 and SC2 in different legs of the H-Bridge the levels vary multiple times in the same instance as shown in fig 9(b) which is similar to the simulation result shown in fig. 6(i). During simultaneous failure of switches SC1 and SC4 in the same leg, one or more voltage levels are missed as shown in fig.9(c). Fault in the input source of H-bridge 2 results in three-level waveform as shown in fig. 9(d) which is in correlation with the simulation result shown in fig. 6(q).

To eliminate randomness, each method was repeated ten times, and the averaged results are obtained. The parameters given in Eq. (15) – (18) are computed for all the proposed methods. Accuracy provides the correctness of predictions. Precision reveals the accuracy of positive predictions, and recall implies the capability of the model to correctly identify all positive instances. The F1 score is the harmonic mean of precision and recall. Specificity provides models with the ability to correctly identify negative instances.

$$\text{Accuracy}\frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} \tag{15}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} \tag{16}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{17}$$

$$\text{F1} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \tag{18}$$

It is observed from Table 3 that SVM is superior in performance. KNN suffers from the issue of overfitting and thus classifies the fault with a relatively low accuracy of 95.54%. Decision trees when used alone, have the lowest accuracy as it is not benefited by ensemble techniques that use numerous decision trees or ML algorithms to boost accuracy.

Thus, bagging technique EML is applied, and accuracy rate of fault diagnosis is 99.32% with the training set and test set in the ratio of 80%: 20%. The results reveal the fact that EML classifier using PPCA dimensionality reduction achieve fault classification with strong characteristic information and high classification accuracy. In order to validate EML the results are compared with existing literature which is presented in Table 4.

Further online fault detection requires high-performance computation with rapid decision making. Fast diagnosis of faults requires Raspberry Pi with a dual-core Arm Cortex-M0+ processor for training of machine learning algorithms. In real time if any of the H-bridges in the system fail, the faulty H-bridge is bypassed and switching pattern reconfiguration is performed to operate the system with reduced THD and balanced voltage in positive and negative half cycle. Additional H-bridge cells with relay circuits are employed as auxiliary sources to clear the fault in real time. Auxiliary H-bridge cells and chopper circuit that generate dc voltage in a 1:2:3 ratio is employed as redundant circuits. In case of fault, the H-bridge can be disconnected, and auxiliary H-bridge cells and voltage of the chopper can be connected through a relay circuit.

Table 3: Parameters obtained in ML Techniques

	SVM	KNN	DT	ELM
Training Accuracy	97.9	95.54	91.77	99.32
Testing Accuracy	96.58	94.44	88.91	98.74
Precision	97.40	93.02	90.51	99.20
Recall	95.74	95.54	89.78	98.41
F1 Score	96.57	94.26	89.82	98.80
Specificity	97.44	93.44	90.37	99.12

8 Conclusion

This article presents ELM machine learning for single and double switch open circuit fault diagnosis in AC-MLI. Simulation and experimental data sets of many samples are taken, which enhance fault features and

Table 4: Comparison with Existing literatures

ANN Techniques	Topology	Techniques	Number of levels	Accuracy
PCA-BP [22]	Symmetric CMLI	FFT-	5-level	88.7%
RPCA-SVM [15]	Symmetric CMLI	FFT	5-level	92.5%
PPCA-SVM [18]	Symmetric CMLI	FFT	5-level	97.6%
DCNN [19]	Symmetric CMLI	Image processing	5-level	98.16%
Combined Optimizer Fault Classifier [20]	Symmetric CMLI	FFT	5-level	95.56%
AIRMAR [21]	Symmetric (Nested Neutral Point Piloted) MLI	Image processing	5-level	99.33%
Multiscale Kernel CNN [22]	Symmetric CMLI	Short-time Fouri-er transform (STFT)	15-level	98.3%
Mean Voltage Decision trees [23]	Neutral Point Clamped MLI	Pole voltage sign	11-level	98.14%
CSA optimized CNN [24]	Symmetric CMLI	Mean and RMS voltage, THD	9-level	99.84%
BPN [25]	Asymmetric CMLI	Mean and RMS volt-age, THD	7-level and 9-level	99.771%
Proposed Method	Asymmetric CMLI	Mean and RMS volt-age, THD	13-level	98.74%

improve the accuracy of fault diagnosis. This paper identifies that faults in different switches in certain scenarios yield identical Vrms and THD values. Thus, it is necessary to take into consideration the mean voltage and FFT harmonic spectrum from (2nd to 19th). Dimensionality technique PPCA is applied to reduce the dimension of harmonics and trace the most significant features in the dataset.

Machine Learning algorithms namely SVM, KNN and DT are applied individually for fault diagnosis in ACMLI, and their performance is analyzed. ML algorithms SVM,

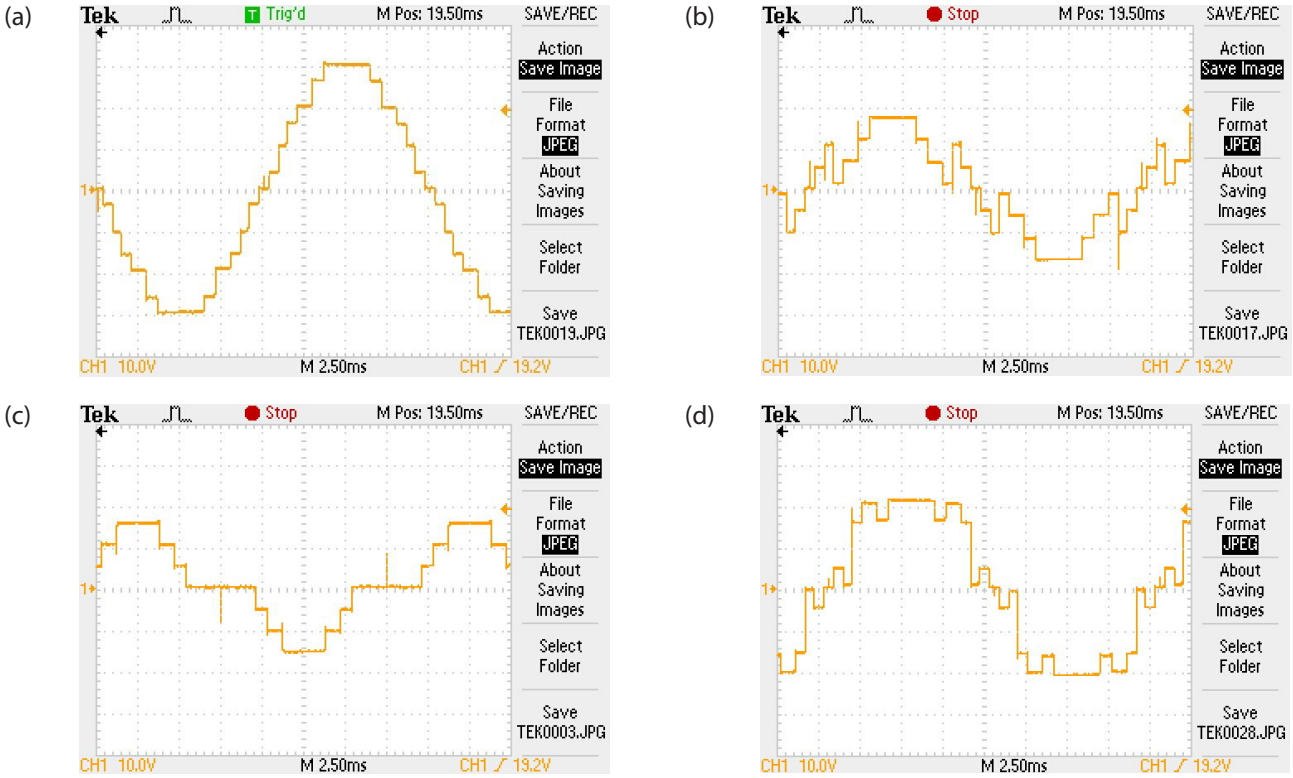


Figure 9: Experimental Results: (a) Normal condition (b) Fault in switch SC3 and SC2 (c) Fault in switch SC1 and SC4 (d) Fault in input source2

KNN and DT are then linked together by a bagging ML approach with maximum voting technique termed as Ensemble Machine Learning (EML). Simulation and experimental investigations are carried out to validate the performance of EML in ACMLI. The experimental results show an accurate categorization rate of 99.32%. ELM provides lower variance within a noisy dataset and thus yields higher accuracy when compared to SVM, KNN and DT algorithms. The study demonstrates that EML classifier with PPCA dimensionality reduction technique achieves the highest classification accuracy in fault diagnosis of ACMLI. The potential limitation of the study is input source fault, and double switch fault in the same leg results in similar kind of fault. To overcome this limitation, additional voltage sensors can be provided to individual input sources to detect source fault. Further research may also be carried out on fault diagnosis based on image processing of the output voltage waveform.

9 References

1. Leon, Jose I., Sergio Vazquez, and Leopoldo G. Franquelo. "Multilevel converters: Control and modulation techniques for their operation and industrial applications." Proceedings of the IEEE 105, no. 11 (2017): 2066-2081. https://doi.org/10.1109/JPROC.2017.2726583
2. B. N. Rao, Y. Suresh, A. K. Panda, B. S. Naik and V. Jammala, "Development of cascaded multilevel inverter based active power filter with reduced transformers," in *CPSS Transactions on Power Electronics and Applications*, vol. 5, no. 2, pp. 147-157, June 2020, https://doi.org/10.24295/CPSSTPEA.2020.00013
3. Babaei, E., Kangarlu, M. F., & Mazgar, F. N. (2012). Symmetric and asymmetric multilevel inverter topologies with reduced switching devices. Electric Power Systems Research, 86, 122-130. https://doi.org/10.1016/j.epsr.2011.12.013
4. A. A. Stonier and B. Lehman, "An Intelligent-Based Fault-Tolerant System for Solar-Fed Cascaded Multilevel Inverters," in *IEEE Transactions on Energy Conversion*, vol. 33, no. 3, pp. 1047-1057, Sept. 2018, doi: 10.1109/TEC.2017.2786299. https://doi.org/10.1109/TEC.2017.2786299
5. Choi, U. M., Blaabjerg, F., & Lee, K. B. (2014). Study and handling methods of power IGBT module failures in power electronic converter systems. IEEE Transactions on Power Electronics, 30(5), 2517-2533. https://doi.org/10.1109/TPEL.2014.2373390
6. S. Khomfoi, L. Tolbert, Fault diagnosis system for a multilevel inverter using a neural network, IEEE Trans. Power Electron. 22 (2007) 1062–1069, https://doi.org/10.1109/TPEL.2007.897128
7. Jung, S. M., Park, J. S., Kim, H. W., Cho, K. Y., & Youn, M. J. (2012). An MRAS-based diagnosis of open-circuit fault in PWM voltage-source inverters for PM synchronous motor drive systems. IEEE Transactions on Power Electronics, 28(5), 2514-2526. https://doi.org/10.1109/TPEL.2012.2212916
8. Mellit, Adel, Giuseppe Marco Tina, and Soteris A. Kalogirou. "Fault detection and diagnosis methods for photovoltaic systems: A review." Renewable and Sustainable Energy Reviews 91 (2018): 1-17. https://doi.org/10.1016/j.rser.2018.03.062
9. Reyes-Malanche, J. A., Villalobos-Pina, F. J., Cabal -Yepez, E., Alvarez-Salas, R., & Rodriguez-Donate, C. (2021). Open-Circuit Fault Diagnosis in Power Inverters Through Currents Analysis in Time Domain. IEEE Transactions on Instrumentation and Measurement, 70, 1-12. https://doi.org/10.1109/TIM.2021.3082325
10. Zhang, Weiwei, and Yigang He. "A hypothesis method for T-type three-level inverters open-circuit fault diagnosis based on output phase voltage model." IEEE Transactions on Power Electronics (2022). Power Electron., vol. 30, no. 12, pp. 7006–7018, Dec. 2015 https://doi.org/10.1109/TPEL.2022.3151731
11. B. Cai, Y. Zhao, H. Liu, and M. Xie, "A data-driven fault diagnosis methodology in three-phase inverters for PMSM drive systems," IEEE Trans. Power Electron., vol. 32, no. 7, pp. 5590–5600, Jul. 2017. https://doi.org/10.1109/TPEL.2016.2608842
12. Xia, Y., & Xu, Y. (2021). A transferrable data-driven method for IGBT open-circuit fault diagnosis in three-phase inverters. IEEE Transactions on Power Electronics, 36(12), 13478-13488. https://doi.org/10.1109/TPEL.2021.3088889
13. Zheng, H., Wang, R., Wang, Y., & Zhu, W. (2017, July). Fault diagnosis of photovoltaic inverters using hidden Markov model. In 2017 36th Chinese Control Conference (CCC) (pp. 7290-7295). IEEE. https://doi.org/10.23919/ChiCC.2017.8028508
14. Wu, Xun, Chun-Yang Chen, Te-Fang Chen, Shu Cheng, Zhi-Hong Mao, Tian-Jian Yu, and Kaidi Li. "A fast and robust diagnostic method for multiple open-circuit faults of voltage-source inverters through line voltage magnitudes analysis." IEEE Transactions on Power Electronics 35, no. 5 (2019): 5205-5220. https://doi.org/10.1109/TPEL.2019.2941480
15. T.Z. Wang, J. Qi, H. Xu, Y.D. Wang, L. Liu, D.J. Gao, Fault diagnosis method based on FFT-RPCA-SVM for Cascaded-Multilevel Inverter, ISA Trans. 60 (2016) 156–163, https://doi.org/10.1016/j.isatra.2015.11.018

16. Wang, Tianzhen, Hao Xu, Jingang Han, Elhoussin Elbouchikhi, and Mohamed El Hachemi Benbouzid. "Cascaded H-bridge multilevel inverter system fault diagnosis using a PCA and multiclass relevance vector machine approach." IEEE Transactions on Power Electronics 30, no. 12 (2015): 7006-7018. https://doi.org/10.1109/TPEL.2015.2393373
17. Yuan, Q., Tu, Q., Yan, L., & Xia, K. (2023). Fault diagnosis of H-bridge cascaded five-level inverter based on improved support vector machine with gray wolf algorithm. Energy Reports, 9, 485-495. https://doi.org/10.1016/j.egyr.2023.03.017
18. Ali, Murad, Zakiud Din, Evgeny Solomin, Khalid Mehmood Cheema, Ahmad H. Milyani, and Zhiyuan Che. "Open switch fault diagnosis of cascade H-bridge multi-level inverter in distributed power generators by machine learning algorithms." Energy Reports 7 (2021): 8929-8942. https://doi.org/10.1016/j.egyr.2021.11.058
19. Du, B., He, Y., & Zhang, C. (2021). Intelligent diagnosis of cascaded H-bridge multilevel inverter combining sparse representation and deep convolutional neural networks. IET Power Electronics, 14(6), 1121-1137. https://doi.org/10.1049/pel2.12094
20. Sudha, V., K. Vijayarekha, Rakesh Kumar Sidharthan, and Natarajan Prabaharan. "Combined Optimizer for Automatic Design of Machine Learning-Based Fault Classifier for Multilevel Inverters." IEEE Access 10 (2022): 121096-121108. https://doi.org/10.1109/ACCESS.2022.3193784
21. Ye, S., Zhang, F., Gao, F., Zhou, Z., & Yang, Y. (2022). Fault diagnosis for multilevel converters based on an affine-invariant riemannian metric autoencoder. IEEE Transactions on Industrial Informatics, 19(3), 2619-2628. https://doi.org/10.1109/TII.2022.3186992
22. A. Sivapriya, N. Kalaiarasi, R. Verma, B. Chokkalingam and J. L. Munda, "Fault Diagnosis of Cascaded Multilevel Inverter Using Multiscale Kernel Convolutional Neural Network," in IEEE Access, vol. 11, pp. 79513-79530, 2023, doi: 10.1109/ACCESS.2023.3299852. https://doi.org/10.1109/ACCESS.2023.3299852
23. Rinsha, V., & Jagadanand, G. (2023). Rolling Average-Decision Tree-Based Fault Detection of Neutral Point Clamped Inverters. IEEE Journal of Emerging and Selected Topics in Industrial Electronics, vol. 4, no. 3, pp. 744-755. https://doi.org/10.1109/JESTIE.2023.3236587
24. Sivapriya, A., & Kalaiarasi, N. (2023). A Novel Enhanced Deep Learning-Based Fault Diagnosis Approach for Cascaded Multilevel Inverter. e-Prime-Advances in Electrical Engineering, Electronics and Energy, 100253. https://doi.org/10.1016/j.prime.2023.100253
25. Raj, N., Jagadanand, G., & George, S. (2018). Fault detection and diagnosis in asymmetric multilevel inverter using artificial neural network. International Journal of Electronics, 105(4), 559-571. https://doi.org/10.1080/00207217.2017.1378382

Arrived: 12. 10. 2023
Accepted: 10. 03. 2024