Abstract—In this paper, a fault diagnostic system in a multilevel-inverter drive (MLID) system using a neural network is developed. It is difficult to diagnose a multilevel-inverter drive (MLID) system using a mathematical model because MLID systems consist of many switching devices and their system complexity has a nonlinear factor. Therefore, a neural network classification is applied to the fault diagnosis of a MLID system. Five multilayer perceptron (MLP) networks are used to identify the type and location of occurring faults from inverter output voltage measurement. The neural network design process is clearly described. The classification performance of the proposed network between normal and abnormal condition is about 90%, and the classification performance among fault features is about 85%. Thus, by utilizing the proposed neural network fault diagnostic system, a better understanding about fault behaviors, diagnostics, and detections of a multilevel inverter drive system can be accomplished. The results of this analysis are identified in percentage tabular form of faults and switch locations.

Index Terms—Diagnostic system, fault diagnosis, multilevel inverter drive (MLID), neural network.

I. INTRODUCTION

In recent years, industry has begun to demand higher power ratings, and multilevel inverter drive (MLID) systems have become a solution for high power applications. A multilevel inverter not only achieves high power ratings, but also enables the use of renewable energy sources. Two topologies of multilevel inverters for electric drive application have been discussed in [1]. The cascade MLID is a general fit for large automotive all-electric drives because of the high VA rating possible and because it uses several level dc voltage sources which would be available from batteries or fuel cells [1].

A schematic of a single phase multilevel inverter system is illustrated in Fig. 1. Because multilevel inverter systems are utilized in high power applications, the reliability of the power electronics equipment is very important. For example, industrial applications such as industrial manufacturing are dependent upon induction motors and their inverter systems for process control. Generally, the conventional protection systems are passive devices such as fuses, overload relays, and circuit breakers to protect the inverter systems and the induction motors. The protection devices will disconnect the power sources from the multilevel inverter system whenever a fault occurs, stopping the operated process. Downtime of manufacturing equipment can add up to be thousands or hundreds of thousands of dollars per hour, therefore fault detection and diagnosis is vital to a company’s bottom line.

Many engineers and researchers have focused on incipient fault detection and preventive maintenance to avert inverter and motor faults. The most important factor, however, is how the system could operate continuously while there is an abnormal condition.

In order to maintain continuous operation for a multilevel inverter system, knowledge of fault behaviors, fault prediction, and fault diagnosis are necessary. Faults should be detected as soon as possible after they occur, because if a motor drive runs continuously under abnormal conditions, the drive or motor may quickly fail.

Research on fault diagnostic techniques initially focused on conventional pulsedwidth modulation (PWM) voltage source inverters (VSI). The various fault modes of a VSI system for an induction motor are investigated in [2]. Then, the integration of a fault diagnostic system into VSI drives is described in [3]. This integrated system introduced remedial control strategies soon after failure occurrences; therefore, system reliability and fault tolerant capability are improved. A noninvasive technique for diagnosing VSI drive failures based on the identification of unique signature patterns corresponding to the motor supply current Park’s Vector is proposed in [4]–[6]. Input motor currents can be reconstructed by measuring a dc link current; so, only one sensor is needed to represent all input motor currents. A new topology for a low-cost VSI where true phase current information exists with the use of only one dc link current sensor has been proposed by [7]. A comparison of features, cost, and
limitations of fault-tolerant three-phase ac motor drive topologies is investigated in [8].

It is possible that artificial intelligence (AI)-based techniques can be applied in condition monitoring and diagnosis. AI-based condition monitoring and diagnosis have several advantages. For instance, AI-based techniques do not require any mathematical models; therefore, the engineering and development time could be significantly reduced. AI-based techniques utilize the data sets of the system or expert knowledge [9]. A general review of recent developments in the field of AI-based diagnostic systems in machine drives has been proposed in [10]. A study of a machine fault diagnostic system by using fast Fourier transform (FFT) and neural networks is clearly explained in [11]. Also, a fault diagnostic system for rotary machines based on fuzzy neural networks is developed in [12]. The possibilities offered by a neural network for fault diagnosis and system identification are investigated in [13]. Furthermore, a new topology with fault-tolerant ability that improves the reliability of multilevel converters is proposed in [14].

Fault diagnosis and neutral point voltage control during a fault condition in a three-level diode-clamped multilevel inverter using Park’s Vector has been proposed in [15]. A method for operating cascaded multilevel inverters when one or more power H-bridge cells are damaged has been proposed in [16]. The method is based on the use of additional magnetic contactors in each power H-bridge cell to bypass the faulty cell and use the neutral shift technique to maintain balanced line-to-line voltages. Thus far, limited research has focused on MLID fault diagnosis. Therefore, a MLID fault diagnostic system is proposed in this paper that only requires measurement of the MLID’s voltage waveforms.

An example of a MLID open circuit fault at switch $S_{A+}$ is represented in Fig. 2. $S_{A+}$ fault will cause unbalanced voltage and current output, while the induction motor is operating. The unbalanced voltage and current may result in vital damage because of overheating to the induction motor if the induction motor is supplied with unbalanced voltages for a long time. The unbalanced condition from fault $S_{A+}$ can be corrected if the fault location is identified. Switching patterns and the modulation index of other active switches in the MLID can be adjusted to maintain output voltage and current in a balanced condition. Although the MLID can continuously operate in a balanced condition, the MLID will not be able to operate at its rated power.

Therefore, the MLID can operate in a balanced condition at reduced power while the fault occurs until the operator knows and repairs the inactive switch.

In this research, we will attempt to diagnose the fault location in a MLID from its output voltage waveform because the output voltages are normally independent from the load and correspond with fault types and locations. MLID open circuit faults at each switch are considered, and for illustrative purposes, one phase of a multilevel inverter is the focus of this paper. Although the MLID system usually consists of three phases of H-bridge inverters and can also have short circuit faults, the fault diagnostic system will be the same topology as a single phase and open circuit case. Moreover, other inverter levels can be extended by using this proposed topology with more training data. It should be noted that added training data could lead to a more neurons in the existing hidden layers or could lead to additional hidden layers which would slow the calculation. The proposed network utilizes output voltage signals of the MLID to train the neural networks. The acquired data is transformed by using FFT technique to rate a signal value as an important characteristic. The signal feature extraction is discussed, and the process of neural network design is clearly described in the following section.

II. GENERAL NOTION OF FAULT DIAGNOSTIC SYSTEM

A. Structure of Fault Diagnostic System

The structure for a fault diagnostic system is illustrated in Fig. 3. The system is composed of four major states: feature extraction, neural network classification, fault diagnosis, and switching pattern calculation with gate signal output. The feature extraction, neural classification, and fault diagnosis are the focus of this research. The feature extraction performs the voltage input signal transformation, with rated signal values as important features, and the output of the transformed signal is transferred to the neural network classification. The networks are trained with both normal and abnormal data for the MLID;
thus, the output of this network is nearly 0 and 1 as binary code. The binary code is sent to the fault diagnosis to decode the fault type and its location. Then, the switching pattern is calculated.

### B. Feature Extraction System

The Simpower Matlab toolbox in Simulink is used to simulate data of fault features with 0.8 modulation index (\(m_{\text{dq}}\)) out of 1.0 as illustrated in Fig. 4. Also, output voltages are shown for an MLID with open circuit faults and short circuit faults in Fig. 5. One can see that all fault features in both open circuit and short circuit cases could be visually distinguished; however, the computation unit cannot directly visualize as a human does. Also, the signals in Fig. 5 are difficult to rate as an important characteristic and have high correlation coefficient for classifying a fault hypothesis. Therefore, a signal transformation technique is required. An appropriate selection of the feature extractor is to provide the neural network with adequate significant details in the pattern set so that the highest degree of accuracy in the neural network performance can be obtained. One possible technique for implementation with a digital signal processing microchip is FFT [18]. Beginning with the discrete Fourier transform (DFT) in (1), and then the FFT using the decimation in time decomposition algorithm is illustrated in (2). Together, the computational savings of the FFT becomes \(O(N \log_2 N)\) compared to quadratic time \(O(N^2)\) for the DFT. This means that if \(N\) is 16, the FFT will execute only 64 times, whereas the DFT will run 256 times. Other popular signal transformation techniques such as Hartley and Wavelet are explained in [18].

\[
F_k = \sum_{n=0}^{N-1} f_n W_N^{nk} \quad \text{for } k = 0, \ldots, N-1 \quad (1)
\]

where

\[
W_N = e^{-j2\pi/N},
\]

\[
N \quad \text{number of harmonic orders};
\]

\[
F_k = G_k + W_N^k H_k \quad \text{for } k = 0, \ldots, \frac{N}{2} - 1
\]

\[
F_{k+(N/2)} = G_k - W_N^k H_k \quad \text{for } k = 0, \ldots, \frac{N}{2} - 1. \quad (2)
\]

\(G_k\) is for even-numbered elements of \(f_n\), whereas \(H_k\) is for odd-numbered elements of \(f_n\). \(G_k\) and \(H_k\) can be computed as shown in

\[
G_k = \sum_{n=0}^{(N/2)-1} f_{2n} W_N^{nk} \quad (3)
\]

\[
H_k = \sum_{n=0}^{(N/2)-1} f_{2n+1} W_N^{nk} \quad (4)
\]

### C. Simulation and Experiment Comparison

The transformed signals of both simulation and experiment are represented in Fig. 6(a) and (b), respectively. Obviously, the results are nearly identical fault features. The FFT technique has a good identical feature to classify normal and abnormal features; therefore, FFT is used to transform voltage output signals in this research in order to rate signal value for important features so that the features for a fault hypothesis can be classified (see Fig. 7).

A three-phase, wye-connected cascaded multilevel inverter using 100-V, 70-A MOSFETs as the switching devices was used in the experiment. An Opal-RT Lab system was utilized to generate gate drive signals and to interface with the gate drive board. The switching angles are calculated by using Simulink based on multilevel sinusoidal PWM as shown in Fig. 8. A 12-V power
supply is supplied to each H-Bridge inverter in both simulation and experiment. It should be noted that the two separate dc sources are used in this experiment set up in each of three phases. Fault occurrence is created by physically removing the switch in the desired position. A Yokogawa DL 1540c is used to measure output voltage signals shown in Fig. 9 as ASCII files. The measured signals are set to $N = 10032$; sampling frequency is 200 kHz. Voltage spectrum is calculated and transferred to the neural network fault classification.

Obviously, the output voltage signals in Fig. 9 could also be used to classify the faults at different H-bridges. The experimental results suggest that if one could visualize the different fault features, the neural network could also perform the fault classification. Additionally, a classification technique using a neural network offers an extra degree of freedom to solve a non-linear problem: the failure of single neuron will only partially degrade performance. If an input neuron fails, the network can still make a decision by using the remaining neurons. In contrast, if only a single input, for instance the dc offset of signals is used as the input data to classify the faults, the diagnostic system could not perform classification when the input data has drifted or the single sensor has failed. Furthermore, a neural network also permits parallel configuration and seasonal changes. Additional H-bridges and fault features (short circuit) could be conveniently extended into the system with more training data and parallel configuration.

D. Reconfiguration Method

The basic principal of the reconfiguration method is to bypass the faulty cell (H-bridge); then, other cells in the MLID are used to compensate for the faulty cell. For instance, if cell 2 of MLID in Fig. 2 has an open circuit fault at $S_{A+}$; accordingly, $S_{A-}$ and $S_{B-}$ need to be turned on (1), whereas $S_{B+}$ needs to be turned off (0) to bypass cell 2. The corrective actions taken for other fault locations are shown in Table I. As can be seen, the corrective action would be the same for cases that have similar voltage waveforms during their faulted mode (for instance, see Fig. 5 for a short circuit fault in $S_{A+}$ and open circuit fault in $S_{A-}$). Therefore, even if the fault may be misclassified (an actual short circuit fault at $S_{A+}$ is misclassified as an open circuit fault at $S_{A-}$ or vice versa), the corrective action taken would still solve the problem.

III. METHODOLOGY OF NEURAL NETWORK FAULT CLASSIFICATION

All fault features, as previously discussed, can be classified based upon their effects upon the output voltages. The transformation of output voltage signals is achieved by using FFT as shown by simulation and experimental results in Figs. 6 and 9. As mentioned before, the systematic mathematical technique
Fig. 7. Experiment set-up.

Fig. 8. Multilevel carrier-based sinusoidal PWM showing carrier bands, modulation waveform, and inverter output waveform \((m_o = 0.8/1.0)\).

may be complicated to implement in a practical real time control system; therefore, a feedforward neural network technique permitting input/output mapping with a nonlinear relationship between nodes will be utilized [9], [17]. Neural networks provide the ability to recognize anomalous situations because of their intrinsic capacity to classify and generalize. Especially, the sensitivity and response time of the original procedure presented for the on-line analysis of fault set repetition enable on-line fault location techniques to be developed [13]. The stages of neural network fault classification are explained as follows:

A. Neural Network Architecture Design

The architecture of the proposed fault diagnostic neural network, NN, is illustrated in Fig. 10. The five multilayer feedforward networks, or multilayer perceptron (MLP), are used in this research because the input data contain continuous features. A network has one hidden layer with 40 input nodes corresponding to harmonic order and magnitude, 2 hidden nodes, and 1 output node. The sigmoid activation function is used: \(tansig\) for hidden nodes and \(logsig\) for an output node. A \(logsig\) activation function is used for an output node because the target output is between 0 and 1. The implementation of the proposed neural network classification system, consisting of five NNs is shown in Fig. 11. It should be noted that the number of nodes for the input and output layers depends on the specific application. The selection of number and dimension in the hidden layer is based on neural network accuracy in preliminary tests. Indeed, optimization of the network architecture is a significant topic in a study of artificial intelligence aspects [17].

B. Input/Output Data

Each network is trained with one set of normal data and four sets of abnormal data, thus the size of the input matrix is five
input data rows with 40 columns, \([5 \times 40]\). The size of the output target is \([5 \times 1]\). The target output corresponding with classification data is represented in Table II. A selection of training data and test data, from the original signal, is discussed in the next section.

C. Neural Network Training

The Levenberg Marquardt training paradigm, \textit{trainlm} [19] is utilized in this research because \textit{trainlm} not only performs very fast training time but also has inherent regularization properties. Regularization is a technique which adds constraints so that the results are more consistent. The 1\% misclassification and 1\% input data error rate are chosen to calculate a sum of square error goal, SSE; therefore, a 0.0334 SSE goal is used to train the network by calculating from (5). The training process will be finished when the SSE goal is met

\[
SSE = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{L} (y_{ij} - d_{ij})^2
\]  

(5)

where

- \(y_{ij}\) output value of neural network;
- \(d_{ij}\) output of training data;

\(N\) number of training data;
\(L\) number of units in the output layer.

1) Training and Testing Data Set Selection: The training data set should also cover the operating region; thus, the test data sets are generated from simulation with various operation points (different modulation indices). The testing sets are measured from experiment. Training and testing sets have 200 kHz sampling frequency. Both data sets are transformed by FFT to a set of 0 to 39 harmonic orders. Zero harmonic order means the dc component of the signals.

2) Scaling Data: The input training data are scaled by using the mean center and unit variance method (Z-score scaling). The scaling data will avoid premature saturation of sigmoidal units and also allow the use of a specific output neuron [17], [19]. The scaling parameters: the mean value \((X_M)\) and the standard
deviation value ($X_S$) are saved with the same data file as weights and biases. The testing data set will be scaled with the same scaling parameters as the training data set when the network is examined.

D. Neural Network Testing

The networks are examined with the test data sets, as mentioned above, when the proposed networks have trained to the desired error goal. Testing the network involves presenting the test set to the network and calculating the error. If the error goal is met, the training is complete.

IV. FAULT CLASSIFICATION RESULTS

The performance of the proposed networks is tested in two categories. First, the networks are tested with the selected test set at the same operating point ($m_{\text{dc}} = 0.8$) as previously mentioned. The tested results along with the testing data set are illustrated in Table III. As can be seen, the error between the actual and target output data after testing the network is less than the SSE goal (0.0334), thus the training process is complete. Clearly, each network has classification performance of more than 98%; therefore, the classification performance of the network is quite satisfactory when the test data is at operating conditions close to the training data.

Second, the new testing data are simulated at different operating points by changing the modulation index ($m_{\text{dc}}$) to 1.0. The new testing data set, at a unity modulation index, is transferred to the neural networks. In the MLID application, the modulation index is dependent on the operation point of the induction motor. Therefore, it would be better if the networks have good classification performance for a wide range of operation. The second category testing results are illustrated in Table IV. Obviously, the classification performance between normal and abnormal conditions is about 90%, which is very good performance, whereas the classification performance among fault features is about 85% in Fault $A-$ and $B-$ and about 75% in Fault $A+$ and $B+$ (The proposed network misclassified only one out of four test sets). It should be noted that the future classification performance is important for the application. The classification performance among fault occurrences is acceptable. The neural network can be trained with a larger, more complete fault data set to get more accurate results when classifying actual faults.

As shown in Table IV, the classification performance decreases when the operation point of MLID is changed. This suggests that an additional training set, or more training data, may be needed to train the fault classification neural network, especially in Fault $A+$ and Fault $B+$ training set. Also, a better data transformation technique may be required; however, the time-consuming process used by a selective algorithm such as the DFT, Hartley, or Wavelet in feature extraction should be considered when implementing it in an on-line fault diagnostic system.

V. CONCLUSION

A study of a fault diagnostic system in a multilevel inverter using neural networks has been proposed, and the feature extraction system has been discussed. An FFT technique is utilized to transform output voltage signals in order to rate the signal value as an important characteristic for classifying a fault hypothesis. FFT has an advantage in that less time is needed to calculate this signal transformation process, and it is possible to implement with a digital signal processing microchip. The feature extraction system is an important process because bad transformation signals will lead to poor classification performance. A study suggests that the feature extraction is a challenging research topic.

The proposed networks perform well with the selected testing data set. The classification performance is high, more than 98%. Obviously, the classification performance between normal and abnormal condition is quite satisfactory in both testing categories; whereas, the classification performance among fault features is acceptable even when tested at modulation indices other than the training data set. The classification performance decreases when the operating point of the MLID is different from the training set. The results indicate that a new training set, or more training data, may be needed to accomplish a wide range of operation.

Although the classification performance decreases when the operating point is changed, the overall classification performance of the proposed fault diagnostic system is acceptable. The proposed networks have the ability to classify normal and abnormal conditions, including the fault location. Additional H-bridges and fault features (short circuit) could be conveniently extended into the system with more training data and parallel neural network configuration. Therefore, by utilizing
the proposed neural network fault diagnostic system in this research, a better understanding of fault behaviors, diagnostics, and detections for multilevel inverter drive systems can be achieved.

REFERENCES