Neural Networks Detect Inter-Turn Short Circuit Faults Using Inverter Switching Statistics

Mustafa Umit Oner, İlker Şahin and Ozan Keysan

***Abstract*—Early detection of an inter-turn short circuit fault (ISCF) can reduce repair costs and downtime of an electrical machine. In an induction machine (IM) driven by an inverter with a model predictive control (MPC) algorithm, the controller outputs are influenced by a fault due to the fault-controller interaction. Based on this observation, this study developed neural network models using inverter switching statistics to detect the ISCF of an IM. The method was non-invasive, and it did not require any additional sensors. In the fault detection task, an area under receiver operating characteristics curve value of 0.9950 (95% Confidence IntervaI: 0.9949 - 0.9951) was obtained. At the rated operating conditions, the neural network model detected and located an ISCF of 2-turns (out of 104 turns per phase) under 0.1 seconds, a speedup of more than two times compared to the thresholding-based method. Moreover, we published the switching vector data collected at various load torque and shaft speed values for healthy and faulty states of the IM, becoming the first publicly available ISCF detection dataset. Together with the dataset, we provided performance baselines for three main neural network architectures, namely, multi-layer perceptron, convolutional neural network, and recurrent neural network.**

***Index Terms*—Condition monitoring, fault diagnosis, induction motor, machine learning, motor drives, multi-layer perceptron, neural networks, predictive control.**

I. Introduction

OWING to its value and significance, fault diagnosis of electrical machines has been a focus of intensive research, as reflected by a plethora of publications over the past years [1]–[4]. The early detection of an incipient fault can enable repair cost and downtime reduction benefits. Furthermore, provided that the machine is fault-tolerant by design and proper identification of the inflicting fault is made, the continuum of operation with a reduced rating is also possible.

Several fault detection methods that address induction motors (IM) have been reported [5]–[7] as the IM is the most commonly used AC machine type due to its low cost and ruggedness. It is estimated that the stator faults constitute 21% of all the faults [8]. Stator faults usually start as inter-turn short circuit faults (ISCF) [9] and quickly develop further into complete phase-to-phase or phase-to-ground faults, which implies the total malfunctioning of the machine. Depending on the machine and the fault's structure, the time between ISCF occurrence and the total loss of insulation is in the order of seconds [10]. Therefore, a swift and effective identification of an incipient ISCF is crucial.

An important distinction regarding the fault detection studies is the control method assumed for the motor. The motor can be line-fed (uncontrolled, open-loop) or closed-loop controlled via an inverter. There exists a complex interaction between the

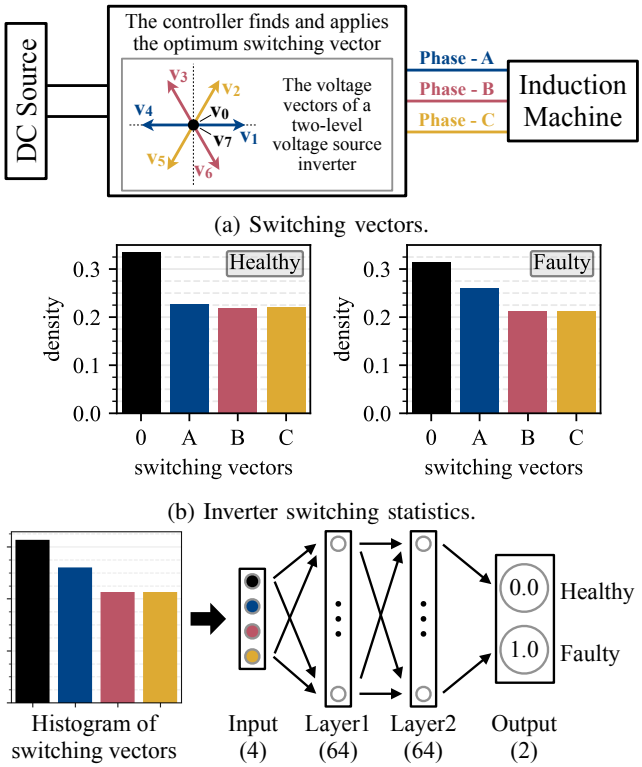


Fig. 1: **Inverter switching statistics and neural network model. (a)** The voltage vectors of a two-level voltage source inverter. The controller finds the optimum voltage vector in view of the control outcomes, and applies it at the next switching instant. **(b)** Histograms of switching vectors over a period for a healthy machine and a machine with inter-turn short circuit fault are given. While aggregated 0-vectors is represented as 0, aggregated active vectors are represented as A, B and C. **(c)** The neural network model is a multi-layer perceptron consisting of an input layer (with 4 nodes), two hidden layers (Layer1 and Layer2 - each with 64 nodes) and an output layer (with 2 nodes). The model takes a histogram of switching vectors at the input and predicts whether the machine is healthy or faulty at the output.

fault and the controller [11], [12]. The controller inherently tries to negate the fault's effect. It is shown in [13] that an IM drive implemented with model predictive control (MPC) continues to exhibit perfectly balanced phase currents under an ISCF of 3-turns (out of 104 turns per phase). However, a significant unbalance is observed for the line-fed operation under the same fault condition. This example implies that most fault detection methods developed considering line-fed machines (such as motor current signature analysis) would be less effective (if not totally useless) for a high-performance control case. Therefore, it is essential to develop a fault detection method in conjunction with the main control algorithm.

Recently, the utilization of artificial intelligence (AI) techniques, such as neural networks (NN), has been gaining increasing momentum in power electronics [14]. A particular area for which the NN approach is very suitable is the fault diagnosis of electrical machines. Several studies have developed AI-based fault detection methods as reviewed in [15]–[18]. They mostly use stator currents or vibration signals from additional sensors to extract the fault data.

While most of these studies are for bearing fault diagnosis [15], few are for ISCF detection. A data-driven online detection method utilizing multiple classifiers is proposed in [19]. The fault information is acquired from phase currents and voltages. ISCFs down to 2% could have been detected. A multi-layer perceptron is trained to detect ISCFs down to 0.6% in [20]. The three-phase shifts are utilized as the input data. An unsupervised learning-based NN using phase currents for fault detection is reported in [21]. Similarly, ISCF detection is achieved in [22] using an NN-based method on stator currents of an IM driven by an inverter via open-loop scalar *V/f* control. While these studies [19]–[22] consider induction machines that work in an open-loop fashion, a closed-loop controlled IM, driven by a model predictive control structure, is considered in this paper, which constitutes a fundamental difference. Although an NN-based method detecting ISCFs down to 4.2% in a closed-loop controlled permanent magnet machine is reported in [23], no details regarding the controller structure or the controller-fault interaction are provided. Phase currents and speed information are the inputs.

This study directly utilizes the switching sequences generated by the finite control-set model predictive controller (FCS-MPC) for ISCF diagnosis, therefore no additional sensor or hardware is required for fault detection purposes. The voltage vectors of a two-level voltage source inverter (2L-VSI) are depicted in Fig. 1a. In the standard FCS-MPC, the controller finds the optimum voltage vector in view of the control outcomes and applies it at the next switching instant. Hence, the controller outcomes are discrete voltage vectors and convenient for statistical approaches. The switching data utilized in this paper, which correspond to the healthy operation and the ISCF case (i.e., the faulty case), are obtained from the experimental setup shown in Fig. 2. The parameters of the IM are presented in Table I. The ISCF condition corresponds to a short-circuiting of 2-turns out of 104-turns in a phase winding.

A motor drive inverter with model predictive control produces an (almost) uniform distribution of active switching vectors while driving a healthy induction machine. However, an inter-turn short circuit in the stator changes the distribution of active switching vectors since the driver tries to compensate for the fault's influence for the proper operation of the motor (Fig. 1b). This observation constitutes the basis for this study. Our primary approach is to train a neural network with the switching vector data collected for healthy and faulty cases so that the trained structure can identify and locate an ISCF. The detection performance results prove the effectiveness of the proposed approach.

There are three main contributions of this paper.

1) A machine learning-based, non-invasive ISCF detection method using inverter switching statistics is introduced.
2) The first publicly available ISCF detection dataset containing switching vector data collected at various load torque and shaft speed values for healthy and faulty states of an induction machine is released.
3) Performance baselines on the released dataset for three main neural network architectures, namely, multi-layer perceptron, convolutional neural network, and recurrent neural network are provided.

TABLE I: **Induction Machine (IM) parameters.**

Name	Symbol	Value
Apparent power	S	640 VA
Stator voltage	V_{ab}	128 V
Stator current	I	2.9 A
Base frequency	f	50 Hz
Torque	T_e	1.2 N.m
Number of poles	p	2
Total turns in one phase	N_s	104
Stator resistance	R_s	2.3 Ω
Rotor resistance	R_r	3.1 Ω
Magnetizing inductance	L_m	98 mH
Stator leakage inductance	L_{ls}	4 mH
Rotor leakage inductance	L_{lr}	2 mH
Stator flux magnitude reference	$\|\Psi_s\|^{ref}$	0.3 Wb

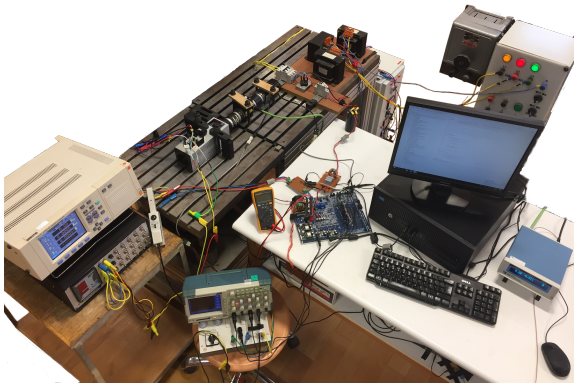


Fig. 2: **Experimental setup.** An FCS-MPC driven IM is used in the experiments. The ISCF condition corresponds to a short-circuiting of 2-turns out of 104-turns in a phase winding.

II. Neural Network Based ISCF Detection

This study designs neural network models detecting inter-turn short circuit faults in an induction machine driven by an inverter with model predictive control (Fig. 1c and Fig. 3).

TABLE II: **Data for ISCF detection.** Inverter switching vectors were collected for different conditions of an induction machine with a rated speed of $w = 3000$ rpm and a rated torque of $T = 1.20$ N·m. For each data series, speed (w), torque (T), measured electrical frequency (f_e), and machine status as "Healthy" or "Faulty" are presented. Each data series in the first part was divided into two such that the first 70% and the remaining 30% were used to create samples for the training and validation sets, respectively. Data series in the second part were used to create samples for the test set.

Training and Validation					**Test**				
id	w (rpm)	T (N·m)	f_e (Hz)	Machine Status	id	w (rpm)	T (N·m)	f_e (Hz)	Machine Status
001	1500	0.30	25.5	Healthy	035	1500	0.30	25.5	Healthy
002	1500	0.30	25.5	Faulty	036	1500	0.30	25.5	Faulty
003	1500	0.90	28.0	Healthy	037	1500	0.90	28.0	Healthy
004	1500	0.90	28.0	Faulty	038	1500	0.90	28.0	Faulty
005	2250	0.30	38.1	Healthy	039	2250	0.30	38.1	Healthy
006	2250	0.30	38.1	Healthy	040	2250	0.30	38.1	Faulty
007	2250	0.30	38.1	Faulty	041	2250	0.30	38.1	Faulty
008	2250	0.30	38.1	Faulty	042	2250	1.25	42.0	Healthy
009	2250	1.25	42.0	Healthy	043	2250	1.25	42.0	Healthy
010	2250	1.25	42.0	Faulty	044	2250	1.25	42.0	Faulty
011	2250	1.25	42.0	Faulty	045	2250	1.25	42.0	Faulty
012	3000	0.30	50.6	Healthy	046	3000	0.30	50.6	Healthy
013	3000	0.30	50.6	Healthy	047	3000	0.30	50.6	Healthy
014	3000	0.30	50.6	Faulty	048	3000	0.30	50.6	Faulty
015	3000	0.30	50.6	Faulty	049	3000	0.30	50.6	Faulty
016	3000	1.20	54.4	Healthy	050	3000	1.20	54.4	Healthy
017	3000	1.30	54.8	Faulty	051	3000	1.30	54.8	Faulty
018	3000	1.35	55.3	Healthy	052	3000	1.35	55.3	Healthy
019	3000	1.35	55.3	Faulty	053	3000	1.35	55.3	Faulty
020	3750	0.30	63.1	Healthy	054	3750	0.30	63.1	Healthy
021	3750	0.30	63.1	Healthy	055	3750	0.30	63.1	Faulty
022	3750	0.30	63.1	Faulty	056	3750	0.30	63.1	Faulty
023	3750	0.30	63.1	Faulty	057	3750	1.25	67.5	Healthy
024	3750	1.25	67.5	Healthy	058	3750	1.25	67.5	Healthy
025	3750	1.25	67.5	Faulty	059	3750	1.25	67.5	Faulty
026	3750	1.25	67.5	Faulty	060	3750	1.25	67.5	Faulty
027	4500	0.30	75.6	Healthy	061	4500	0.30	75.6	Healthy
028	4500	0.30	75.6	Healthy	062	4500	0.30	75.6	Faulty
029	4500	0.30	75.6	Faulty	063	4500	1.15	79.8	Healthy
030	4500	0.30	75.6	Faulty	064	4500	1.15	79.8	Faulty
031	4500	1.15	79.8	Healthy					
032	4500	1.15	79.8	Healthy					
033	4500	1.15	79.8	Faulty					
034	4500	1.15	79.8	Faulty					

We formulate ISCF detection as a classification problem using inverter switching statistics.

A. Problem Formulation

Let $\boldsymbol{X} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N\}$ be a set of training samples such that each sample $\boldsymbol{x}_i \in \mathbb{R}^D$ has a corresponding ground-truth label $\boldsymbol{y}_i = [y_i^1, \cdots, y_i^K] \in \{0,1\}^K$ where $\sum_{k=1}^{K} y_i^k = 1$. Given a model is represented as a function parameterized by θ, $f_\theta : \mathbb{R}^D \rightarrow \mathbb{R}^K$, it predicts the label of an input sample $\boldsymbol{x}_i$ as $\hat{\boldsymbol{y}}_i = f_\theta(\boldsymbol{x}_i) = [\hat{y}_i^1, \cdots, \hat{y}_i^K] \in \mathbb{R}^K$ such that $\hat{y}_i^k \geq 0 \; \forall k$ and $\sum_{k=1}^{K} \hat{y}_i^k = 1$. We train a model end-to-end using categorical-cross entropy as the loss function (1).

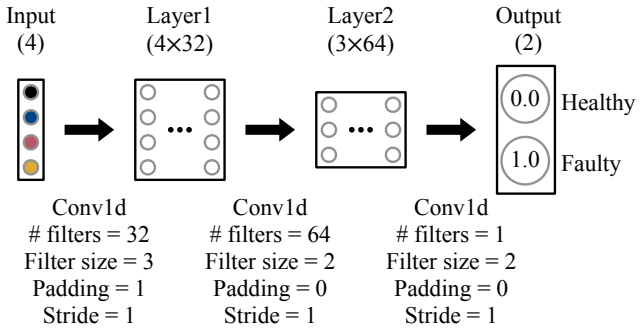
$$loss = \frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K} y_i^k \log \hat{y}_i^k \tag{1}$$

B. Neural Network Model Architectures

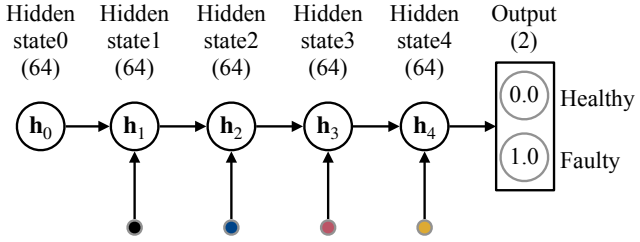
We constructed three different models using multi-layer perceptron (MLP), convolutional neural network (CNN), and recurrent neural network (RNN) architectures. Models are designed such that they have almost the same number of learnable parameters, i.e. 'capacity' (MLP: 4612, CNN: 4417, and RNN: 4418 learnable parameters). A model accepts a histogram of inverter switching statistics at the input and predicts the machine's status (healthy or faulty) at the output.

The multi-layer perceptron model consists of an input layer with 4 nodes, two hidden layers with 64 nodes, and an output layer with 2 nodes (Fig. 1c). Each layer computes a weighted sum of its inputs ($s_j = \sum_i w_{ji} x_i + b_j$, where $\mathbf{s} = [s_j]$ is the output vector, $\mathbf{x} = [x_i]$ is the input vector, $\mathbf{W} = [w_{ji}]$ is the learnable weight matrix, and $\mathbf{b} = [b_j]$ is the learnable bias vector), followed by a non-linear activation function. Hidden layers have a ReLU activation function ($f(s) = \max(0, s)$) followed by a dropout with a rate of 0.5. The output layer has a softmax activation function producing normalized probability values, i.e., adding up to 1.

The convolutional neural network model consists of three convolutional layers with 32, 64, and 2 filters, respectively (Fig. 3a). Each convolutional layer computes a cross-correlation of its inputs and filter weights ($s_j = \sum_i w_i x_{j+i} + b$, where $\mathbf{W} = [w_{ji}]$ is the learnable and shared filter weights and b is the learnable bias). Except the last layer, each convolutional layer is followed by a ReLU activation function



(a) Convolutional neural network.



(b) Recurrent neural network.

Fig. 3: **Neural network models.**

and a dropout with a rate of 0.5. Similar to the MLP model, outputs at the last layer are normalized using a softmax activation function.

The recurrent neural network model consists of a recurrent cell containing 64 hidden nodes and a fully connected layer as a linear classifier on top (Fig. 3b). A recurrent cell computes an affine transformation of the input and the previous hidden state, adds them up and passes through a non-linear activation function to compute the current hidden state ($h_t = f(\mathbf{W}_{ih}\mathbf{x} + \mathbf{b}_{ih} + \mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{b}_{hh})$, where $\mathbf{W}_{ih}$ and $\mathbf{W}_{hh}$ are learnable weight matrices, and $\mathbf{b}_{ih}$ and $\mathbf{b}_{hh}$ are learnable bias vectors). Note that non-linear activation function is ReLU, and after each recurrent step a dropout with a rate of 0.5 is applied on hidden state.

C. Experimental Setup and Preparation of the Machine Learning Dataset

The switching vectors of the motor drive inverter were collected from the same experimental setup utilized in [13]. A photo of the setup is provided in Fig. 2. The parameters of the IM, on which intentional ISCFs can be created for tests, are given in Table I as utilized in the MPC loop. The motor drive development kit TMDXIDDK379D from Texas Instruments is used as the motor drive inverter. The interested reader is referred to [24] for detailed descriptions regarding the FCS-MPC structure, the laboratory implementation, and various test results, including motor drive operation and the ISCF detection through switching vector analysis based on a simple thresholding method.

For several different combinations of speed and torque at healthy and faulty states (see Table II), the voltage vectors (decided by the controller and executed by the inverter) are recorded as time series of 22000 elements. The FCS-MPC algorithm has a control frequency of 40 kHz, therefore the record for the switching vector array of 22000 elements corresponds to a total of 0.55 second time interval. In the creation of Table II, frequency and torque values are read from the waveform analyzer and torque sensor respectively, which are involved in the experimental setup shown in Fig. 2. ISCFs were introduced over 2 out of 104 turns in a phase winding of a star connected IM.

Collected data series were segregated into two sets such that the first set was for training and validation, and the second set was for the test. Each data series in the first set was further split into two. The first 70% of data points constituted a data series for training, and the remaining 30% constituted a data series for validation. Then, machine learning datasets of training, validation, and test were prepared by creating sample and label pairs over respective data series.

Over a data series, multiple samples were created in a sliding window fashion with a step size of one. A sample was created by calculating the histogram of switching vectors over a window of five electrical periods (≈ 92 ms at the rated speed and torque) (Fig. 1b). Note that switching vectors were aggregated as 0-vectors ($v_0 - v_7$), phase A vectors ($v_1 - v_4$), phase B vectors ($v_3 - v_6$), and phase C vectors ($v_2 - v_5$). The machine's status (i.e., healthy or faulty) was assigned as the sample's label.

D. Training and Testing of The NN Model

A neural network model was trained on samples created from training data series using the loss function given in Eq. 1. Early stopping based on loss in the validation dataset was employed to avoid overfitting. Finally, the model's performance was evaluated on the unseen test dataset.

The area under the receiver operating characteristic curve (AUROC) was used as the performance metric. We also calculated the 95% confidence interval (CI) of the AUROC using the percentile bootstrap method [25].

E. Data and Code Availability

All original code and dataset have been deposited at Zenodo under the https://doi.org/10.5281/zenodo.6774360 and made publicly available [26].

III. Results

A. The NNs Detect ISCF

We checked the performance of neural network models on ISCF detection. For each sample in the test dataset, we obtained a prediction from each trained model and plotted the receiver operating characteristic curves. We obtained AUROC values of 0.9946 (95% CI: 0.9945 - 0.9947), 0.9942 (95% CI: 0.9940 - 0.9943), and 0.9950 (95% CI: 0.9949 - 0.9951) for the MLP, CNN, and RNN models, respectively.

Moreover, we checked the effect of the model's capacity on the model's performance by varying the hidden layer sizes in the MLP model. As expected, we observed that the model's performance decreased with decreasing model capacity (Fig. 4). Nevertheless, the performance decrease was

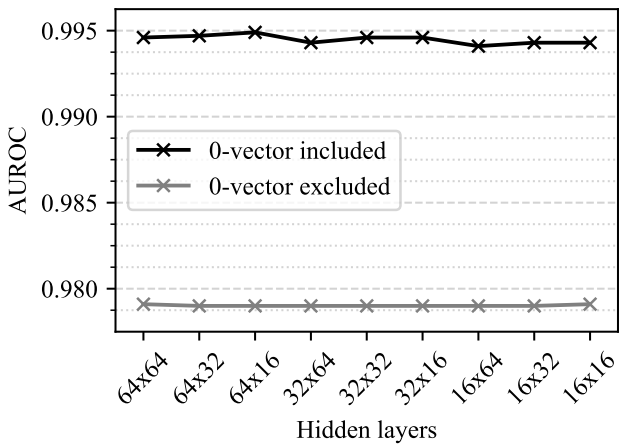


Fig. 4: **The neural network models detect ISCF.** Different NN models using the same architecture with different number of nodes in the hidden layers were trained and tested on the ISCF detection task. The area under receiver operating characteristics curve (AUROC) calculated on the test set was used as performance metric. We trained and tested the models using histogram of all vectors (0-vector included) and active vectors only (0-vector excluded).

not drastic. Hence, smaller networks could be preferable for real-world deployment since they require less computational power.

Besides, we checked the model's performance on identifying if a data series was collected at the healthy or faulty state of the machine. Prediction for a data series was obtained as the average of predictions of all samples created from the data series. Although we observed wrong predictions at the sample level in the few data series, the model *perfectly* identified the machine's status at the data series level (Fig. 5).

B. 0-vectors Help The NN Detect ISCF

We know that the proportion of 0-vectors provides information about the speed and torque of the machine, which can be valuable for the model. To test the effect of 0-vectors on the performance of the neural network model, we excluded 0-vectors from histogram calculation and reran our experiments. As expected, we observed a performance decrease (Fig. 4). Hence, we concluded that 0-vectors helped the neural network models detect ISCF.

Furthermore, we observed a decrease in fault detection performance for the speed values beyond 3750 rpm (Fig. 5). This region corresponds to the verge of overmodulation in an inverter control system with a carrier-based modulation, where the percentage of zero vectors significantly decreases. The switching vector statistics were also not as responsive to the occurrence of the fault as they were at the slower speeds.

C. Statistics Over Longer Intervals Improve Performance

We trained and tested our models with sample and label pairs. A sample was created by calculating the histogram of switching vectors over a window of five electrical periods ($\approx$ 92 ms at the rated speed and torque). In an ideal system (where there is no noise), an ISCF can be detected using inverter switching statistics calculated over one electrical period. However, the system is noisy in practice, and the noise directly affects the model's performance.

We investigated the effects of the interval length, over which switching vector histograms were calculated, on the performance of neural network models. We observed that as the length of the interval decreased (multiples of electrical period: 5, 4, 3, 2, and 1), the model's performance also decreased (Fig. 6). In short, we concluded that statistics over longer intervals improved the model's performance.

One interesting observation was that the performance decay was gradual from five to two periods. Nevertheless, it was drastic from two periods to a single period. Our observation was also consistent with the findings of [13], which uses a thresholding-based method over switching vector statistics. Besides, the contribution of 0-vectors in the model's performance was evident (Fig. 6).

D. The NN Identifies Faulty Phase in a Machine with ISCF

After we showed that a neural network model successfully detected ISCF, we also checked if it could identify the faulty phase. We prepared a small dataset using data series collected at the rated speed and around the rated torque of the induction machine (Table III).

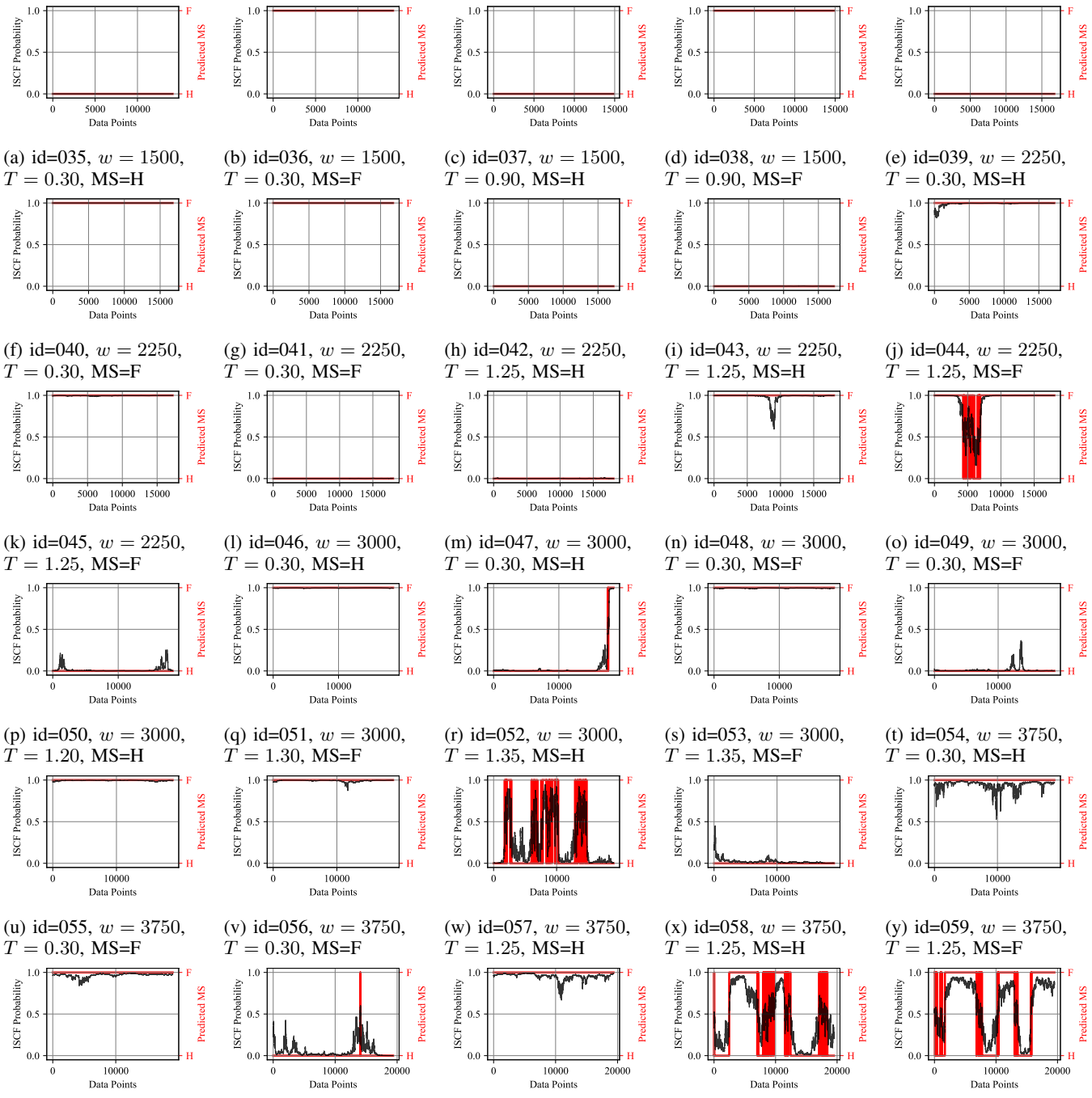
We modified the neural network architecture used in ISCF detection to a multi-class classification model with four classes corresponding to the machine's status of healthy, faulty (A), faulty (B), and faulty (C). Then, we trained the model on the training set with early stopping based on loss in the validation set and evaluated its performance on the test set. The model successfully detected ISCF and identified the faulty phase (Table IV). It achieved an accuracy of 0.9995. As in the ISCF detection task, our neural network model perfectly detected ISCF and identified the faulty phase at the data series level.

E. The NN Outperforms Thresholding Based Method

The duration of high electrical currents passing through the shorted turns during an ISCF is critical for the repair and possible fault-tolerant operation of the machine. As the detection time takes longer, the ISCF condition will evolve further into the unmanageable situations such as complete phase to phase or phase to ground shorts. Therefore, we compared the performance of our neural network model with the thresholding-based method of [13] in terms of ISCF detection time. At various load torque and shaft speed values, while the thresholding-based method detected ISCF in between 0.5 to 2 seconds, the neural network model detected ISCF in between 0.074 to 0.196 seconds. There was a speedup of more than two times at the rated operating conditions ($\approx$ 0.2 second for the thresholding-based method in [24] and 0.092 seconds for our neural network model).

IV. Conclusion

Early detection of an ISCF in an electrical machine is vital for its maintenance. This study developed neural network



(a) id=035, $w = 1500$, $T = 0.30$, MS=H
(b) id=036, $w = 1500$, $T = 0.30$, MS=F
(c) id=037, $w = 1500$, $T = 0.90$, MS=H
(d) id=038, $w = 1500$, $T = 0.90$, MS=F
(e) id=039, $w = 2250$, $T = 0.30$, MS=H
(f) id=040, $w = 2250$, $T = 0.30$, MS=F
(g) id=041, $w = 2250$, $T = 0.30$, MS=F
(h) id=042, $w = 2250$, $T = 1.25$, MS=H
(i) id=043, $w = 2250$, $T = 1.25$, MS=H
(j) id=044, $w = 2250$, $T = 1.25$, MS=F
(k) id=045, $w = 2250$, $T = 1.25$, MS=F
(l) id=046, $w = 3000$, $T = 0.30$, MS=H
(m) id=047, $w = 3000$, $T = 0.30$, MS=H
(n) id=048, $w = 3000$, $T = 0.30$, MS=F
(o) id=049, $w = 3000$, $T = 0.30$, MS=F
(p) id=050, $w = 3000$, $T = 1.20$, MS=H
(q) id=051, $w = 3000$, $T = 1.30$, MS=F
(r) id=052, $w = 3000$, $T = 1.35$, MS=H
(s) id=053, $w = 3000$, $T = 1.35$, MS=F
(t) id=054, $w = 3750$, $T = 0.30$, MS=H
(u) id=055, $w = 3750$, $T = 0.30$, MS=F
(v) id=056, $w = 3750$, $T = 0.30$, MS=F
(w) id=057, $w = 3750$, $T = 1.25$, MS=H
(x) id=058, $w = 3750$, $T = 1.25$, MS=H
(y) id=059, $w = 3750$, $T = 1.25$, MS=F
(z) id=060, $w = 3750$, $T = 1.25$, MS=F
(aa) id=061, $w = 4500$, $T = 0.30$, MS=H
(ab) id=062, $w = 4500$, $T = 0.30$, MS=F
(ac) id=063, $w = 4500$, $T = 1.15$, MS=H
(ad) id=064, $w = 4500$, $T = 1.15$, MS=F

Fig. 5: **Sample-level ISCF predictions on the data series in the test dataset.** At each point of a data series (i.e., for a sample), inter-turn short circuit fault (ISCF) probability was obtained from the trained model. ISCF probability and predicted machine status (MS) obtained by thresholding the predicted probability value with 0.5 are presented for data series. For each data series, id, speed (rpm), torque (N·m), and MS ((H)ealthy, (F)aulty) are also given.

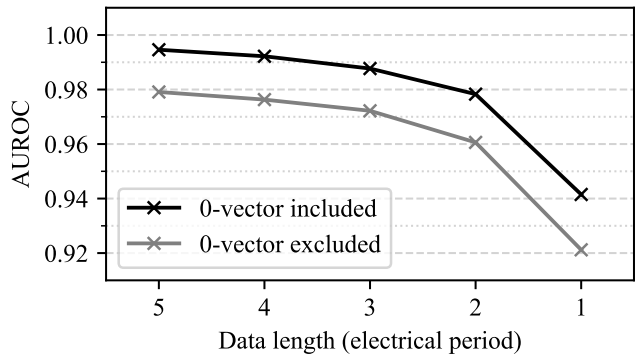


Fig. 6: **Statistics over longer intervals improve performance.** Switching vector histograms were calculated over an interval of multiples of electrical frequency (5,4,3,2, and 1) while training the NN models. We compared the effect of interval length on the model's performance using AUROC on the test set both for 0-vector included and 0-vector excluded cases.

TABLE III: **Data for faulty phase detection.** For each data series, speed (w), torque (T), measured electrical frequency (f_e), and machine status as "Healthy" or "Faulty" with phase info indicated in parentheses (A/B/C) are presented. Each data series in the top part was divided into two such that the first 70% and the remaining 30% were used to create samples for the training and validation sets, respectively. Data series in the bottom part were used to create samples for the test set.

Training and Validation				
id	w (rpm)	T (N·m)	f_e (Hz)	Machine Status
016	3000	1.20	54.4	Healthy
018	3000	1.35	55.3	Healthy
017	3000	1.30	54.8	Faulty (A)
065	3000	1.25	54.6	Faulty (B)
066	3000	1.25	54.6	Faulty (C)
Test				
id	w (rpm)	T (N·m)	f_e (Hz)	Machine Status
050	3000	1.20	54.4	Healthy
052	3000	1.35	55.3	Healthy
051	3000	1.30	54.8	Faulty (A)
067	3000	1.25	54.6	Faulty (B)
068	3000	1.25	54.6	Faulty (C)

TABLE IV: **Faulty phase detection.** The same NN architecture was modified to a multi-class classification model detecting faulty phase as well. The confusion matrix obtained on the test set is presented.

		Predicted			
		Healthy	Faulty (A)	Faulty (B)	Faulty (C)
Truth	Healthy	1.000	0.000	0.000	0.000
	Faulty (A)	0.000	1.000	0.000	0.000
	Faulty (B)	0.000	0.000	1.000	0.000
	Faulty (C)	0.003	0.000	0.000	0.997

models that detect an ISCF in an IM driven by an inverter with an MPC algorithm. The models accepted the histogram of inverter switching vectors, which are readily available, as input and predicted the machine's status (healthy or faulty) at the output (Fig. 1). An ISCF in the IM was successfully detected under 0.1 seconds with an almost perfect performance (Fig. 4). Besides, the faulty phase was identified with an accuracy of 0.9995 (Table IV).

In our experiments, while the large networks performed slightly better in the ISCF detection task, the performance of the smaller networks were also good enough for real-world deployment (Fig. 4). Moreover, a small network requires less memory and computational resources, facilitating its deployment in the same processor alongside the controller algorithm.

Our experiments validated that 0-vectors contained valuable information for ISCF detection (Fig. 4), and statistics over longer intervals improved the performance (Fig. 6). Nevertheless, there was a trade-off between better performance and faster ISCF detection in determining the optimum interval for statistics calculation. We concluded that an interval of three to five electrical periods was reasonable.

Lastly, we observed that the model's performance started to degrade beyond the rated speed and torque values (Fig. 5), which corresponds to the operation on the verge of overmodulation, where the utilization of zero vectors significantly decreases. This could be due to the limited available data around these operation regions of the IM. We had around 300k samples in our training set; however, they were created from only 26 independent data series (Table II). Since data collection is quite expensive, our dataset was minimal compared to traditional deep learning datasets containing millions of independent samples [27].

Hence, the collection of an extensive dataset with broad coverage of the IM's operation regions and the real-world deployment of our ISCF detection models are reserved for future work.

ACKNOWLEDGMENTS

We thank Öztürk Şahin Alemdar for constructive comments on the manuscript.

REFERENCES

[1] S. Nandi, H. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—a review," *IEEE Transactions on Energy Conversion*, vol. 20, no. 4, pp. 719–729, 2005.

[2] H. Henao, G.-A. Capolino, M. Fernandez-Cabanas, F. Filippetti, C. Bruzzese, E. Strangas, R. Pusca, J. Estima, M. Riera-Guasp, and S. Hedayati-Kia, "Trends in fault diagnosis for electrical machines: A review of diagnostic techniques," *IEEE Industrial Electronics Magazine*, vol. 8, no. 2, pp. 31–42, 2014.

[3] A. Gandhi, T. Corrigan, and L. Parsa, "Recent advances in modeling and online detection of stator interturn faults in electrical motors," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 1564–1575, 2011.

[4] Y. Chen, S. Liang, W. Li, H. Liang, and C. Wang, "Faults and diagnosis methods of permanent magnet synchronous motors: A review," *Applied Sciences*, vol. 9, no. 10, p. 2116, 2019.

[5] A. Bellini, F. Filippetti, C. Tassoni, and G.-A. Capolino, "Advances in diagnostic techniques for induction machines," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 12, pp. 4109–4126, 2008.

[6] A. Siddique, G. Yadava, and B. Singh, "A review of stator fault monitoring techniques of induction motors," *IEEE Transactions on Energy Conversion*, vol. 20, no. 1, pp. 106–114, 2005.

[7] R. T. Purushottam Gangsar, "Signal based condition monitoring techniques for fault detection and diagnosis of induction motors: A state-of-the-art review," *Mechanical Systems and Signal Processing*, vol. 144, 2020.

[8] A. H. Bonnett and C. Yung, "Increased efficiency versus increased reliability," *IEEE Industry Applications Magazine*, vol. 14, no. 1, pp. 29–36, 2008.

[9] A. Bonnett and G. Soukup, "Cause and analysis of stator and rotor failures in three-phase squirrel-cage induction motors," *IEEE Transactions on Industry Applications*, vol. 28, no. 4, pp. 921–937, 1992.

[10] C. Gerada, K. Bradley, M. Sumner, P. Wheeler, S. Picker, J. Clare, C. Whitley, and G. Towers, "The results do mesh," *IEEE Industry Applications Magazine*, vol. 13, no. 2, pp. 62–72, 2007.

[11] S. Cheng, P. Zhang, and T. G. Habetler, "An impedance identification approach to sensitive detection and location of stator turn-to-turn faults in a closed-loop multiple-motor drive," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 1545–1554, 2011.

[12] M. Zafarani, E. Bostanci, Y. Qi, T. Goktas, and B. Akin, "Interturn short-circuit faults in permanent magnet synchronous machines: An extended review and comprehensive analysis," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 6, no. 4, pp. 2173–2191, 2018.

[13] İ. Şahin and O. Keysan, "Model predictive controller utilized as an observer for inter-turn short circuit detection in induction motors," *IEEE Transactions on Energy Conversion*, vol. 36, no. 2, pp. 1449–1458, 2021.

[14] B. K. Bose, "Artificial intelligence techniques: How can it solve problems in power electronics?: An advancing frontier," *IEEE Power Electronics Magazine*, vol. 7, no. 4, pp. 19–27, 2020.

[15] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, "Deep learning algorithms for bearing fault diagnostics—a comprehensive review," *IEEE Access*, vol. 8, pp. 29 857–29 881, 2020.

[16] M. Seera, C. P. Lim, S. Nahavandi, and C. K. Loo, "Condition monitoring of induction motors: A review and an application of an ensemble of hybrid intelligent models," *Expert Systems with Applications*, vol. 41, no. 10, pp. 4891–4903, 2014.

[17] W. Lang, Y. Hu, C. Gong, X. Zhang, H. Xu, and J. Deng, "Artificial intelligence-based technique for fault detection and diagnosis of ev motors: A review," *IEEE Transactions on Transportation Electrification*, pp. 1–1, 2021.

[18] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5990–5998, 2018.

[19] Z. Xu, C. Hu, F. Yang, S.-H. Kuo, C.-K. Goh, A. Gupta, and S. Nadarajan, "Data-driven inter-turn short circuit fault detection in induction machines," *IEEE Access*, vol. 5, pp. 25 055–25 068, 2017.

[20] M. B. K. Bouzid, G. Champenois, N. M. Bellaaj, L. Signac, and K. Jelassi, "An effective neural approach for the automatic location of stator interturn faults in induction motor," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 12, pp. 4277–4289, 2008.

[21] J. F. Martins, V. Ferno Pires, and A. J. Pires, "Unsupervised neural-network-based algorithm for an on-line diagnosis of three-phase induction motor stator fault," *IEEE Transactions on Industrial Electronics*, vol. 54, no. 1, pp. 259–264, 2007.

[22] M. Skowron, T. Orlowska-Kowalska, M. Wolkiewicz, and C. T. Kowalski, "Convolutional neural network-based stator current data-driven incipient stator fault diagnosis of inverter-fed induction motor," *Energies*, vol. 13, no. 6, p. 1475, 2020.

[23] H. Lee, H. Jeong, G. Koo, J. Ban, and S. W. Kim, "Attention recurrent neural network-based severity estimation method for interturn short-circuit fault in permanent magnet synchronous machines," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 4, pp. 3445–3453, 2021.

[24] İ. Şahin, "Model predictive torque control of an induction motor enhanced with an inter-turn short circuit fault detection feature," Ph.D. dissertation, Middle East Technical University, 2021.

[25] B. Efron, "Bootstrap methods: another look at the jackknife," in *Breakthroughs in statistics*. Springer, 1992, pp. 569–593.

[26] M. Ümit ÖNER, İlker ŞAHİN, and O. KEYSAN, "Inter-turn short circuit fault (iscf) detection," Jun. 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6774360

[27] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

Mustafa Umit Oner got his B.S. degree in 2013 and M.S. degree in 2016 from the Electrical and Electronics Engineering Department at the Middle East Technical University, Ankara, Turkey. Then, he got his Ph.D. degree in 2021 from the Computer Science Department at the National University of Singapore, Singapore. He continues his research at the Bahcesehir University, Istanbul, Turkey. He is interested in developing novel machine learning models and machine learning-based information systems for digital histopathology and integrative multi-omics to support diagnostic and therapeutic decision-making in cancer.

İlker Şahin received the B.Sc., M.Sc., and Ph.D. degrees in electrical and electronics engineering, in 2010, 2014, and 2021 respectively, from Middle East Technical University (METU), Ankara, Turkey. From 2011 to 2020, he was a Teaching Assistant with METU, then he joined Aselsan in 2020, as a Lead Design Engineer. His current research interests include high-performance motor drives, predictive control, and fault diagnosis.

Ozan Keysan received the master's degree from Middle East Technical University (METU), Ankara, Turkey, in 2008, and the Ph.D. degree from the University of Edinburgh, Edinburgh, Scotland, in 2014. He is currently an Associate Professor with the Electrical and Electronics Engineering Department, METU. His current research interests include renewable energy, design, and optimization of electrical machines, smart grids, superconducting machines, and permanent-magnet machines.