ITSC Fault Diagnosis in Permanent Magnet Synchronous Motor Drives Using Shallow CNNs

Vera Szabo, Saeed Hasan Ebrahimi, Martin Choux, and Morten Goodwin

Department of Engineering Sciences
University of Agder
NO-4879 Grimstad, Norway

Abstract – Due to its features, permanent magnet synchronous motor (PMSM) has gained popularity and is used in various industrial applications, including those with high downtime costs like offshore equipment. Inter-turn short-circuit (ITSC) fault is one of the most typical PMSM faults and therefore is its early diagnostics in real-time highly valuable. Solving the problem using conventional signal, model-based, or data-driven approaches faces challenges such as computational complexity, time demand, or need for detailed domain expertise. This paper presents a computationally simple, robust, and accurate method based on the 2D convolutional neural network (CNN). The proposed data-driven model has first been validated with the help of experimental data obtained from an inverter fed PMSM subject to ITSC faults in different time intervals, and secondly its performances have been compared to a model-based structural analysis approach using Dulmage-Mendelsohn decomposition tool. The comparison is based on the same data. Results show that the accuracy of the CNN model for diagnosing early faults is more than 98% without doing additional comprehensive fine-tuning. In addition, the paper presents a robust method that can be successfully used as a metric for fast fault detection benchmark.

G.1 Introduction

Permanent magnet synchronous motors (PMSMs) are deployed in various industrial systems, such as offshore equipment, wind generators, robotics or electric vehicles. While having conventional three phase windings in the stator, PMSMs produce their rotor magnetic flux by the mean of permanent magnets, either embedded tangentially around the rim of the rotor as seen in Fig. G.1, or buried radially for higher performances. Their efficiency (92% - 97%) is significantly higher compared to traditional asynchronous motors (75% - 92%) [1], while low reactive power consumption, improved dynamic performance, light weight, and small dimensions are further reasons for their increased popularity. More than 40% of all faults in synchronous motors start as stator related [2]. Among those, the inter-turn short-circuit (ITSC) faults are the most common, however difficult to detect automatically [3], which is partially caused by ITSC faults having little effect to the motor performance in early stages. However, if not discovered and mended in time, the ITSC fault can quickly grow into severe motor damage and consequently lead to total failure

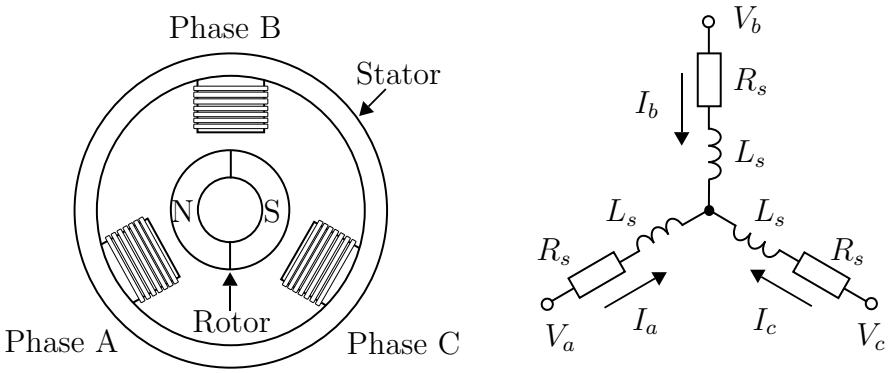


Figure G.1: Basic structure of a PMSM

of the system [4]. The drive to cut operating and maintenance costs and increase operational safety is pushing the agenda in the industry towards the adoption of predictive maintenance strategies. In this process, fault diagnosis, i.e. fault detection and isolation, represents an important part. A proven method for diagnosis of ITSC faults in their early phase, that is easy to implement in practice, is therefore in great demand.

This paper focuses on a simple detection and diagnosis method for ITSC faults. While deep CNNs have several layers, often of various types, shallow CNNs have only one besides the input and the output layer. One objective of this study is to explore whether the simplest CNNs can be successfully used as ITSC fault classifiers, i.e with high enough accuracy, as they do not overfit on small datasets and require less computational time and energy consumption than deep CNNs. Indeed, awareness about CO_2 emission in machine learning research started to arise lately [5]. According to the EU Annual Report on SMEs (2019), in the EU just 6% of the SMEs use AI, although they represent 99.8% of all enterprises in the EU-27, with lack of skill to be one of the main obstacles. Therefore, any model that is easy to use, does not require high computational power and shows robustness is of huge demand by the industry. Following a short literature overview in section 2, the proposed method is described in section 3 while section 4 details the results and compares the performances with a model-based method using the same experimental setup and dataset. Finally, conclusions are drawn in Section 5.

G.2 Literature Review

Fault diagnosis can be divided from data processing into model-based, signal-based and data-driven methods. The literature on PMSMs show that faults, as for example inter-turn short circuit or demagnetisation faults, can be early detected using either of the methods. Comprehensive reviews of methods for detection and diagnosis of ITSC faults in PMSMs are presented in [3] and [6].

The model-based methods establish a mathematical model based on principles of physics that describes the actual machine. The most accurate results have been achieved with finite element analysis (FEA) models which compared to other models also have the highest computational cost as well [6]. Other types of models, such as equivalent circuit,

field reconstruction and linear PMSM models, are beneficial in understanding how the fault behaves assuming that they are detailed enough [3]. The signal-based and data driven methods use statistical tools and mathematical transformations to identify and extract fault patterns from signals such as current, voltage, vibrations and so on. Motor current signal analysis (MCSA) is the most common model and is extensively studied [6].

Artificial intelligence (AI) and machine learning (ML) -based approaches show increased performance compared to conventional signal-based models providing a solution for the complexity introduced by increased data quantity. However, it is often not easy to apply traditional ML techniques in practice, due to lack of efficient methods to obtain training data, and specific knowledge needed to train the models [7, 8]. The traditional ML with tailor made and handcrafted features — typically used by applying feature extraction and learning algorithms such as support vector machine (SVM), random forest (RF), principle component analysis (PCA) or linear decrement analysis (LDA) [9] — has been used for many years while deep learning (DL) methods emerged in 2006.

DL represents a breakthrough in the field of AI and shows state-of-the-art performance when compared to traditional machine learning in many fields. Constructing a ML system needs careful engineering and high domain expertise to design a feature extractor that transforms the raw data into a suitable representation from which the learning model can detect or classify patterns [10].

In contrast, when it comes to DL, the features are learned automatically from raw data. DL models used in motor fault detection and diagnosis include for instance deep belief networks [11], generative adversarial networks (GAN) [12, 13], long short term memory models (LSTM) [14]. Among DL methods for fault diagnosis, extensively used are CNNs [15]. 1D CNNs are, among others, used with direct input of time-domain signals collected in motors [16,17], while 2D CNNs are, among others, used by converting the time-domain raw signals into 2D grey images without further feature extraction [18].

G.3 The proposed Method

This section presents the proposed data-driven method based on a shallow 2D CNN. The input to the model is obtained from an experimental setup where a PMSM is run through a healthy and three faulty sequences. Switching the faults on and off is done by controllable relays placed between the winding taps. The three faulty sequences represent the three ITSC faults. Each fault is applied on a designated phase by short-circuiting different numbers of turns, resulting in different fault percentage, enabling establishing a fault of less than 1% in terms of number of short-circuited turns per total number of turn in one phase. Altogether, ten different features (four voltages, four currents, position of the rotor and speed) in their raw form, without any prepossessing, have been used as input to the model.

CNNs, primarily used for pattern recognition tasks, especially within images, usually consist of three types of layers: convolutional, pooling, and fully-connected layers. As the name indicates, the convolutional layers play the most important role, where the learnable parameters origin from kernels. The structure of the proposed classification

Table G.1: Motor parameters

Parameter	Value
Rated DC bus voltage	280 V
Rated rms phase current	5 A
Rated output torque	7 Nm
Rated speed	1500 rpm
Stator resistance	0.8 Ω
Stator inductance	0.5 mH
Rotor inertia	30.065 kgm^2
Pole pairs	2

model is outlined in Fig. G.2. It has only one convolutional layer. The output of the model is one of the 4 classes: no-fault and ITSC faults at phases A, B and C. We use SHAP

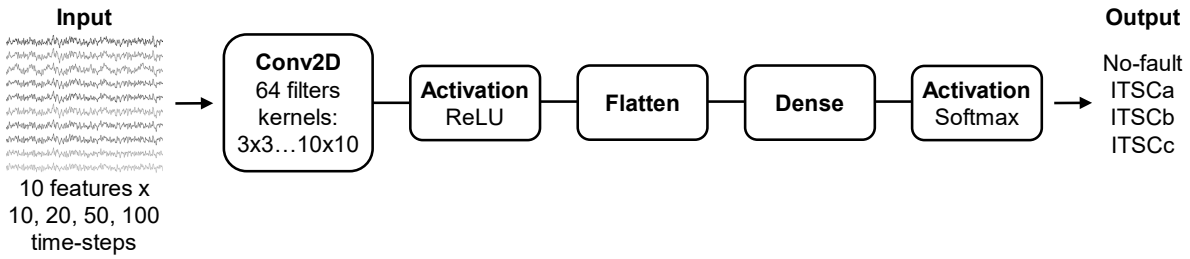


Figure G.2: The proposed shallow 2D CNN architecture

(Shapley Additive exPlanations), a method introduced in 2017 [19] to explain individual predictions of models on global and local level. On the global level it can show which features contribute to the model output and how significant their contribution is. On the local level it can examine each data point and investigate why the model made a certain decision.

G.4 Experiment and Results

The proposed method has been validated on the experimental setup used in [20], i.e. with a 4-pole PMSM whose parameters are given in Table. G.1. Each of the motor phase windings consists in two coils of 51 turns in series, with hence 102 turns per phase. As shown in Fig. G.3, ITSC faults have been applied on each phase by short-circuiting different number of turns resulting in a different fault percentage at each phase as shown in Table. G.2. The experiment lasted for 20s with sampling time for data acquisition of 50μs. ITSC faults in phase A, B and, C were applied in the time intervals $t = 4.471 - 7.238$s, $t = 9.613 - 12.760$s and, $t = 15.600 - 18.410$s respectively, see Fig. G.4. The 10 inputs of the model are shown in Fig. G.5 around the transition between the no fault region (green) and the ITSC fault in phase A (red) at t=4.471s. Samples of different

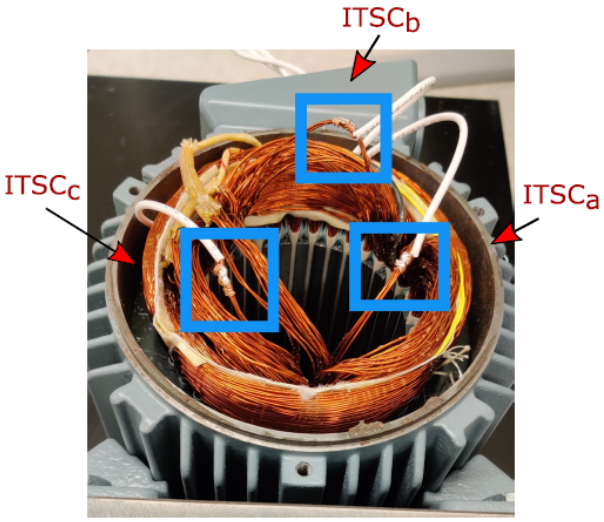


Figure G.3: Applied ITSC faults

Table G.2: Applied ITSC faults per phase

Fault type	Phase	Nr. of short-circuited turns	Applied ITSC fault in %	Nr. of records
ITSCa	A	1	0.89	55300
ITSCb	B	3	2.94	62800
ITSCc	C	5	4.90	56000
No-fault	—	0	0.00	185600

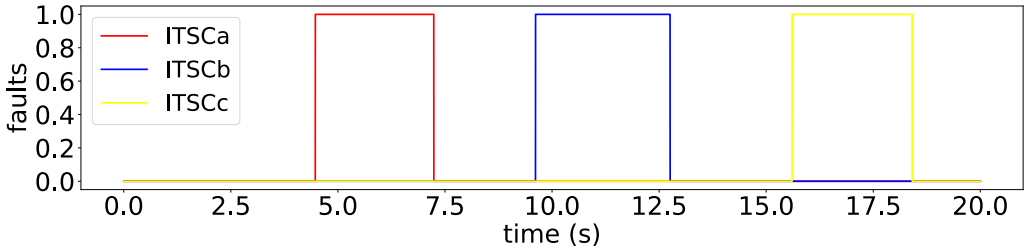


Figure G.4: Timeline of applied ITSC faults

Table G.3: Model attributes - approach 1

Sample length [time-steps]	No. of training samples	Batch size	Learning rate	Average training time [min]
10	25 177	32	0.001	6.14
20	12 588	32	0.001	3.48
50	5 035	32	0.001	2.08
100	2 517	32	0.001	1.37

lengths have been stacked one under the other making a 2D input of size 10 times sample length. The proposed method is evaluated using two approaches:

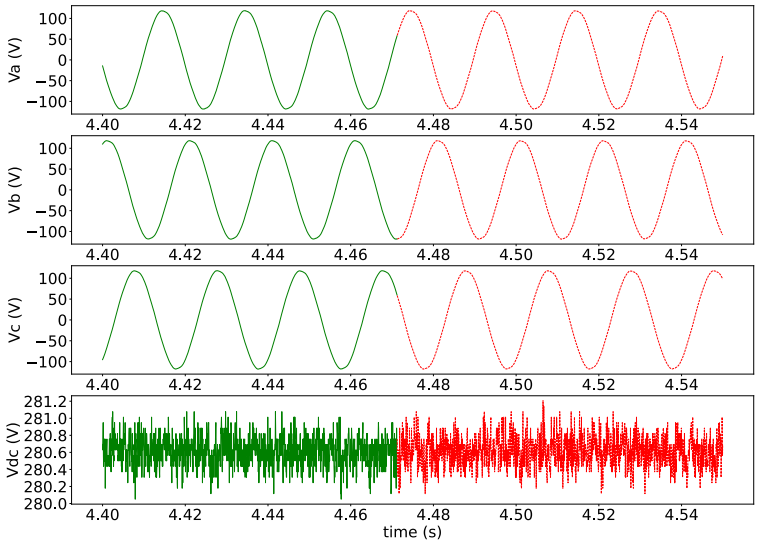
- Approach 1 - input data as separate non-overlapping samples.
- Approach 2 - input data as overlapping samples using sliding windows.

Sample sizes range over 10, 20, 50 and 100 time-steps corresponding to 0.5, 1, 2.5 and 5 ms. Each model is trained on 500 epochs using kernel sizes from 3x3 to 10x10 in order to achieve optimal results in terms of accuracy, simplicity, computational and energy efficiency. The number of filters has been set to 64. Adam optimizer with learning rate of 0.001 is used for all model configurations. Data has been divided into train and validation/test set in 70:30 ratio after random shuffling, resulting in train and validation/test sets being different for each training session. The final accuracies for the different model configurations have been determined as the average value of accuracies obtained after 30 trainings.

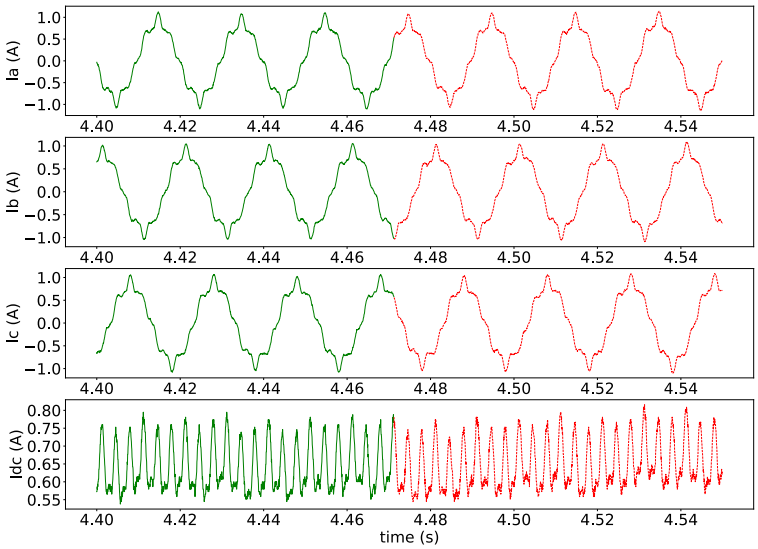
Trainings are performed on 4 NVIDIA Tesla V100 GPUs using Uber's horovod framework for distributed learning on TensorFlow. All available data has been used for training and testing, which results in slightly imbalanced classification due to data size in ratios of 52% (no fault), 15% (ITCSa), 17% (ITSCb), 16% (ITSCc).

G.4.1 Approach 1 - Non-overlapping Samples

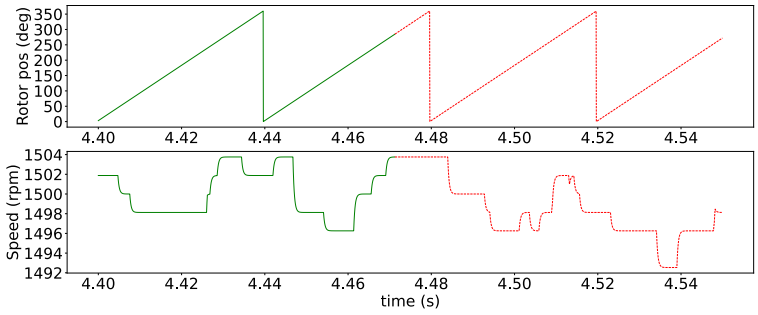
This subsection investigates what is the optimal length of the input samples. We start with the simplest approach, slicing the sequences into non-overlapping segments of 10, 20, 50 and 100 time-steps. By using 2D inputs into the convolutional network we expect from the model to find pattern between the different features sampled at the same time. In case of clear patterns, we expect that shorter lengths can deliver as good results as longer ones, potentially even better. The attributes of the model variations, together with the average time needed for training are given in Table G.3 while Table G.4 shows the corresponding validation accuracies. The max accuracies achieved for the best performing models are given in Fig. G.6 and Fig. G.7.



(a) Voltages

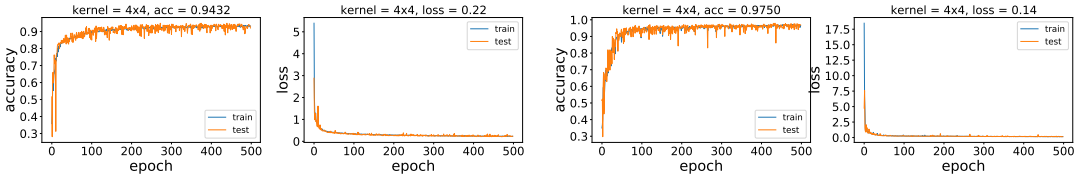


(b) Currents



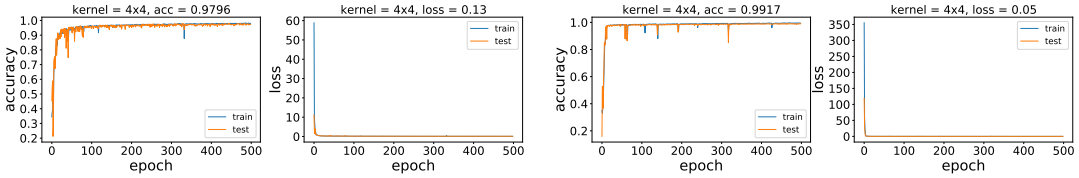
(c) Rotor position and speed

Figure G.5: 10 features - input to 2D CNN.



(a) Time-steps = 10

(b) Time-steps = 20



(c) Time-steps = 50

(d) Time-steps = 100

Figure G.6: Best performing models - max accuracies - appr. 1

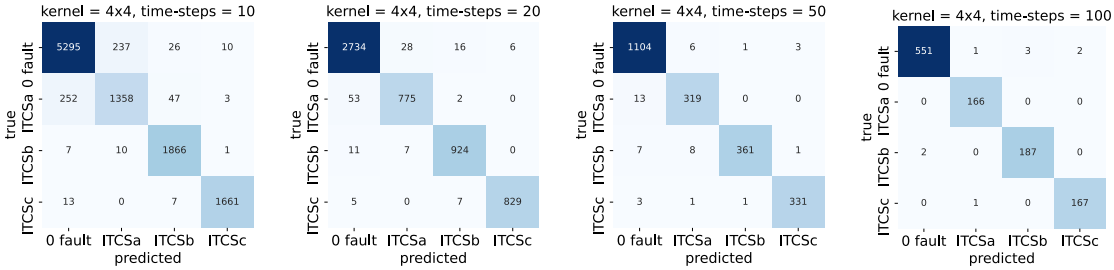
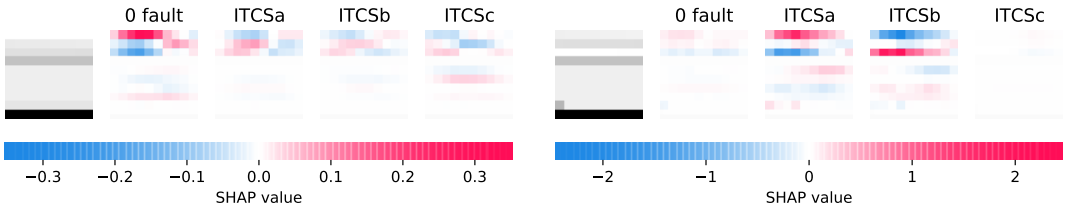
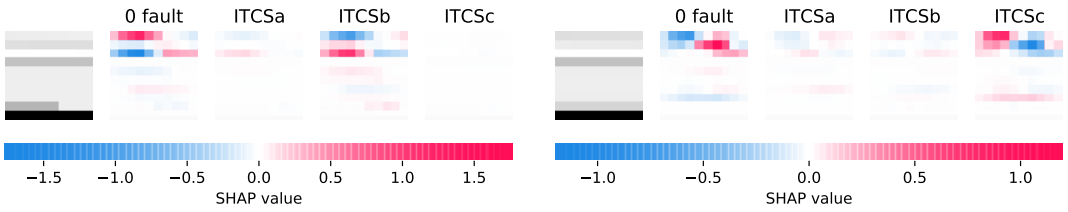


Figure G.7: Best performing models - confusion matrices - appr. 1



(a) No fault

(b) ITSCa fault



(c) ITSCb fault

(d) ITSCc fault

Figure G.8: SHAP values: kernel = 4x4, time-steps = 10, appr.1

Table G.4: Validation acc. [%] - approach 1

Kernel size	Sample length [time-steps]							
	10		20		50		100	
	μ	σ	μ	σ	μ	σ	μ	σ
3x3	91.00	1.81	94.92	1.76	94.95	2.46	97.91	0.90
4x4	92.46	1.23	96.33	0.95	96.74	0.61	98.42	0.72
5x5	88.16	2.68	94.49	2.53	93.73	4.64	97.32	1.54
6x6	88.26	2.45	94.18	2.82	94.49	2.45	97.04	1.85
7x7	87.81	2.27	94.25	3.74	95.41	0.93	96.36	3.57
8x8	88.00	1.56	94.60	1.12	94.82	2.57	96.60	2.80
9x9	86.68	1.44	92.79	2.15	93.84	2.23	86.11	11.59
10x10	67.20	12.42	61.13	14.73	66.29	15.11	58.96	10.41

G.4.1.1 Discussion

The proposed method results in high accuracies. The accuracy of the models shows general increase with the length of timesteps for all kernels, except for kernels 9x9 and 10x10 that show slight deviation. The best performing model is the one based on 100 time-steps and kernel size 4x4. It achieves an average accuracy of 98.42%. Fig. G.8 shows SHAP values corresponding to four different outcomes: no-fault, and three ITSC faults. As seen, the voltages in phases A, B and C play an important role together with the currents. However, the last two features (rotor position and speed) have a minimum or no impact on the results. Fig. G.8 shows the best performing model for time-steps length of 10, however the conclusions are valid for all models.

G.4.2 Approach 2 - Sliding Windows

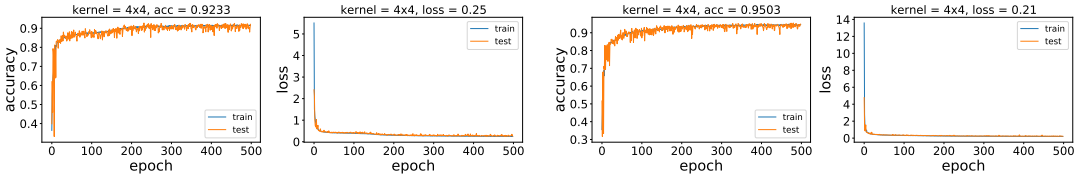
In this subsection we investigate whether we can get better results by using overlapping segments of 10, 20, 50 and 100 time-steps. We approximately double the number of train input samples and test whether introducing additional sequences of data gives more information. The model attributes and the validation accuracies are given in Table G.5 and Table G.6. The max accuracies achieved for the best performing models (kernel size 4x4) are given in Fig. G.9 and Fig. G.10.

G.4.2.1 Discussion

This approach shows similar results as approach 1, however it generally achieves slightly lower accuraccies for the same number of epochs (97.47% compared to 98.42%). The best performing model is again the one based on 100 time-steps and kernel size 4x4. Fig. G.11 shows SHAP values for four different outcomes with the same conclusions as earlier. The voltages and currents in phases A, B and C play an important role while the last two features contribute less.

Table G.5: Model attributes - approach 2

Sample length [time-steps]	No. of training samples	Batch size	Learning rate	Average training time [min]
10	50350	32	0.001	12.93
20	25171	32	0.001	7.03
50	10065	32	0.001	3.45
100	5028	32	0.001	2.31



(a) Time-steps = 10

(b) Time-steps = 20

(c) Time-steps = 50

(d) Time-steps = 100

Figure G.9: Best performing models - max accuracies - appr. 2

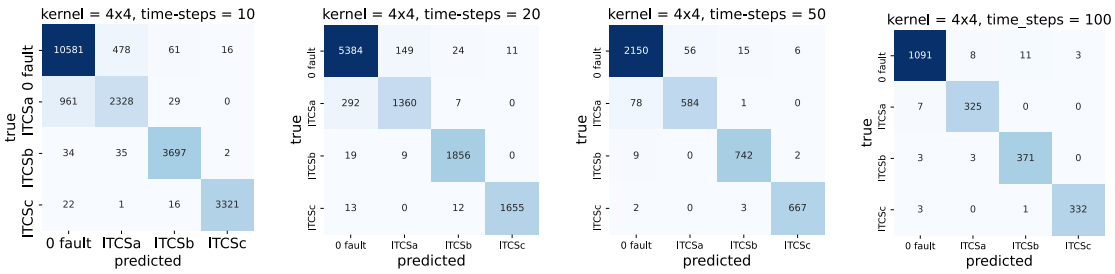
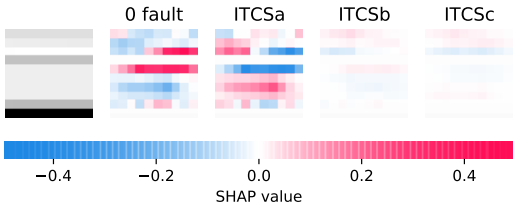


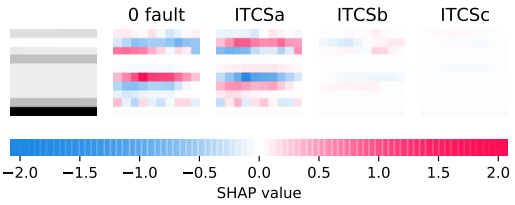
Figure G.10: Best performing models - confusion matrices - appr. 2

Table G.6: Validation acc. [%] - approach 2

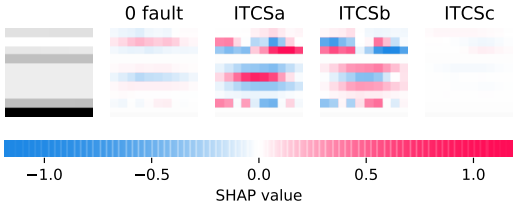
Kernel size	Sample length [time-steps]							
	10		20		50		100	
	μ	σ	μ	σ	μ	σ	μ	σ
3x3	89.89	1.42	92.43	1.02	89.52	2.12	95.75	0.89
4x4	90.92	0.90	93.80	0.86	93.83	1.54	97.47	0.51
5x5	86.78	1.72	88.93	2.77	86.96	2.48	92.64	5.87
6x6	87.40	1.67	88.67	2.68	87.76	2.23	95.07	1.74
7x7	87.52	2.13	90.42	0.87	89.64	1.65	94.44	3.10
8x8	87.34	1.79	90.20	1.67	90.32	1.58	95.28	1.24
9x9	87.17	1.74	87.98	1.06	87.77	2.63	88.99	7.20
10x10	67.15	12.33	56.60	10.14	51.61	0.00	58.90	11.16



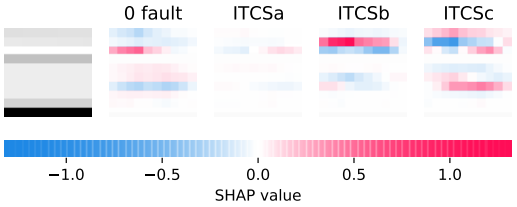
(a) No fault



(b) ITSCa fault



(c) ITSCb fault



(d) ITSCc fault

Figure G.11: SHAP values: kernel = 4x4, time-steps = 10, appr.2

G.4.3 Comparison with Model-based Approach

The main difference between data-driven (NNs) approach and signal- or model-based approach is the need of a priori understanding of the system. While both signal- and model-based approaches require a deep domain knowledge of the underlying system, data-driven approach discovers dependencies automatically. However, large amount of historical data for training the models, both healthy and faulty, is needed which is usually not available in such scale. Moreover, producing such data comes with high cost.

A model-based approach developed on the same underlying data [20] is used in this secton to allow for direct comparison. This method relies on structural analysis, where a dynamic mathematical model of the system is presented in matrix form, and where Dulmage-Mendelsohn decomposition tool has been used to extract small redundant parts and to design the error residuals further used for detection of the three ITSC faults

Table G.7: Best performing model - appr. 1 - performance metrics

	Binary classification	Multiclass classification			
Metrics	No-fault/fault	ITSCa	ITCSb	ITSCc	W. avg
Precision	0.9843	0.9722	0.9759	0.9900	0.9793
Sensitivity	0.9883	0.9960	0.9713	0.9839	0.9832
Specificity	0.9852	0.9948	0.9949	0.9982	0.9959
F1 score	0.9863	0.9839	0.9736	0.9870	0.9812
P_D	0.9883	0.9960	0.9713	0.9839	0.9832
P_{FA}	0.0148	0.0052	0.0051	0.0018	0.0041
Support	557/523	166	189	168	-

through a statistical test based on the generalized likelihood ratio test (GLRT). This approach has achieved detection rates (P_D) of 60.93% for ITSCa, 98.13% for ITSCb and 100% for ITSCc fault, given that the probability of false alarm (P_{FA}) has been set to 2%. It should be noted that this approach only detects the presence of the fault but does not distinguish among types of faults.

The achieved overall detection rate of the 2D CNN model presented in this paper is 98.83% when calculating on the best performing model. The overall and the detection rates for ITSCa, ITSCb and ITSCc faults, together with other performance metrics are shown in Table. G.7. The main limitation of the model is the need for sufficient amount of training data, especially faulty data that can be challenging to obtain outside of experimental setup.

G.5 Conclusions

This paper presented a straightforward method for detection and diagnosis of ITSC faults in PMSMs based on shallow 2D CNNs that compared to a model-based method showed a few advantages. The main advantage shown is the ability to deliver high accuracies without high calculation cost and without need for any feature pre-processing. In the future work, we intend to implement this type of approach to real-time monitoring of the motors located on an offshore rig. The input data is available, however not used and offered to customers as a service, mainly due to lack of a robust and easy to implement modeling. In addition, companies face a challenge during the official accreditation of the service due to the inability to explain the results of the model used. This challenge can be successfully faced with methods such as SHAP briefly outlined in this paper.

References

[1] Karolis Dambrauskas, Jonas Vanagas, Tomas Zimnickas, Artūras Kalvaitis, and Mindaugas Ažubalis. A method for efficiency determination of permanent magnet synchronous motor. *Energies*, 13(4):1004, 2020.

[2] RN Bell, CR Heising, P O'donnell, C Singh, and SJ Wells. Report of large motor reliability survey of industrial and commercial installations. ii. *IEEE Transactions on Industry applications*, 21(4):865–872, 1985.

[3] Yuan Qi, Emine Bostanci, Mohsen Zafarani, and Bilal Akin. Severity estimation of interturn short circuit fault for pmsm. *IEEE Transactions on Industrial Electronics*, 66(9):7260–7269, 2018.

[4] Austin H Bonnett and George C Soukup. Cause and analysis of stator and rotor failures in three-phase squirrel-cage induction motors. *IEEE Transactions on Industry applications*, 28(4):921–937, 1992.

[5] Eva García-Martín, Crefeda Faviola Rodrigues, Graham Riley, and Håkan Grahn. Estimation of energy consumption in machine learning. *Journal of Parallel and Distributed Computing*, 134:75–88, 2019.

[6] Yong Chen, Siyuan Liang, Wanfu Li, Hong Liang, and Chengdong Wang. Faults and diagnosis methods of permanent magnet synchronous motors: A review. *Applied Sciences*, 9(10):2116, 2019.

[7] Andrew KS Jardine, Daming Lin, and Dragan Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, 20(7):1483–1510, 2006.

[8] Sven Myrdahl Opalic, Morten Goodwin, Lei Jiao, Henrik Kofoed Nielsen, Ángel Álvarez Pardiñas, Armin Hafner, and Mohan Lal Kolhe. Ann modelling of co2 refrigerant cooling system cop in a smart warehouse. *Journal of Cleaner Production*, 260:120887, 2020.

[9] Rui Zhao, Ruqiang Yan, Zhenghua Chen, Kezhi Mao, Peng Wang, and Robert X Gao. Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, 115:213–237, 2019.

[10] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

[11] Xiaoli Zhao, Minping Jia, and Zheng Liu. Semisupervised graph convolution deep belief network for fault diagnosis of electormechanical system with limited labeled data. *IEEE Transactions on Industrial Informatics*, 17(8):5450–5460, 2020.

[12] Sen Wang, Jieqiu Bao, Siyang Li, Hongkui Yan, Tianyao Tang, and Di Tang. Research on interturn short circuit fault identification method of pmsm based on deep learning. In *2019 22nd International Conference on Electrical Machines and Systems (ICEMS)*, pages 1–4. IEEE, 2019.

[13] Yuanjiang Li, Yanbo Wang, Yi Zhang, and Jinglin Zhang. Diagnosis of inter-turn short circuit of permanent magnet synchronous motor based on deep learning and small fault samples. *Neurocomputing*, 442:348–358, 2021.

[14] Fatima Husari and Jeevanand Seshadrinath. Sensitive inter-tum fault identifcation in induction motors using deep learning based methods. In *2020 IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020)*, pages 1–6. IEEE, 2020.

[15] Jinyang Jiao, Ming Zhao, Jing Lin, and Kaixuan Liang. A comprehensive review on convolutional neural network in machine fault diagnosis. *Neurocomputing*, 417:36–63, 2020.

[16] Turker Ince, Serkan Kiranyaz, Levent Eren, Murat Askar, and Moncef Gabbouj. Real-time motor fault detection by 1-d convolutional neural networks. *IEEE Transactions on Industrial Electronics*, 63(11):7067–7075, 2016.

[17] I-Hsi Kao, Wei-Jen Wang, Yi-Horng Lai, and Jau-Woei Perng. Analysis of permanent magnet synchronous motor fault diagnosis based on learning. *IEEE Transactions on Instrumentation and Measurement*, 68(2):310–324, 2018.

[18] Long Wen, Xinyu Li, Liang Gao, and Yuyan Zhang. A new convolutional neural network-based data-driven fault diagnosis method. *IEEE Transactions on Industrial Electronics*, 65(7):5990–5998, 2017.

[19] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.

[20] Saeed Hasan Ebrahimi, Martin Choux, and Van Khang Huynh. Real-time detection of incipient inter-turn short circuit and sensor faults in permanent magnet synchronous motor drives based on generalized likelihood ratio test and structural analysis. *Sensors*, 22(9):3407, 2022.