FPGA-based Degradation Evaluation for Traction Power Module with Deep Recurrent Autoencoder

Shuai Zhao[1], Jiahong Liu[1], Kaiqi Chu[2], Shujia Mu[2], Huai Wang[1]

[1]*AAU Energy, Aalborg University*
Aalborg 9220, Denmark
{szh, jliu, hwa}@energy.aau.dk

[2]*CRRC Qingdao Sifang Co., Ltd.*
Qingdao 266111, China
{chu762418459, mushujia8247}@163.com

***Abstract*—The timely and quantitative evaluation of the degradation is crucial for traction inverter systems in railway applications. The implementation in the industry is impeded by two major challenges including the varying operational profiles and the scalability for system-level applications. This paper proposes a deep recurrent autoencoder-based degradation evaluation method, to assess the degradation level of the traction power module online. The recurrent structure is embedded for processing multivariate time series condition monitoring data stream, in order to exploit the inherent time dependence to improve the accuracy and robustness. The autoencoder-based framework enables the scalability of the proposed method to system-level applications and can be applied under varying operating conditions. The method is experimentally demonstrated on an FPGA-based hardware platform.**

***Index Terms*—Power module, inverter system, deep learning, autoencoder, degradation, prognostics and health management**

I. Introduction

Predictive maintenance of traction inverter systems is crucial to the safe operation of high-speed railways. An industry survey [1] indicates that the power semiconductor is among the most fragile components. It demands the timely detection and accurate quantification of the degradation status of the power module in the railway traction inverter, as a basis for safety-oriented decision making [2].

Extensive studies have been devoted to this area that covers for example the physics-of-failure analysis, degradation precursor determination, sensory circuit design, degradation evaluation, etc. Among them, most of the existing efforts focus on the physics of the failure analysis of power devices and modules [3], identifying candidate degradation precursors [4], followed by the design of the hardware circuit to collect the degradation precursors in real-time [5]. With increasing data availability of the degradation precursors, how to evaluate the aging effects via degradation evaluation is the next demanding but very daunting task.

As illustrated in Fig. 1, along with traction inverter operation, the *in-situ* condition monitoring system will collect multiple sensor channels consisting of degradation precursors and operation profiles, resulting in high-speed data streams. Essentially, the specific pattern, e.g., the increasing trend, of the degradation precursors, can be applied to access the inverter degradation, while the operation profiles record the inverter missioning modes. The variation of the degradation precursors can be essentially attributed to two factors including 1) inverter aging effects leading to the degradation; and 2) the variation of external operational profiles. It is worth mentioning that the degradation precursors are coupled with the operational profiles via complex nonlinear and time-varying relationships. For example, in condition monitoring of MOSFETs [6], the degradation precursor drain-source on-state resistance $R_{\mathrm{ds(on)}}$ highly depends on the operational profile device case temperature T_{c}. For degradation evaluation, however, only the variation of the degradation precursors as a result of the aging effects is of interest. Formally, the objective for degradation evaluation is to quantitatively isolate the aging effects in the degradation precursors, from the operational profiles.

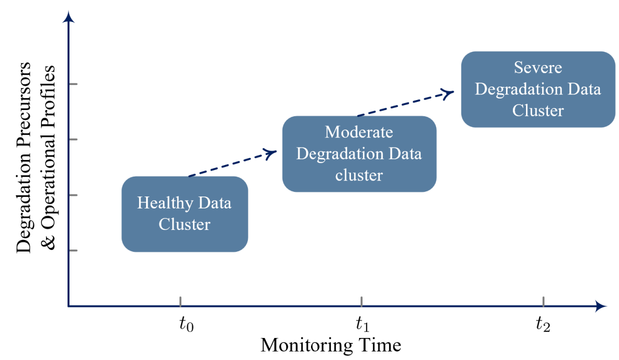

Fig. 1. Condition monitoring schemes for the degradation evaluation of the inverter systems. The variations among the health data cluster at t_0, the moderate degradation data cluster at t_1, and the severe degradation data cluster at t_2 are due to both the aging and the operational profiles.

To this end, one direction is to separate the operational profiles from the degradation precursors separately, leaving the variations due to degradation solely. However, the nonlinear and time-varying between these two factors are very difficult to characterize for decoupling purposes explicitly. Another direction is to evaluate the variation of the degradation precursor, under the *same operational profiles*. Since the operational profiles are identical at different evaluation time epochs, the effects of operational profiles on the degradation precursors can be suppressed, leaving only variations due to the aging effects for the degradation evaluation. In this way, variations on the degradation precursors due to aging can be accessed virtually and indirectly, which are equivalently independent of the operational profiles.

For the selection of the same operational profiles, there are three options including

1) Find an specific working point in the operational profiles that the degradation precursor is invariant. As a result, at different monitoring epochs, the degradation precursors at this specific working point are utilized for comparison so that we can investigate the aging effect only. As an example, for the degradation detection of insulated gate bipolar transistor (IGBT) package in [7], an inflection point at which the precursor on-state collector-emitter voltage $v_{\mathrm{ce(on)}}$ is irrespective of the junction temperature, is used for condition monitoring purposes. Similar example can be found in [8] for condition monitoring of capacitor, where elaborated transformation is performed on the degradation precursors to make them invariant to the operational profiles. Note that this technological method requires expert experience and experimental evidence to identify this invariant working point.
2) Identify a specific working point at which the variations of the degradation precursors due to the aging effects is more evident and sensitive. When accessing the degradation, only executing the degradation evaluation procedure when the data at these specific operational points are available. However, the identified most evident and sensitive working points is complicate and case-by-case, which is challenging to be generalized. Moreover, when the number of the degradation precursors is relatively large, it will be difficult to determine the most evident and sensitive working points, since each degradation precursor may resulting different working points. Therefore, it will be difficult to apply to the system-level cases.
3) From the data-driven perspective, the degradation precursors and the comprehensive operational profiles are collected. As a result, it assumes that the overall operational profiles are available. The next step is evaluating the variation effects of the degradation precursors, in the presence of the same characterisitics and distribution of the operational profiles at any condition monitoring times. From a system-level perspective, this technological method can be well generalized to complex cases without extensive expert experience. The paid cost, however, is that the data collection part should be well-designed to cover the comprehensive operational profiles as much as possible. For system-level applications, this method is preferred since the method can be generalized well. The above limitations on the data collection requirement, from the industrial implementation perspective, can be well fulfilled since the degradation of power electronic systems is very slow. There is enough time for the data logging system to collect the data at comprehensive operational profiles.

For the degradation evaluations for the traction inverter system, another key feature is the fast system dynamics that require high-speed data collection, resulting in a large volume of condition monitoring streams. It makes the determination of the specific invariant or the most sensitive operational profile even more challenging. As a result, considering the generalization capability for dealing with multiple degradation precursors of the degradation evaluation methods, the endeavors of this paper focus on developing the degradation evaluation method from the data-driven perspective. For this paradigm, exemplary cases include a Principal Component Analysis (PCA)-based degradation detection of SiC MOSFETs [9], a Mahalanobis distance-based technique identifying the anomalous behavior in non-punch through (NPT) and field stop (FS) IGBTs [10], etc. However, these conventional statistical methods such as PCA-based methods and Mahalanobis distance-based methods, are based on the monitoring information only at the current time instance, which is less robust and results in a high percentage of false alarms, especially considering that the monitoring precursors from the traction inverter are very noisy.

To address the above challenges and ease field implementation, we resort to the reconstruction-based scheme for degradation detection, which is facilitated by the deep learning tool of autoencoder [11]. Specifically, in this paper, we propose a new deep recurrent autoencoder-based degradation evaluation method for traction power modules. By exploiting the time dependence among the condition monitoring data streams, the proposed framework has industry-favorable features including unnecessary specific operational profile selection and robustness. Moreover, the method is irrelevant to the varying operating conditions and has the potential to be scalable to system-level applications with a large number of degradation precursors and complex operational profiles. The method implementation on a Field Programmable Gate Array (FPGA) platform is investigated for demonstration.

The remainder of the paper is organized as follows. Section II presents the problem setting and the proposed method, illustrated with the field dataset of the traction inverter system. Section II-A demonstrates the method by using the implementation on an FPGA hardware. Finally, the findings and insights are summarized in Section IV.

II. Methodology

A. Degradation Emulation for Traction Power Module

An experimental setup of traction inverter systems in railway applications is developed for the data collection. Note that the degradation of the traction inverter in the field is very slow, which is difficult to collect field degradation data. Instead, for illustration, the traction inverter is operated with two modes for emulating the healthy and degraded behavior of the power module. Specifically, when the inverter is normally operated, the cooling fan of the power module is working

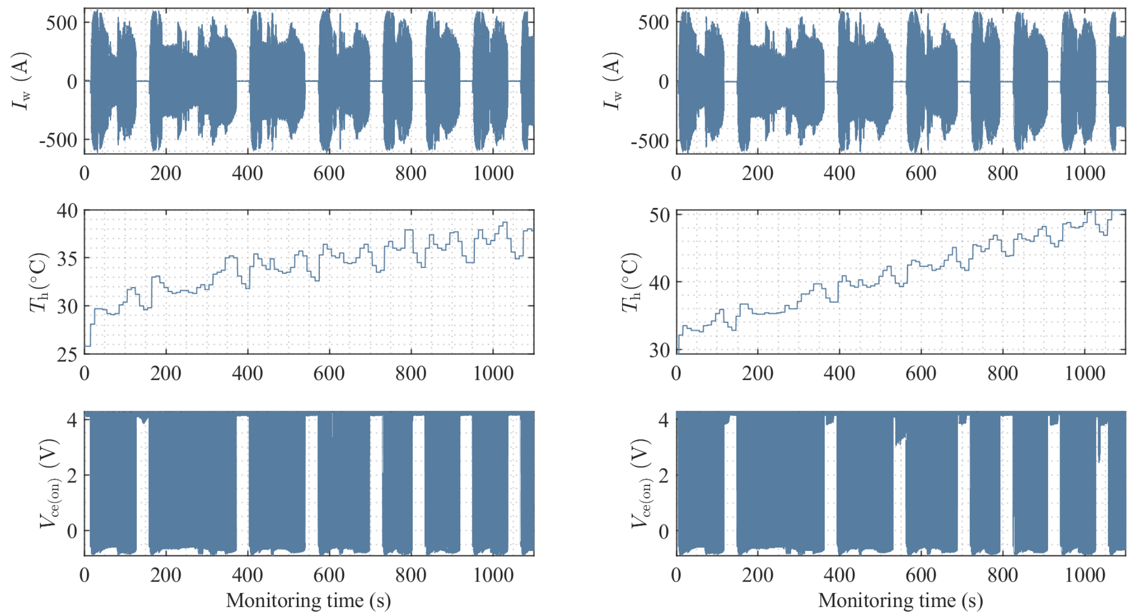

Fig. 2. Data collection for traction power module with different status of the cooling fan. (a) Healthy dataset when cooling the fan at a high speed; and (b) Degraded dataset when the cooling fan stops working.

at a high speed. The dataset collected in this condition is considered a healthy dataset. To emulate the degraded behavior of the power module, the cooling fan is manually stopped without any cooling wind flow through the heatsink. The different operational conditions of the cooling fan can be used to emulate the degradation due to the thermal path deteriorating. As a result, the dataset collected in this case is considered the degraded dataset. In these two cases, the operational profiles including phase current I_w, and the degradation precursors including the heatsink temperature T_h and the on-state collected-emitter voltage $v_{\mathrm{ce(on)}}$ of the IGBT module, are collected, as shown in Fig. 2. There is an evident difference in the heatsink temperature T_h for these two cases. Note that the operational profiles are as diversified as possible so that they can cover all of the possible cases in the field. For the degradation evaluation, the primary task is to evaluate the data cluster difference quantitatively between the healthy and the degraded datasets of the power module in the traction inverter. Mathematically, the collected datasets consist of multi-channel information resulting in multivariate time series signals.

Fig. 3 shows the healthy and degraded datasets from a point-wise perspective. It can be clearly seen, however, that a large proportion of these two datasets are extensively mixed, suggesting that the healthy and the degraded datasets are challenging to be separated from the point-wise point. Therefore, the overall difference cannot be evidently quantified. To address this challenge, the inherent time-dependence information in the monitoring precursors is exploited, in addition to the features in the space coordinate $[T_\mathrm{h}, I_\mathrm{w}, v_{\mathrm{ce(on)}}]$. Theoretically, these multivariate time series data are from a dynamic system that will generate and output system states progressively along with time. As a result, the former and the latter time steps are highly dependent on each other. Therefore, incorporating this time-dependence information will increase the probability to separate these two datasets quantitatively, and the following analysis confirms this assumption that such a time-dependence mechanism can significantly improve the accuracy and robustness of degradation evaluation. In this case, the multivariate time series $\mathbf{X}_k = [T_\mathrm{h}, I_\mathrm{w}, v_{\mathrm{ce(on)}}]_k^{(k+s)}$ is considered as the inputs rather than just the single vector $[T_\mathrm{h}, I_\mathrm{w}, v_{\mathrm{ce(on)}}]_k$ values at any particular time instance t_k, where $s = 20$ is the sequence length.

B. Degradation Evaluation with Deep Recurrent Autoencoder

The proposed method applies the reconstruction scheme [11], [12] for the degradation evaluation with the inputs of multivariate time series, by using a deep recurrent autoencoder. Given the multivariate time series, the objective is to acquire a low-dimensional feature representation space that effectively captures the essential characteristics of the given data streams. This technique is commonly employed for data compression and dimension reduction purposes. In the context of degradation evaluation, the rationale behind utilizing this approach is that the learned feature representations are constrained to capture significant regularities of the data streams, resulting in minimal reconstruction errors. Consequently, degraded datasets, being inherently dissimilar from normal ones, exhibit substantial reconstruction errors, making them difficult to accurately reconstruct from the derived representations. Summarizing, it is assumed that normal multivariate time series

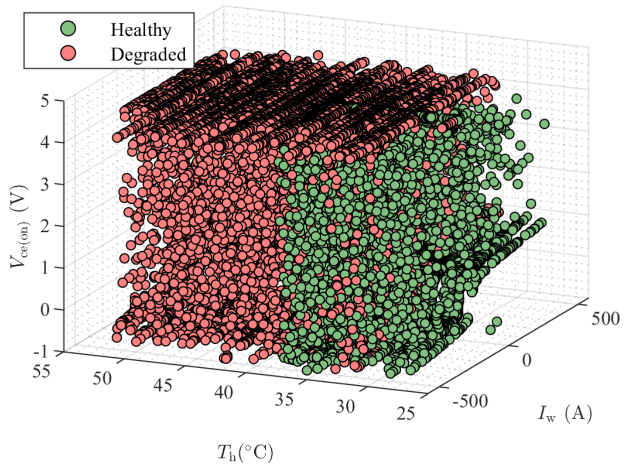

Fig. 3. Point-wise illustration for the healthy and degraded datasets. Note that most of the data points from these two categories are extensively overlapped.

can be more effectively reconstructed from the compressed space compared to degraded data streams.

For the implementation, the deep autoencoder is designed to fulfill this purpose. It consists of the encoder part and the decoder part. The encoder part will transform the multivariate time series into a compact representation, i.e., the latent space, while the decoder will restore the signals. Due to the reduced expressive capability of the encoder part, the architecture will enforce the network to learn the most relevant information in the data and suppress the irrelevant ones such as noise. Note that the autoencoder is an unsupervised learning method [13], suggesting only the healthy dataset is required for the method training and is more applicable to field applications.

However, the encoder and decoder of the conventional autoencoder are implemented with the vanilla neural network, which has no capability to ingest the sequential information in the datasets. Deep recurrent autoencoders have emerged as a powerful tool for the representation learning tasks in time series modeling domains. The key ability of this powerful data-driven tool is to capture temporal dependencies and preserve sequential information. By employing recurrent architecture in the encoder and decoder networks, it can effectively model time series data and learn compact representations with temporal information. In this case, the encoder and decoder of the model are implemented with recurrent structure, i.e., the Long Short-Term Memory (LSTM), in particular. Fig. 4 shows the architecture of the deep recurrent autoencoder. It is worth mentioning that one of the key advantages of this technological approach is that there is no limitation on the number of information channels. It can be easily scalable to system-level applications without extensive manual configuration.

For degradation evaluation in this paper, this deep recurrent autoencoder is applied to establish the healthy model of the traction power module. Assume that the model is trained by using the inverter healthy data only, it therefore can reconstruct these healthy data streams in the outputs. However, since the model cannot characterize the inverter behavior in other states,

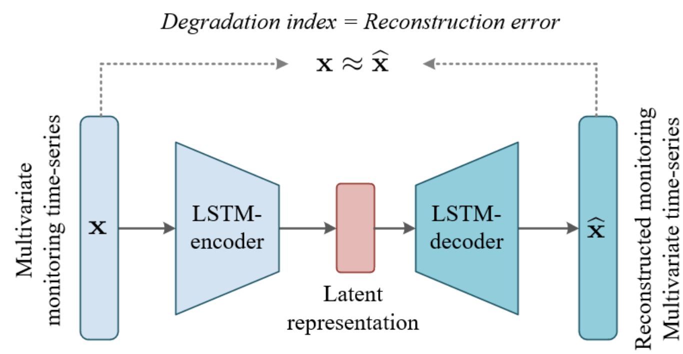

Fig. 4. Deep recurrent autoencoder for degradation evaluation with the condition monitoring of multivariate time series for the traction power module. The reconstruction error is calculated with the Mean Absolute Error (MAE).

e.g., the different levels of degraded states, it will not be able to reconstruct the datasets. Therefore, when the data streams of the degraded power module are fed to the established model, the model will reconstruct the signals with a large reconstruction error. The more degraded status, the higher the reconstruction error. In this way, the reconstruction error can be applied as an indicator to illustrate the degradation status of the traction power module quantitatively.

C. Statistical Tool for Suppressing Operational Profile Variations

Another challenge for field degradation evaluation is that the calculated degradation index is highly sensitive to the operational profiles. To mitigate the affecting factor of operational profiles, one feasible approach is to compare the data characteristic difference in the presence of all possible operational profiles. It is noted that collecting comprehensive operational profiles is feasible at a specific condition monitoring epoch since the degradation is very slow. To evaluate the overall variation effect, the Box-plot statistical tool [6] is applied to obtain the median value of the overall variation. Since the data characteristics of the operational profiles are almost the same at different evaluation time instances, the variations due to the operational profiles can be suppressed.

Specifically, the reconstruction errors of a large number of multivariate time series $[T_{\mathrm{h}}, I_{\mathrm{w}}, v_{\mathrm{ce(on)}}]_k^{(k+s)}$ are applied to quantify the degradation levels of the traction power module. The median of the reconstruction errors, obtained by using the Box-plot analysis, is applied as the degradation index. Regarding the healthy dataset and the degraded dataset as in Fig. 2, the histogram of the reconstruction errors is shown in Fig. 5. It can be seen that there is a clear gap between the histograms based on the healthy and the degraded datasets, suggesting an acceptable limit can be applied to differentiate these two cases. Once the reconstruction errors are obtained, the box-plot tool is applied to calculate the statistical median of the reconstruction errors, as the overall degradation index of the traction power module. As a result, based on the box-plot analysis, the medians of these reconstruction errors are applied as a quantitative index for assessing the degradation levels,

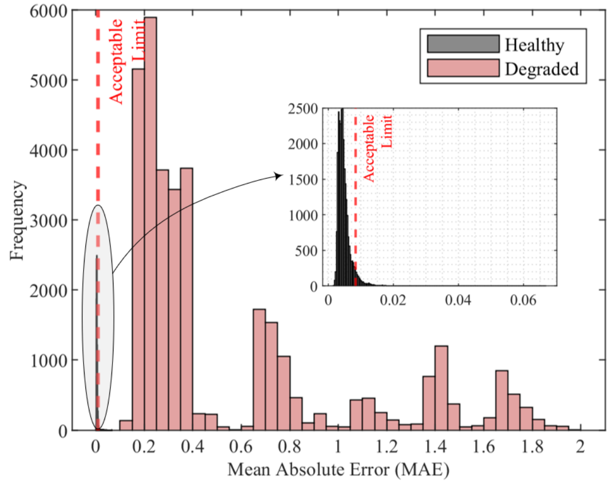

Fig. 5. The histogram of the reconstruction errors for the healthy and degraded cases.

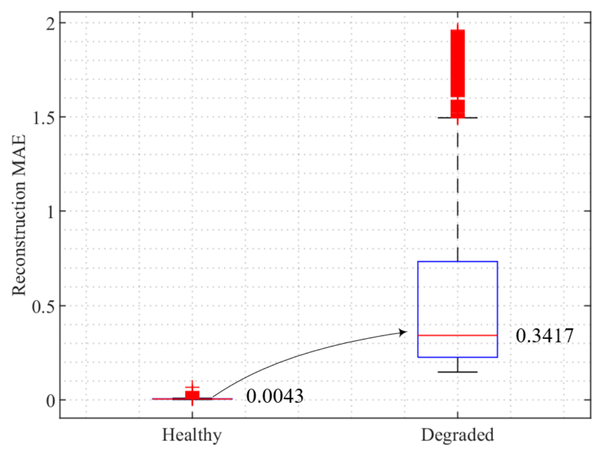

Fig. 6. The median statistics using the Box-plot for quantifying the reconstruction errors for different degradation levels.

which are 0.0043 and 0.3417 for the healthy and degraded cases, respectively, as shown in Fig. 6. It can be seen that the difference between the healthy and the degraded cases can be evidently quantified.

It is worth mentioning that one of the key requirements of the proposed method is the completeness of the operational profiles, which should be as diversified as possible and can cover most of the representative operational profiles in the field. This assumption can be feasibly fulfilled in field applications. For example, the healthy dataset can be collected in the field for several months when a new traction inverter is installed for commissioning. In this period, the degradation of the inverter at the very beginning of the service life cycle is negligible. During these months, the operational profiles that the system is subjected to will be as comprehensive as possible, which can cover all the possible cases in the following operations. The above data collection procedure can be iteratively performed when requesting the degradation evaluation result. Although the data collection is a little bit time-consuming, considering the degradation of power electronics is very slow in practice, the above procedures can be readily implemented in the field.

III. Hardware Implementation with FPGA

A. Algorithm Modules and FPGA Architecture

The method implementation on an edge platform is key for industrial applications. The design of the degradation evaluation system endeavors to attain low complexity and acceptable utilization of resources in the context of resource-limited edge implementation. In this paper, the proposed method is implemented on an FPGA platform (Xilinx Zynq ZC706). It is based on the Zynq System-on-Chip (SoC) platform of Xilinx, which integrates a Dual-core ARM Cortex-A9 processor system (PS) and a programmable logic (PL) unit of FPGA within a single chip. The combination of the PL and PS enables the partitioning of the functionalities into time-critical components that are mapped onto the FPGA, freeing the processor to handle less critical and potentially more complex functions. Implementation of both critical and non-critical functions is accomplished through hardware description language (HDL) development in the PL. While developing the proposed architecture on the designated FPGA development board, specific hardware blocks within the PL, such as distributed memory, arithmetic units, logic, and various interconnections, are utilized.

As the degradation of a well-designed traction inverter system is a gradual process over the normal operating year, degradation evaluation is not considered a regular task that requires frequent execution. Hence, the computing speed of the designed architecture is not a significant performance metric, and constraints on speed can be relaxed to facilitate implementation simplicity. The architecture of the implemented hardware is shown in Fig. 7. It has been designed to be scalable for deep learning tool implementation in terms of the number of network layers, hidden units, and input dimension if sufficient hardware resources are available. The system framework performs four primary functions including the initialization of the neural network, normalization of the input data, neural network inference, and calculation of the MAE value for the reconstruction error in terms of the degradation index. These functions were implemented using Verilog hardware language in the programmable logic PL section of the hardware, with a focus on implementing both critical and noncritical functions efficiently.

The implementation of a deep learning tool (e.g., neural network) involves a significant number of weights and bias parameters, making manual coding a challenging and time-consuming task. To overcome these difficulties, the MATLAB Deep Learning HDL Toolbox was employed to construct a deep learning processor optimized for FPGA implementation. The data type employed is in accordance with the IEEE-754 single-precision floating-point standard. With the Matlab Deep Learning HDL Toolbox, it allows to generate the synthesizable HDL code from deep learning models directly, which can then

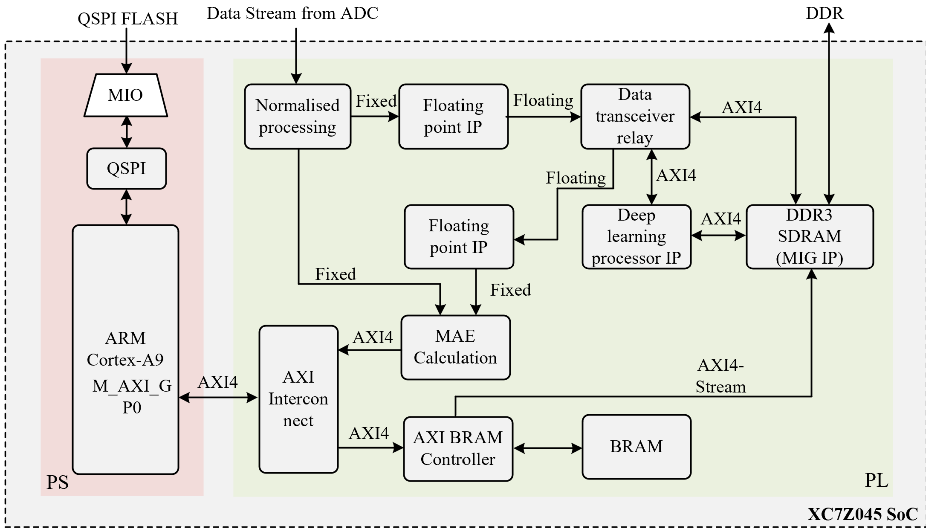

Fig. 7. The architecture of the FPGA hardware and the implementation of the deep recurrent autoencoder-based degradation evaluation. The autoencoder inference is implemented in the deep learning processor IP module and the degradation index calculation is implemented in the MAE calculation module. More details can be found in [14].

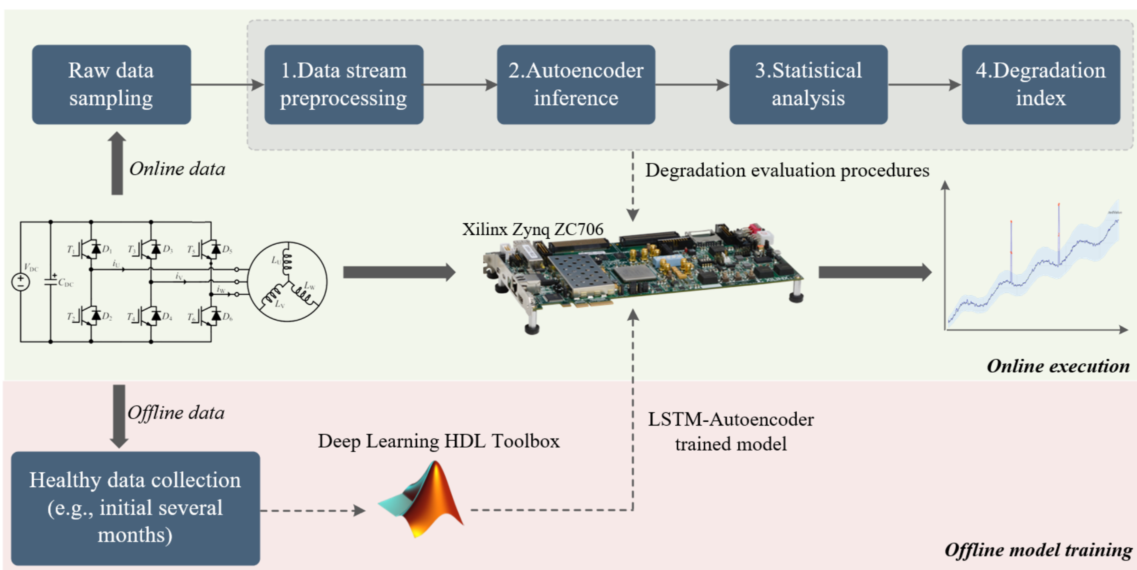

Fig. 8. The implementation pipeline of the degradation evaluation of the power module in the traction inverter system with FPGA hardware.

be integrated into digital circuits for efficient execution on FPGAs and other hardware platforms. It offers a more focused approach to implementing deep neural networks compared to the HDL Coder Toolbox. Moreover, the former emphasizes the implementation of crucial components for deep learning tools such as processing modules, memory access arbitrator modules, a top-level scheduler module, and profiler and debugger utilities. Another key advantage of the Deep Learning HDL Toolbox is its ability to adapt to the size and scale of the deep learning models being implemented. This adaptation can result in reduced hardware resource usage by optimizing the input and output memory size, disabling convolutional processing modules, adjusting the number of fully connected layer threads, etc.

B. Implementation Pipeline

The entire pipeline and procedures are illustrated in Fig. 8. The implementation of the proposed degradation evaluation method consists of the offline model training part and the online execution for the degradation evaluation part. For the offline model training, the new traction power module system will be monitored for data collection for several months, so that to collect the data at varying operating conditions as diversified as possible. These data are considered as the healthy dataset, which will be used to train the deep recurrent

TABLE I
RESOURCE UTILIZATION OF OVERALL ARCHITECTURE OF FPGA (XC7Z045 FFG900)

Resource	Utilization	Available	Percentage
LUT	95843	218600	43.84%
LUTRAM	9658	70400	13.72%
Flip-Flops	105727	437200	24.18%
BRAM	51.5	545	9.45%
DSP	73	900	8.11%

autoencoder offline and then embedded on the FPGA with the MATLAB Deep Learning HDL Toolbox.

For the online execution for calculating the degradation index part, by feeding the raw multivariate time series data stream $\mathbf{X}_k = [T_{\mathrm{h}}, I_{\mathrm{w}}, v_{\mathrm{ce(on)}}]_k^{(k+s)}$ into the pipeline, the steps of 1) data stream preprocessing, 2) autoencoder inference, 3) statistical analysis tool, and 4) degradation index calculation are implemented on FPGA and will be subsequently executed in real-time. In this way, the degradation level of the inverter system can be continuously evaluated and monitored. The details of resource utilization of FPGA are given in Table I. It can be seen that the method occupies less than 50% of the total resources, which can be further reduced with computation-light measures such as network quantizing and pruning [15].

One of the key advantages of the proposed method is that it is a tailored degradation detection solution for a specific inverter system. Therefore, it is unnecessary for transferring the trained model to the specific system. The paid cost, however, is that there is no detection capability when collecting healthy data at the very beginning of several months. It is acceptable in practice since the new system degradation at the beginning stage would be negligible.

IV. CONCLUSIONS

A deep recurrent autoencoder-based degradation detection method is proposed for the traction power module in this paper. The multivariate time series of the traction power module in a healthy state is applied to train a deep recurrent autoencoder model, formulating the healthy behavior model of the traction inverter system. The reconstruction error is applied as a quantitative index for degradation evaluation. The entire data-driven pipeline is implemented and verified on FPGA-based hardware. It demonstrates that the deep recurrent autoencoder-based framework holds great potential to address the challenges of scalability and affecting operational profiles in the degradation evaluation for power electronic systems.

REFERENCES

[1] J. Falck, C. Felgemacher, A. Rojko, M. Liserre, and P. Zacharias, "Reliability of power electronic systems: An industry perspective," *IEEE Ind. Electron. Mag.*, vol. 12, no. 2, pp. 24–35, 2018.

[2] S. Zhao, F. Blaabjerg, and H. Wang, "An overview of artificial intelligence applications for power electronics," *IEEE Trans. Power Electron.*, vol. 36, no. 4, pp. 4633–4658, 2021.

[3] E. Ugur, F. Yang, S. Pu, S. Zhao, and B. Akin, "Degradation assessment and precursor identification for SiC MOSFETs under high temp cycling," *IEEE Trans. Ind. Appl.*, vol. 55, no. 3, pp. 2858–2867, 2019.

[4] Z. Ni, X. Lyu, O. P. Yadav, B. N. Singh, S. Zheng, and D. Cao, "Overview of real-time lifetime prediction and extension for SiC power converters," *IEEE Trans. Power Electron.*, vol. 35, no. 8, pp. 7765–7794, 2020.

[5] Y. Peng, Y. Shen, and H. Wang, "A converter-level on-state voltage measurement method for power semiconductor devices," *IEEE Trans. Power Electron.*, vol. 36, no. 2, pp. 1220–1224, 2021.

[6] Y. Z. Peng, S. Zhao, and H. Wang, "A digital twin based estimation method for health indicators of dc-dc converters," *IEEE Trans. Power Electron.*, vol. 36, no. 2, pp. 2105–2118, 2021.

[7] A. Singh, A. Anurag, and S. Anand, "Evaluation of V_{ce} at inflection point for monitoring bond wire degradation in discrete packaged igbts," *IEEE Trans. Power Electron.*, vol. 32, no. 4, pp. 2481–2484, 2017.

[8] M. Rigamonti, P. Baraldi, E. Zio, D. Astigarraga, and A. Galarza, "Particle filter-based prognostics for an electrolytic capacitor working in variable operating conditions," *IEEE Trans. Power Electron.*, vol. 31, no. 2, pp. 1567–1575, 2016.

[9] W. Chen, L. Zhang, K. Pattipati, A. M. Bazzi, S. Joshi, and E. M. Dede, "Data-driven approach for fault prognosis of SiC MOSFETs," *IEEE Trans. Power Electron.*, vol. 35, no. 4, pp. 4048–4062, 2020.

[10] N. Patil, D. Das, and M. Pecht, "A prognostic approach for non-punch through and field stop IGBTs," *Microelectron. Reliab.*, vol. 52, no. 3, pp. 482–488, 2012.

[11] G. S. Pang, C. H. Shen, L. B. Cao, and A. Van den Hengel, "Deep learning for anomaly detection: A review," *ACM Comput. Surv.*, vol. 54, no. 2, pp. 1–38, 2021.

[12] J. Wen and H. Gao, "Degradation assessment for the ball screw with variational autoencoder and kernel density estimation," *Adv. Mech. Eng.*, vol. 10, no. 9, p. 1687814018797261, 2018.

[13] M. Rigamonti, P. Baraldi, E. Zio, A. Alessi, D. Astigarraga, and A. Galarza, "A self-organizing map-based monitoring system for insulated gate bipolar transistors operating in fully electric vehicle," in *Annu. Conference PHM Soc.*, vol. 7, no. 1, 2015.

[14] AMD Xilinx. (2023) Xilinx Zynq 7000 SoC ZC706 evaluation kit description. [Online]. Available: https://www.xilinx.com/products/boards-and-kits/ek-z7-zc706-g.html

[15] S. Deng, H. Zhao, W. Fang, J. Yin, S. Dustdar, and A. Y. Zomaya, "Edge intelligence: The confluence of edge computing and artificial intelligence," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 7457–7469, 2020.