Preprints.org

Article

Not peer-reviewed version

Hybrid Deep Learning for Si MOSFET Fault Diagnosis and Prediction

JinJu Lee and HyunJun Choi *

Posted Date: 14 January 2026

doi: 10.20944/preprints202601.1021.v1

Keywords: Si MOSFET; power switch; semiconductor; fault diagnosis; fault prediction; gate-oxide degradation; packaging-related degradation

Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2026 doi:10.20944/preprints202601.1021.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Hybrid Deep Learning for Si MOSFET Fault Diagnosis and Prediction

JinJu Lee and HyunJun Choi *

Department of Electrical Engineering, Kumoh National Institute of Technology, Gumi 39177
* Correspondence: hyunjun.choi@kumoh.ac.kr

Abstract

Si MOSFETs are widely used in power conversion systems; however, long-term operation under repetitive switching and electro-thermal stress leads to progressive degradation and eventual failure. Two representative failure modes are commonly observed: gate-oxide degradation and packaging-related degradation, which often exhibit different evolution patterns. This paper proposes an AI-based diagnosis and prognostics framework that jointly leverages steady-state time-series information and fixed-length features extracted from turn-off transients. The study utilizes the NASA Open Accelerated-Aging dataset and reorganized/preprocessed data supported by MATLAB/Simulink measurement circuit modeling. Physics-informed rule-based labeling is applied to discriminate normal, gate-oxide, and packaging-related conditions based on degradation indicators such as R_{ds_on} evolution. The trained model is further interpreted via permutation importance to quantify whether gradual/abrupt degradation indicators and transient features contribute to decision-making. Performance is assessed on held-out tests and synthesized cases sampled from baseline operating distributions to examine consistency under previously unseen conditions.

Keywords: Si MOSFET; power switch; semiconductor; fault diagnosis; fault prediction; gate-oxide degradation; packaging-related degradation

---

1. Introduction

Since power semiconductor switches regulate current flow and enable conversion between direct current (DC) and alternating current (AC), they are core components in power conversion equipment. They are indispensable in a wide range of power electronics applications, including inverters, converters, chargers, and electric-vehicle traction systems [1, 2]. Because the switching device directly influences conversion efficiency and operational robustness, it is widely considered as one of the most critical elements in a power conversion system [3].

Among various power semiconductor devices, silicon (Si) MOSFETs are extensively used. Their manufacturing processes are mature and cost-effective, making them attractive for high-volume applications. In low-voltage operating regions, Si MOSFETs can achieve low on-state resistance drain–source on-resistance (R_{ds_on}), thereby effectively reducing conduction losses. In addition, since Si MOSFETs are voltage-driven devices, they exhibit fast switching behavior and can be operated at relatively high switching frequencies. They also avoid the tail-current phenomenon, which is inherent to IGBT-based switches. Owing to these favorable characteristics, Si MOSFETs have been widely adopted across diverse power-conversion topologies and industrial systems [4].

Despite these advantages, long-term operation of Si MOSFETs leads to the accumulation of electrical stress and thermal stress due to repetitive switching, high temperature environments, and electrical transients, as shown in Figure 1. Such stress accumulation induces device degradation and performance deterioration within the internal structure. In particular, two representative fault modes are commonly observed: gate-oxide degradation and packaging-related degradation [5].

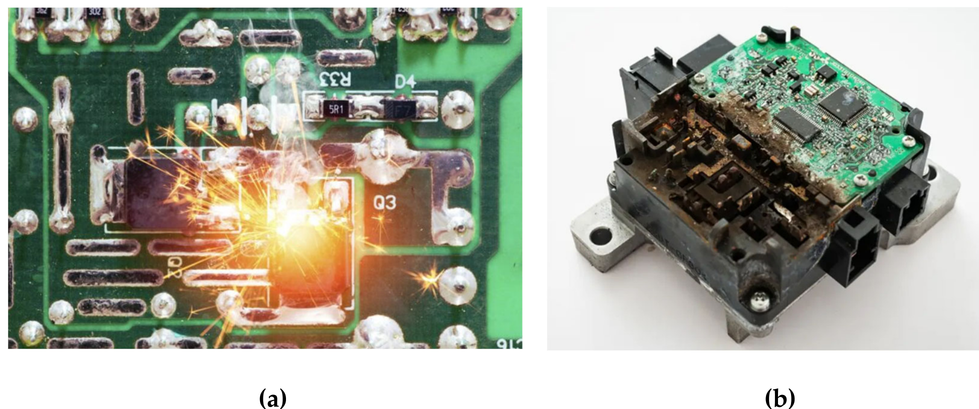
© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2026 doi:10.20944/preprints202601.1021.v1

Conventional fault diagnosis for power semiconductor devices relies on parameter measurements obtained using dedicated sensing and measurement circuits. Key physical parameters such as threshold voltage (V_{th}), drain–source voltage (V_{ds}), drain current (I_d), R_{ds_on}, source–drain voltage (V_{SD}), and junction temperature (T_j) are directly measured through these circuits and used to determine fault conditions. In many practical implementations, fault diagnosis is implemented in a post-event manner, where abnormal variations are analyzed after a fault has already occurred [5]. This diagnosis approach has limitations: subtle early-stage degradation indicators are difficult to detect, and diagnostic accuracy is frequently degraded by measurement noise and electromagnetic interference (EMI). Moreover, since the response is only possible after a fault occurrence, the system may experience shutdowns or irreversible device damage, making proactive fault prevention and effective maintenance challenging [4-7].

These limitations can be mitigated by AI-based approaches that learn fine-grained patterns from multiple sources of information, including measured parameters, transient switching waveforms, statistical characteristics, and temporal trends. In particular, degradation and fault mechanisms that evolve over long durations are often difficult to infer from single time point values. Consequently, time series learning methods are suitable for assessing degradation progression and predicting impending faults [8].

This paper proposes a novel fault diagnosis and prognostics framework for power semiconductor switches that integrates two representative AI algorithms: an Multi-Layer Perceptron (MLP) and an Long Short-Term Memory (LSTM) network. The proposed algorithm enhances the accuracy and timeliness of fault identification while enabling predictive assessment of degradation progression.



(a) **(b)**

Figure 1. (a) Power Semiconductor switch degradation. (b) Power conversion system degradation cause by component fault.

2. Fault Mode Analysis and Data-Set Collection for Si MOSFET

2.1. Analysis of Gate-Oxide Degradation Fault

Gate-oxide degradation in Si MOSFETs occurs as a result of accumulated electrical stress under conditions such as high gate electric fields, elevated ambient or device temperature, and repetitive switching operation. These stress factors promote charge trapping within the oxide layer $SiO^{2,}$ and increase interface trap density at the Si/SiO^2 interface [3, 9]. As degradation progresses, the effective gate electric field and channel formation conditions are altered, leading to V_{th} drift and reduced channel mobility. Consequently, both switching and conduction characteristics gradually deteriorate. With further accumulation of damage, the gate insulation capability weakens, which can manifest as increased gate leakage current and can ultimately lead to oxide breakdown [10].

Si MOSFETs affected by gate-oxide degradation also exhibit observable changes in the transient voltage current response. When switching speed is reduced during turn-on and turn-off transitions, the rise slope of I_d decreases under identical gate drive conditions, and the peak current tends to be

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2026 doi:10.20944/preprints202601.1021.v1

reduced. The on-state resistance R_{ds_on} gradually increases due to diminished channel conduction capability and the associated rise in junction temperature. In addition, an increase in V_{SD} leads to higher conduction losses in the body diode, thereby increasing the overall power loss. Overall, the combined increase in R_{ds_on} and switching losses not only reduces system efficiency but also decreases the thermal margin of the MOSFET under the same load conditions, thereby increasing the likelihood of fault occurrence during long-term operation [10]. Furthermore, changes in switching dynamics can interact with circuit parasitics, causing gradual variations in the overshoot and ringing characteristics of the V_{ds} transient response [11]. Depending on the observation conditions, overshoot and ringing may appear to increase slowly over time. Importantly, these behaviors typically evolve as a smooth and continuous drift driven by accumulated electro thermal stress, rather than emerging abruptly within a short time interval [12, 13].

2.2. Analysis of Packaging-Related Degradation Fault

Packaging-related degradation faults primarily evolve due to thermos mechanical fatigue within the device structure, which accumulates during repetitive thermal cycling and power cycling. In Si MOSFETs, damage can occur in various interconnect and attachment regions such as bond wires, solder joints, die-attach layers, metallization, and lead frames, appearing as lift-off, cracks, voids, solder fatigue, and non-uniform bond or joint thickness. Such package-level damage increases the series resistance along the current conduction path, raises parasitic inductance, and degrades heat dissipation by increasing thermal resistance [3, 12].

From an electrical perspective, accumulated packaging degradation increases the conduction path resistance, resulting in a rise in the on-state resistance R_{ds_on}. The increase in R_{ds_on} leads to higher conduction losses, which in turn raises the T_j. The elevated T_j further increases R_{ds_on}, forming a thermal feedback loop between R_{ds_on} and T_j. In the early stage, parameter variations may remain relatively gradual; however, once degradation exceeds a critical threshold, self-heating can increase rapidly. At this point, R_{ds_on}, T_j, and other device parameters often exhibit abrupt and nonlinear changes. Degradation is then accelerated, electro thermal stability deteriorates quickly, and the device may transition to a sudden fault state, potentially culminating in catastrophic breakdown [13].

A rapid increase in R_{ds_on} amplifies conduction path losses under the identical drive conditions, which can reduce I_d and, depending on the load, distort the current waveform. In addition, because a higher on-state resistance produces a larger voltage drop at the same load current, V_{ds} tends to increase. As packaging damage increases parasitic inductance and resistance, V_{ds} overshoot and ringing during switching transients can also become more pronounced, thereby degrading transient behavior and imposing additional loss and electrical stress [14].

V_{SD} is observed when the MOSFET body diode conducts in the off-state. It reflects the combined contributions of the diode's forward voltage component and the voltage drops caused by series resistance and parasitic inductance in the current path. As packaging degradation increases these parasitic components, V_{SD} generally rises under identical current conditions. An increased V_{SD} can increase losses and self-heating during reverse-conduction intervals; depending on operating conditions, increased body-diode losses can further intensify thermal stress. Consequently, parameters such as R_{ds_on}, V_{ds}, V_{SD}, and T_j interact and mutually reinforce degradation progression. When a packaging-related degradation fault develops, abrupt parameter fluctuations are therefore more likely to be observed [12-14].

2.3. Collection of the NASA Open Data-Set

The NASA open dataset was collected from accelerated aging experiments conducted on a practical Si MOSFET, enabling the analysis of degradation progression and fault evolution over time. In the accelerated aging protocol, repetitive switching operation is performed by applying a 15 V square wave gate drive signal. As aging accumulates, voids can form in solder joints, and the resulting increase in T_j is accompanied by a rise in R_{ds_on}. Using the open dataset, both gate-oxide

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2026 doi:10.20944/preprints202601.1021.v1

degradation fault behavior and packaging-related degradation fault behavior were identified under long-term operation, in both steady-state and transient-state measurements.

The steady-state data include operating point information recorded at each stress cycle, such as supply voltage, package temperature, V_{ds}, and I_d. Because these variables jointly reflect the device's electro thermal condition, they are useful for quantifying degradation-induced changes in resistive components and temperature rise. In particular, the temporal trend of R_{ds_on} can be tracked from V_{ds} and I_d, providing a key indicator of accumulated degradation. The transient-state data contain voltage and current waveforms during the turn-on and turn-off intervals, offering information on switching dynamic changes. Variations in V_{ds} and I_d at the switching instants, including overshoot and undershoot behavior, are influenced by changes in internal device conditions, package parasitics, and interconnect integrity.

This dataset was obtained by repeatedly executing the same stress and measurement procedure on the same device over an extended period. In this study, the data structure was reorganized to align the time evolution of an individual device and to support long-term degradation trend analysis. In addition, the dataset was preprocessed to mitigate the influence of missing values, outliers, and measurement noise on subsequent analysis and learning.

2.4. MATLAB/Simulink Circuit Design and Data Collection

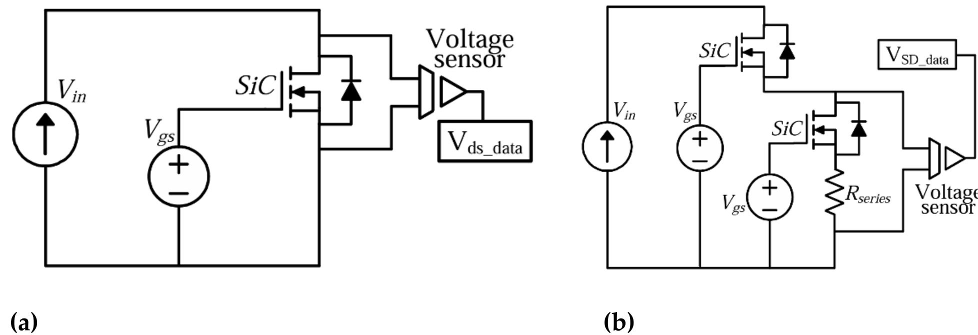
A measurement circuit was designed in MATLAB/Simulink to observe V_{ds} under conditions where V_{th} increases. As shown in Figure 2 (a), when the MOSFET is in the turn-on state, an increased V_{th} weakens channel conduction at the same V_{gs}, resulting in a reduced I_d. This behavior is consistent with gate-oxide degradation, which typically induces a gradual upward drift in V_{th} and a reduction in carrier mobility. Subsequently, I_d was calculated by substituting the measured V_{ds} and V_{th} into (1).

$$I_d = k(V_{gs} - V_{th})^n \tag{1}$$

The resulting parameter trend, a decrease in I_d, reflects the progression of a gate-oxide degradation fault. Using the measured V_{ds} and the calculated I_d, R_{ds_on} was obtained through (2), and it increases as degradation progresses.

$$R_{ds_on} = \frac{V_{ds}}{I_d} \tag{2}$$

In Figure 2 (b), V_{SD} under increased R_{ds_on} is measured during the turn-off state. The measured V_{ds} from Figure 4 (a) and the calculated I_d and R_{ds_on} from (1) and (2) were applied to the circuit in Figure 2 (b). As a packaging-related degradation fault develops, V_{SD} exhibits a rapid increase.



(a) **(b)**

Figure 2. (**a**) V_{ds} Measurement Circuit. (b) V_{SD} Measurement Circuit.

3. Proposed Fault Diagnosis Algorithm Based on MLP with LSTM

In this study, a multi-input, multi-task deep learning model is developed for fault diagnosis of Si MOSFETs by jointly exploiting steady-state time-series features and transient-state features. The

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2026 doi:10.20944/preprints202601.1021.v1

proposed model utilizes two inputs: a sequence dataset that captures parameter evolution along the degradation trajectory. The second input is a fixed-length feature vector extracted from turn-off switching waveforms, which includes voltage and current slew rates, and voltage-current correlation coefficients. The network extracts representations from the time series branch and the static feature branch independently, fuses them, and then performs fault type classification and future parameter prediction simultaneously.

The MLP consists of an input layer, hidden layers, and an output layer, where each neuron applies a nonlinear activation function to learn complex nonlinear interactions among input features [15]. In this work, the transient waveform information is represented as a fixed length vector without an explicit temporal structure, which is well suited for te MLP-based modeling. Due to its ability to express nonlinear coupling effects among heterogeneous physical quantities, the MLP can effectively learn relationships embedded in the extracted transient features [16, 17].

Conv1D layer extracts local patterns from time series data through learnable kernels, enabling efficient learning of abrupt changes and short duration morphological characteristics [17]. Here, Conv1D is applied before the LSTM to transform short term fluctuations and noise contaminated patterns in the degradation time series into a more compact and refined representation. This preprocessing step facilitates the extraction of subtle parameter variations and improves the reliability of the time series input for subsequent learning [18, 19].

LSTM networks consist of an input gate, forget gate, and output gate, and they preserve long range temporal dependencies via a gating mechanism [20]. Because Si MOSFET degradation and fault behavior often evolves gradually over extended operation rather than as a single short duration event temporal evolution provides essential information for both diagnosis and prediction [21]. Therefore, LSTM based modeling is appropriate for capturing degradation progression and its relation to fault occurrence [22].

Moreover, instead of uniformly averaging all time steps, the proposed model assigns larger weights to diagnostically important intervals [23]. This design enables clearer discrimination between rapidly varying segments associated with packaging-related degradation faults and slowly drifting segments associated with gate-oxide degradation faults, thereby improving decision making under long-term degradation trajectories.

3.1. Data Matching, Preprocessing, and Labeling

The steady-state and transient-state data were first matched under the identical experimental conditions, and the features extracted from both datasets were then combined to form the final training dataset. Each experiment was assigned a unique index extracted from the file name. Using this index, the steady-state and transient-state files were paired one-to-one paring, preventing any mixing across MOSFET devices and test runs.

After standardizing column names and aligning the time series order, data preprocessing was performed. When computing R_{ds_on}, a minimum current threshold was applied to eliminate unrealistically large R_{ds_on} values, which can occur in low current regions. To suppress abrupt fluctuations caused by measurement noise and outliers, a moving median based smoothing method was applied, and extreme values were clipped to prevent outliers from dominating the model training process. Finally, to reduce discrepancies in units and dynamic ranges across input variables, the training data were normalized to a common scale using a MinMaxScaler, as described in (3).

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{3}$$

Data labeling was performed using physics-informed rules. A packaging-related degradation fault was defined as a case in which the target indicator exhibits a continuous increase of at least 30% relative to a predefined baseline. In contrast, a gate-oxide degradation fault was defined as a case in which the indicator increases by at least 30% while showing a small jump component followed by a gradually increasing slope, reflecting a progressive drift rather than an abrupt change.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2026 doi:10.20944/preprints202601.1021.v1

3.2. Classifier Design and Training Pipeline

To achieve stable generalization performance with a limited dataset, a tree-based ensemble model was employed. Extremely Randomized Trees (ExtraTrees) constructs an ensemble of decision trees using randomized split selection, which mitigates overfitting while effectively capturing nonlinear interactions among features. Within each tree, a split at a node is chosen to maximize the reduction in impurity, where the Gini impurity is commonly used, defined as where D denotes the set of samples at a node and p_k is the proportion of class k in D. The ensemble prediction probability was obtained by averaging the class probability outputs across all trees.

$$\hat{p}(y=k \mid x)=\frac{1}{T}\sum_{t=1}^{T}\hat{p}_t(y=k \mid x) \tag{5}$$

The training pipeline consisted of missing-value handling, input scaling, and the classifier. Because some features may be missing in a subset of samples, median imputation was applied. Standardization was then used to reduce scale discrepancies across input variables. To suppress class imbalance, balanced class weighting was adopted. For reproducibility, a fixed random seed was used during model training.

3.3. Data Splitting and Performance Metrics

Model evaluation was conducted in a manner that prevents data leakage. The dataset was partitioned into training, validation, and test sets using stratified splitting, such that the class proportions were maintained consistently across all subsets. To enable fair performance assessment under class imbalanced conditions, balanced accuracy was adopted as the primary metric. Balanced accuracy is defined as the average of recall values computed for each class, as expressed in (6)

$$BA_{cc}=\frac{1}{K}\sum_{k=1}^{K}\frac{TP_k}{TP_k+FN_k} \tag{6}$$

where TP_k and FN_k denote the true positives and false negatives for class k, respectively. In addition, a classification report including per class precision, recall, and F1-score was provided, and a confusion matrix was used to analyze misclassifications among the normal, gate-oxide degradation fault, and packaging-related degradation fault classes.

3.4. Data Matching, Preprocessing, and Labeling

To examine the decision basis of the trained classifier, permutation importance was employed. This method evaluates the contribution of a feature j by measuring the decrease in model performance when that feature is randomly permuted, and it is defined as follows:

$$I_j=S(D_{test})-S(\pi_j(D_{test})) \tag{7}$$

Here, S denotes the performance metric, and π_j represents the operation that applies a random permutation only to feature j in the test set. This analysis enabled a quantitative verification of whether the model indeed exploits indicators of gradual versus abrupt degradation as well as transient-state features when making fault type decisions.

4. Simulation Results

To validate the performance of the proposed Si MOSFET fault diagnosis and prediction algorithm, AI model training and evaluation were conducted using the NASA open data set and data designed and collected using MATLAB/Simulink. From the steady-state segments of each experiment, degradation indicators were extracted to summarize the nominal operating point and the rate of resistance change. From the transient-state segments, additional features were constructed by incorporating the slope components of the turn-on waveforms and the voltage-current correlation

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2026
doi:10.20944/preprints202601.1021.v1

characteristics. For labeling, a rule-based criterion was adopted to distinguish gate-oxide degradation faults from packaging-related degradation faults. In the NASA open data set, the resulting class composition was confirmed as 50 normal, 31 packaging-related, and 20 gate-oxide samples.

Figure 3 shows the correlation structure among the input features, indicating which parameter combinations exhibit strong coupling and are therefore jointly informative for model learning. Highly correlated parameters were treated as mutually reinforcing descriptors, enabling the model to reflect inter parameter dependencies even when only subtle variations appear in an individual variable.

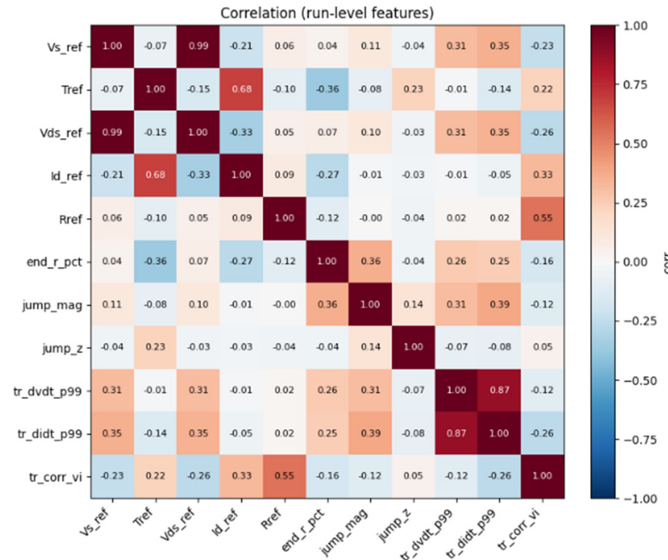


Figure 3. Pairwise parameter correlations.

The convergence and generalization tendencies during training were examined using loss, accuracy, and regression error metrics. In the loss curves of Figure 4, both the training and validation losses decrease with increasing epochs and gradually converge after approximately 40 epochs, with no substantial divergence between the two curves, suggesting limited overfitting. The accuracy increases rapidly in the early epochs and then stabilizes at a high level, indicating that the model consistently learned the separation among normal, packaging-related, and gate-oxide classes. The prediction performance in Figure 5, reported via mean squared error (MSE) and mean absolute error (MAE), also shows that both training and validation errors drop sharply at the beginning and then converge to low values. In particular, MAE directly reflects the average magnitude of prediction deviation in practical settings; the validation MAE decreases and stabilizes as training proceeds, implying that the model reduces prediction error across the overall data distribution rather than fitting only a specific region.

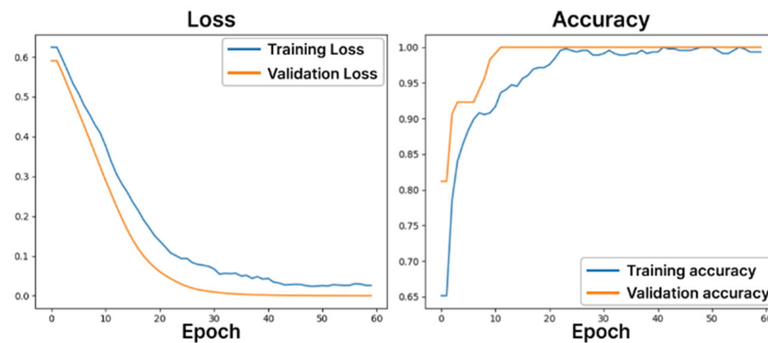


Figure 4. Fault Diagnosis AI model loss and accuracy.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2026 doi:10.20944/preprints202601.1021.v1

Figure 5. Fault Prediction AI model's learning and validation loss.

Table 1 presents the simulation results of the proposed AI algorithm on the test dataset, which was not used during training. The parameters were sampled from the baseline operating point distributions extracted from real measurements. Based on the rule-based criteria for gradual degradation and abrupt degradation, three scenarios normal, packaging-related, and gate-oxide were generated. For the 10 synthesized cases, the rule-based labels were compared with the model predictions, and the two results matched in 9 out of 10 cases. This indicates that the proposed AI algorithm can perform fault diagnosis and prediction clearly on previously unseen data. For the single-mismatched case, the rule-based criterion classified the sample as a packaging-related degradation fault, whereas the model predicted a gate-oxide degradation fault. In this case, the prediction confidence was relatively low, suggesting that the sample lies near the boundary between gradual and abrupt degradation behaviors. Overall, this result implies that the proposed model does not rely solely on temperature or the absolute magnitude of a single parameter; instead, it makes decisions by jointly considering a combination of characteristics, including the evolution pattern of R_{ds_on}, the operating point, transient switching features, and cycle-to-cycle jump-like variations in the parameters.

Table 1. Fault type Diagnosis and Prediction with key Parameter change indicators.

Case	True	Prediction	ΔR_{ds_on} (%)	Jump_mag (%)	Jump_z
1	Normal	Normal	14.83	7.69	4.94
2	Gate Oxide	Gate Oxide	118.79	3.19	17.44
3	Normal	Normal	4.65	18.91	19.88
4	Gate Oxide	Gate Oxide	104.9	8.08	24.35
5	Packaging	Gate Oxide	18.97	65.13	99.87
6	Normal	Normal	19.14	10.58	3.29
7	Normal	Normal	6.88	8.31	1.54
8	Packaging	Packaging	122.03	13.88	22.81
9	Gate Oxide	Gate Oxide	118.52	13.26	7.94
10	Normal	Normal	4.56	9.1	2.83

5. Discussion

The joint convergence of the loss and MSE/MAE curves, together with the stabilized accuracy in the learning curves, indicates that the model was trained not only to fit a classification boundary but also to produce stable predictive outputs without excessive fluctuations. When a predominantly gradual increase pattern contains sporadic spike-like cycle jumps, the resulting features can partially activate abrupt degradation indicators, leading to boundary confusion and a blurred separation between fault modes. This observation suggests that the ambiguity of the labeling rules, measurement spikes, and data quality factors may act simultaneously in the boundary region.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2026 doi:10.20944/preprints202601.1021.v1

Therefore, to improve performance near the decision boundary, it is necessary to further suppress spike effects and to enhance gate-oxide specific indicators that more strongly reflect the pre-jump trend and long-term progressive drift.

6. Conclusions

In this study, a hybrid neural network framework integrating an LSTM and an MLP was proposed to simultaneously perform fault diagnosis and parameter prediction for power MOSFETs. Key electrical parameters that vary with MOSFET degradation were used as model inputs. The LSTM branch was designed to learn long-term temporal evolution, while the MLP branch captured nonlinear coupling effects among heterogeneous parameters, enabling effective representation of degradation progression. Using simulation data constructed in a MATLAB/Simulink environment to emulate normal operation, gradual degradation, and abrupt degradation conditions, the proposed model was trained and evaluated. The results demonstrate that the proposed model can accurately classify fault types and stably predict degradation-related parameter trajectories under diverse operating conditions and temperature variations. In addition, auxiliary polynomial regression analysis confirmed that exploiting nonlinear parameter relationships enables earlier and more reliable detection than conventional single-parameter threshold-based methods. Finally, the proposed hybrid LSTM–MLP framework demonstrates robust generalization on previously unseen data, with misclassification occurring only near the boundary between gradual and abrupt degradation behaviors.

Author Contributions: "Conceptualization, Jin-ju Lee. And Hyun-jun Choi.; software, Jin-ju Lee.; validation, Jin-ju Lee.; formal analysis Jin-ju Lee. And Hyun-jun Choi.; investigation, Jin-ju Lee.; writing—original draft preparation, Jin-ju Lee.; writing—review and editing, Jin-ju Lee. And Hyun-jun Choi.; visualization, Jin-ju Lee.; supervision, Hyun-jun Choi.; project administration, Hyun-jun Choi.; All authors have read and agreed to the published version of the manuscript."

Funding: This work was supported by Korea Institute of Energy Technology Evaluation and Planning(KETEP) grant funded by the Korea government(MCEE) (NO. RS-2024-00509127)

References

1. Jiya, Immanuel N., and Rupert Gouws. "Overview of power electronic switches: A summary of the past, state-of-the-art and illumination of the future." *Micromachines* 11.12 (2020): 1116.Author 1, A.; Author 2, B. Title of the chapter. In *Book Title*, 2nd ed.; Editor 1, A., Editor 2, B., Eds.; Publisher: Publisher Location, Country, 2007; Volume 3, pp. 154–196.
2. Rafin, SM Sajjad Hossain, et al. "Power electronics revolutionized: A comprehensive analysis of emerging wide and ultrawide bandgap devices." *Micromachines* 14.11 (2023): 2045.
3. She, Xu, et al. "Review of silicon carbide power devices and their applications." *IEEE transactions on industrial electronics* 64.10 (2017): 8193-8205.
4. Tariq, Ammar, et al. "Optimizing Silicon MOSFETs: The Impact of DTCO and Machine Learning Techniques." *Electronics* 15.1 (2025): 166.
5. Isa, Ridwanullahi, et al. "Junction temperature optical sensing techniques for power switching semiconductors: A review." *Micromachines* 14.8 (2023): 1636.
6. Ren, Hai, et al. "Quasi-distributed temperature detection of press-pack IGBT power module using FBG sensing." *IEEE Journal of Emerging and Selected Topics in Power Electronics* 10.5 (2021): 4981-4992.
7. Issa, Walid, Jose Ortiz Gonzalez, and Olayiwola Alatise. "Switching and Frequency Response Assessment of Photovoltaic Drivers and Their Potential for Different Applications." *Micromachines* 15.7 (2024): 832.
8. Akbar, Ghulam, et al. "Comprehensive Review of Wide-Bandgap (WBG) Devices: SiC MOSFET and Its Failure Modes Affecting Reliability." *Physchem* 5.1 (2025): 10.
9. Wang, Peng, and Zhigang Zhao. "A Gate Oxide Degradation and Junction Temperature Evaluation Method for SiC MOSFETs Based on an On-State Resistance Model." *Electronics* 14.11 (2025): 2278.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2026 doi:10.20944/preprints202601.1021.v1

10. Kim, Jaechang, Sangshin Kwak, and Seungdeog Choi. “Impacts of sic-mosfet gate oxide degradation on three-phase voltage and current source inverters.” *Machines* 10.12 (2022): 1194.
11. Cioni, Marcello, et al. “Evaluation of V TH and R ON drifts during switch-mode operation in packaged SiC MOSFETs.” *Electronics* 10.4 (2021): 441.
12. Baker, Nick, Haoze Luo, and Francesco Iannuzzo. “Simultaneous on-state voltage and bond-wire resistance monitoring of silicon carbide MOSFETs.” *Energies* 10.3 (2017): 384.
13. Gao, Xu, et al. “Review on Power Cycling Reliability of SiC Power Device.” *Electronic Materials* 5.2 (2024): 80-100.
14. Ibrahim, Mesfin Seid, et al. “Long-term lifetime prediction of power MOSFET devices based on LSTM and GRU algorithms.” *Mathematics* 11.15 (2023): 3283.
15. Kim, Heonkook. “Fault Diagnosis in Robot Drive Systems Using Data-Driven Dynamics Learning.” *Actuators*. Vol. 14. No. 12. MDPI, 2025.
16. Zhong, Chaochun, et al. “Improved MLP energy meter fault diagnosis method based on DBN.” *Electronics* 12.4 (2023): 932.
17. Shestakov, Alexander, et al. “Hybrid CNN–MLP for Robust Fault Diagnosis in Induction Motors Using Physics-Guided Spectral Augmentation.” *Algorithms* 18.11 (2025): 722.
18. Lee, Hyeonsu, and Dongmin Shin. “Beyond Information Distortion: Imaging Variable-Length Time Series Data for Classification.” *Sensors* 25.3 (2025): 621.
19. Sun, Jiawen, et al. “Fusion of multi-layer attention mechanisms and CNN-LSTM for fault prediction in marine diesel engines.” *Journal of Marine Science and Engineering* 12.6 (2024): 990. → conv1d
20. Correia, Sérgio D., Pedro M. Roque, and João P. Matos-Carvalho. “LSTM Gate Disclosure as an Embedded AI Methodology for Wearable Fall-Detection Sensors.” *Symmetry* 16.10 (2024): 1296.
21. Qiu, Yunfeng, Zehong Li, and Shan Tian. “GA-LSTM-Based Degradation Prediction for IGBTs in Power Electronic Systems.” *Energies* 18.21 (2025): 5574.
22. Zhou, Xin, Xin Meng, and Zhenyu Li. “ANN-LSTM-A water consumption prediction based on attention mechanism enhancement.” *Energies* 17.5 (2024): 1102.
23. Liu, Haopeng, and Lufeng Yang. “A Comparative Study of CNN-sLSTM-Attention-Based Time Series Forecasting: Performance Evaluation on Data with Symmetry and Asymmetry Phenomena.” *Symmetry* 17.11 (2025): 1846.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.