Deep Learning-Based Failure Prognostic Model for PV Inverter Using Field Measurements

Liming Liu, *Student Member*, *IEEE*, Yi Luo, *Student Member*, *IEEE*, Zhaoyu Wang, *Senior Member*, *IEEE*, Feng Qiu, *Senior Member*, *IEEE*, Shijia Zhao, *Senior Member*, *IEEE*, Murat Yildirim, *Senior Member*, *IEEE*, Rajarshi Roychowdhury, *Senior Member*, *IEEE*

***Abstract*—This study presents a novel approach for the precise monitoring and prognosis of photovoltaic (PV) inverter status, which is crucial for the proactive maintenance of PV systems. It addresses the gaps in traditional model-based methods, which tend to neglect the overall reliability of inverters, and the limitations of data-driven approaches that largely depend on simulated data. This research presents a robust solution applicable to real-world scenarios. The proposed data-driven model for PV inverter failure prognosis employs actual inverter measurements, integrating various operational and weather-related factors based on domain knowledge. This approach effectively represents inverter stressors and operational status. Utilizing an Enhanced Siamese Convolutional Neural Network (ESCNN), the model merges operational data with domain knowledge features, redefining the prognosis challenge as a classification task. Furthermore, the paper discusses an ESCNN-based real-time inverter failure monitoring method developed on the well-trained model. The proposed models are rigorously trained and tested with real inverter data and a novel filtering method is included to address accidental failures in practical scenarios. The results validate the model's efficacy, and the directions for future research are also outlined.**

***Index Terms*—PV inverter, failure prognosis, deep learning, Siamese neural network, field measurements, domain-knowledge features**

I. Introduction

In the context of the escalating global energy crisis, renewable energy sources have garnered considerable attention from the research community [1] [2]. Photovoltaic (PV) solar energy, in particular, has emerged as a key player in the sustainable energy landscape. Projections suggest that by 2030, PV power generation in distribution systems could reach an impressive annual output of 650 GW [3]. However, the realization of this potential is constrained by operational and environmental challenges that can significantly impact the efficiency and longevity of PV systems [4]. PV inverters, which are vital components responsible for converting solar energy into usable electricity, are especially prone to these stressors. Their sensitivity underscores the necessity for robust and effective maintenance strategies. Ensuring the resilience of these inverters through proactive maintenance, vigilant monitoring, and failure prognosis is crucial. Such measures are essential to enhance the reliability and sustainability of PV installations, thereby supporting the broader objective of meeting global renewable energy goals. To effectively implement these strategies, a precise and reliable method for failure prognosis and monitoring is essential.

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technologies Office Award Number 38458. The views expressed herein do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

The failure prognosis research has obtained significant interest across various fields, including applications in battery [5], rolling bearings [6] and wind turbines [7]. Similarly, the issue of solar inverter failure prognosis has emerged as a focal point of research, attracting considerable attention. Generally, the existing work can be divided into two primary categories: empirical model-based methods and data-driven methods.

Regarding the empirical model-based methods, the most prominent approach is the lifetime model of power electronics derived from accelerated power cycling test data, which estimates the device's life expectancy based on thermal cycling parameters [8]. A survey of lifetime models for power semiconductors is provided in the report [9], highlighting several key models: the Coffin-Manson Model [10], the Norris-Landsberg Model [11], and the Bayerer Model [12]. Along with these models, a general flowchart for assessing the reliability and useful life of PV inverters is described in the report. Extending the basic models, researchers explore various scenarios and potential points. [13] investigates the impact of panel positioning and degradation on PV inverter lifetime, finding significant effects from panel orientation but minimal from tilt angles. [14] discusses the reliability of three types of DC-link capacitors in power electronics, presenting the representative lifetime models. This study also highlights challenges in current methodologies, including the resource-consuming nature of accelerated lifetime testing, variations in constant parameters within lifetime models, and the limited range of stressors considered. [15] examines the benefits of local reactive power provision on system performance and its implications for PV inverter longevity. The study utilizes a parameter-based capacitor degradation model present in [14] to assess inverter lifetimes and predict potential failure points. [16] introduces a Monte Carlo-based analysis method for predicting the lifetime consumption of bond wires in insulated gate bipolar transistor (IGBT) modules of PV inverters. This approach differs from previous studies by incorporating variations in IGBT parameters, such as the on-state collector–emitter voltage, diverse lifetime models, and a range of environmental and operational stresses into the lifetime prediction process. Most prior studies on PV inverter reliability have focused solely on the effects of thermal cycling on power semiconductors and DC-link capacitors in the power stage. However, the overall system reliability

and lifetime are also influenced by other components. A more comprehensive model that accounts for the overall reliability of inverters, considering all these components is essential for a complete understanding and accurate prediction of inverter lifespan.

In terms of the data-driven method, [17] introduces an innovative framework utilizing a multi-evaluation index parameter optimization model and Mahalanobis distance. This framework features a unique inverter health indicator, which is developed using a fault feature library encompassing various time-domain, frequency-domain, time-frequency domain characteristics, and circuit performance parameters. However, this health index has yet to be validated with real inverter measurement data, which could limit the application of the index. [18] employs the LightGBM model as an alternative to traditional thermoelectric coupling models, enhancing the efficiency of calculating IGBT junction temperatures and reducing the reliance on IGBT model parameters for reliability evaluations. However, this study focuses exclusively on the reliability of IGBTs, not addressing the reliability of the entire inverter system. Similar to LightGBM, another ensemble algorithm, random forests (RFs), is utilized alongside the Concordia transform to enhance the robustness of the fault diagnosis classifier in [19]. [20] presents a failure diagnostic scheme using the AlexNet deep learning model to classify various fault types in PV systems. While the results are promising, the model is only tested on the simulated data. Based on the traditional convolutional neural network (CNN) structure, [21] introduced the global average pooling layer to replace the fully connected layer to realize fast fault diagnosis of the DC-DC inverter. Similarly, [22] details a fault diagnosis strategy using an attention-assisted recurrent neural network, focusing on extracting key information from signal context. In a different approach, [23] improves support vector machine (SVM) classification accuracy by using a radial basis kernel function and grid-search optimization. However, the performance of SVM is significantly influenced by the choice of kernel function and its parameters. Likewise, principal component analysis (PCA) is employed in [24] as a feature selection model, followed by the proposal of a secondary classification fault diagnosis strategy based on PCA-SVM. While previous studies have achieved impressive results in data-driven inverter failure prognosis, a common issue persists: most rely on simulated data, which tend to be idealized and free from noise or accidental external disturbances. This dependence on computer-simulated PV measurements could potentially diminish the applicability and credibility of these models in real-world scenarios.

Moreover, accurate prediction and analysis of inverter lifetime require understanding the stress factors affecting it. Oversizing PV arrays, also known as under-sizing PV inverters, is a common strategy to maximize inverter capacity and reduce energy costs, but it can significantly affect inverter longevity. [25] assesses the impact of PV array sizing on inverter reliability and lifetime. Their findings indicate that variations in PV array sizing can markedly alter the expected reliability and lifespan of inverters. Additionally, environmental stressors like solar irradiance, temperature, and humidity's effects on inverter reliability and lifespan warrant more research, as suggested by [26]. [27] and [28] further highlight how environmental conditions, particularly temperature extremes and humidity, influence inverter performance and longevity, with extreme temperatures inducing thermal stress and potentially exceeding operational limits. Despite their significance, these internal and external stressors are often overlooked in current failure prognosis research.

In light of these challenges, this study proposes a data-driven PV inverter failure prognosis model that incorporates multiple designed domain knowledge features to evaluate the overall reliability of PV inverters. Rather than utilizing the simulation or lab-generated data, the field measurements are utilized to build and test the model, ensuring its effectiveness in real-world settings. The main contributions of this work are summarized as follows:

- This research develops a data-driven prognosis model for PV inverter failures using an Enhanced Siamese Convolutional Neural Network (ESCNN). It incorporates multiple operational and weather-related factors grounded in domain knowledge to reflect inverter stressors and operational status. This integration of specific features boosts the model's predictive accuracy.
- In this study, an inverter sample filtering process is adopted, which employs a two-stage clustering framework based on the Gaussian Mixture Models (GMM). By identifying and excluding accidental external disturbance failures, the filtering process can minimize the impact of such anomalies, ensuring optimal use of the dataset for enhanced model accuracy and reliability.
- The proposed prognosis model is developed, validated, and tested using a dataset of around 200 real-world inverters. Distinct from simulation data and lab testing results, this industrial dataset incorporates genuine noise and accidental external disturbances. Building and testing the model on such real-world data significantly enhances its reliability in practical applications.

The rest of the paper is organized as follows. Section II provides the problem statement and the descriptions of the dataset. The inverter dataset filtering method and inverter stress feature extraction are formulated in Section III. Section IV presents the ESCNN-based solar inverter failure prognosis framework. Numerical results on the real inverter measurements are given in Section V and the paper is concluded in Section VI.

II. Problem Statement and Dataset Description

A. Dataset Description

The PV inverters data used in this study were sourced from two industrial companies. The dataset encompasses information gathered from more than 700 PV inverters, which are distributed across 440 different locations. These sites include a mix of residential and commercial customers, providing a broad and diverse range of operational data for analysis. The dataset in this project covers data from the date of installation of each inverter up to the end of 2022, with a resolution of 5 minutes. For each inverter, we recorded a variety of operational parameters including total active power, AC side current and voltage, frequency, and DC side voltage. Fig. 1 illustrates

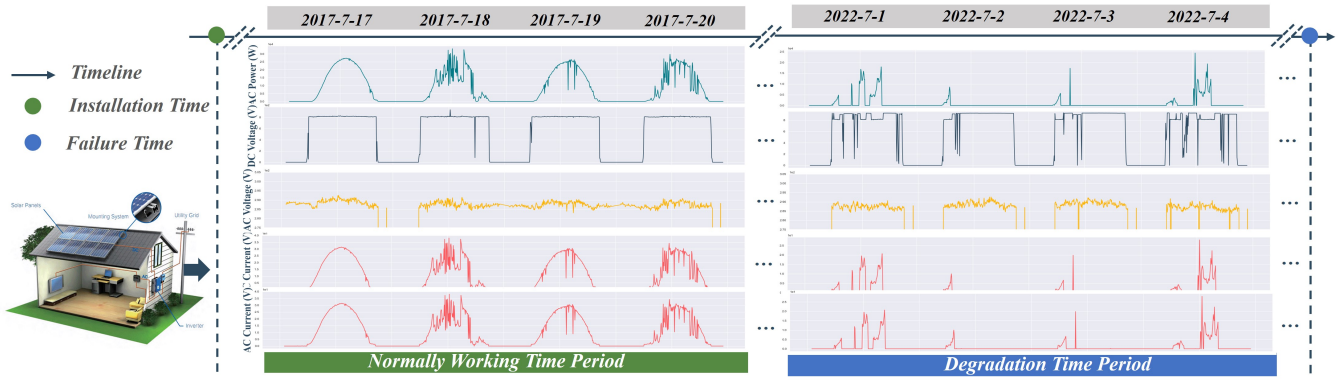

Fig. 1. The operational data snaps of a selected inverter with degradation progress

a segment of operational data from the lifetime curve of a selected inverter, showcasing its degradation progression. Initially, the data indicates excellent functioning of the inverter. However, after approximately five years, clear signs of degradation appear in its operational status. This is evidenced by erratic fluctuations in the AC side current and power, indicating inconsistencies in solar generation, and by the DC voltage exhibiting variations rather than maintaining a steady value as seen at the outset. Due to the grid-connected nature of the system, the AC side voltage and frequency remain largely unchanged. A key distinction of this operational data, as opposed to laboratory or simulated data, is its lower fidelity, characterized by outliers, noise, and reduced observability due to lower resolution. These traits are common in inverter operational data collected from industrial environments.

Besides the historical operational data, the alert records of the inverters are also included in the monitoring system. The alert codes capture and record historical anomaly events of the inverters, offering detailed information on abnormal operations. These alert codes include various alert types categorized into four main areas: equipment, grid, storage, and communication. Among these, inverter fault-related alerts—such as *Inverter Production Issues Detected, DC Isolation*, and *Residual Current Device Signal*—are specifically used to assess abnormal working conditions of inverters. Alerts that could be unrelated to the faults, like *No Communication*, are excluded from the analysis. Out of all the inverters in the dataset, approximately 85 have failed at least once, leading to 103 complete replacement events. While we have logs of all these replacement events, the logged dates do not necessarily align with the actual dates of inverter failure owing to possible delays in repair and replacement. Thus, we combine operational data with replacement records to more precisely identify the failure time of each inverter. Determining failure time is crucial for developing an accurate data-driven failure prognostic model.

B. Problem Restatement

Accurate inverter failure prognostic is vital for optimizing inspection schedules, enhancing repair efficiency, and reducing maintenance costs. Predicting the exact failure time is challenging due to complex failure reasons and uncertainties in field operation conditions. To make the failure prognosis more feasible, our approach strikes a balance between "informativeness" and "reliability" by framing it as a classification task. This approach maximizes the support for solar energy companies with our prognostic results. The classification details are defined as follows:

$$\boldsymbol{y}^s = M(\boldsymbol{X}, \mathcal{A}, \mathcal{W})$$

where $\boldsymbol{X}$ represents the operational data of inverters; $\mathcal{A}$ contains alert information that may include valuable degradation details of inverters, and $\mathcal{W}$ encompasses weather information from the respective inverter locations. $\boldsymbol{y}^s$ contains labels for each inverter. For each equipment i in $\boldsymbol{y}^s$, $y_i^s = 1$ indicates the anticipated failure of inverter i in a future time slot s; conversely, y_i^s is labeled as zero. The time slot in this study is set at one month, considering the practical needs of solar energy companies. In summary, our work aims to model the mapping M using an industrial dataset, employing deep learning techniques to construct the framework. With the well-trained model M_f, we can predict the operational status of inverters in future time periods.

III. Feature Extraction and Dataset Filtering Based on Domain Knowledge

In this section, several key features are generated based on domain knowledge, focusing on inverter degradation and external stressors, enabling the model to detect signs of inverter degradation accurately. Additionally, we establish a framework for filtering the failure samples, which helps refine the dataset for the subsequent failure prognostic task.

A. Operational Feature Extraction

Inverters accumulate vast operational data over the years. Utilizing all this data directly can be inefficient and cumbersome, as redundant data might obscure crucial degradation indicators, particularly in limited datasets. Hence, it's essential to pinpoint key features related to degradation. This section highlights two such features derived from domain expertise, incorporating external stress factors like severe weather into our analysis.

1) Equivalent Under-sizing Rate: Over-sizing PV arrays (or under-sizing inverters) is a common strategy to maximize the utilization of the inverter capacity, reduce energy costs, and increase the PV system's return on investment [25]. While under-sizing an inverter leads to more efficient use by frequently operating at rated power, it also means prolonged high-power operation, which can shorten its lifespan. This is due to increased thermal stress on components like power devices and capacitors, impacting their reliability [25]. Thus, considering the effects of under-sizing is vital for inverter failure analysis. Traditionally, the calculation of the under-sizing rate is based on nominal values or standard test parameters. However, this approach does not account for variable field operation conditions and weather factors, such as cloudy or rainy days, which affect PV panels output. Therefore, the standard method for calculating under-sizing may not accurately reflect real operational stresses. To accurately reflect inverter under-sizing in real-world conditions, we introduce a new metric, the equivalent under-sizing rate (EUR) [29], denoted as R_{e}. It captures the true operational scenarios of inverters under diverse field conditions, thus offering a better assessment of under-sizing. This metric is based on industry data from our dataset. An example is shown in Fig. 2, illustrating an inverter's operational characteristics during under-sizing. The main symptoms that are observed are high generation levels, leading to AC power clipping and elevated DC voltage due to excess PV generation. These phenomena, further discussed in the following subsection, form the basis of our R_{e} definition as:

$$
\begin{aligned}
R_{\mathrm{e}} &= \frac{\sum_{d=1}^{D}\sum_{t=T_{\mathrm{s}}}^{T_{\mathrm{e}}} F\left(v_{d,t}^{\mathrm{dc}}, p_{d,t}^{\mathrm{ac}}\right)}{D \times (T_{\mathrm{e}} - T_{\mathrm{s}})} \\
F\left(v_{d,t}^{\mathrm{dc}}, p_{d,t}^{\mathrm{ac}}\right) &= \mathbf{1}_{\mathbb{R}_+}\left(v_{d,t}^{\mathrm{dc}} - v_{\mathrm{trh}}^{\mathrm{dc}}\right) \odot \mathbf{1}_{\mathbb{R}_+}\left(p_{d,t}^{\mathrm{ac}} - \theta \cdot \max\left(p_{\mathrm{ac}}\right)\right) \\
v_{\mathrm{trh}}^{\mathrm{dc}} &= C_{.75}\left(v^{\mathrm{dc}}\right) + 1.5\left[C_{.75}\left(v^{\mathrm{dc}}\right) - C_{.25}\left(v^{\mathrm{dc}}\right)\right] \\
p_{\mathrm{ac}} &= \left\{p_{d,t}^{\mathrm{ac}}, d = 1, 2, \ldots D, t = T_{\mathrm{s}}, T_{\mathrm{s}}+1, \ldots, T_{\mathrm{e}}\right\},
\end{aligned} \tag{1}
$$

where, $v_{d,t}^{\mathrm{dc}}$ and $p_{d,t}^{\mathrm{ac}}$ represent the DC voltage and AC active power of the inverter at time t on day d, respectively. The symbol $\odot$ denotes element-wise multiplication. The functions $C_{.25}(\cdot)$ and $C_{.75}(\cdot)$ refer to the 25th and 75th percentiles of the target set. T_{s} and T_{e} indicate the start and end times of the selected analysis period, which are 9 a.m. and 3 p.m., respectively. The variable θ represents the load level. The metric R_{e} primarily measures the percentage of time an inverter operates in an undersized state during the specified intervals. Given that heavy loading is a condition for an undersized inverter, R_{e} also highlights periods of significant load stress, which are crucial for the degradation of inverter components. Therefore, by incorporating stresses from capacitor fatigue and component overuse, R_{e} serves as a robust indicator for assessing the operational longevity of inverters.

To give a mechanism analysis of R_{e} metric, a typical three-phase PV inverter circuit is adopted without loss of generality, as shown in Fig. 3. Based on the structure, the voltage-current interrelation on the DC side can be defined as:

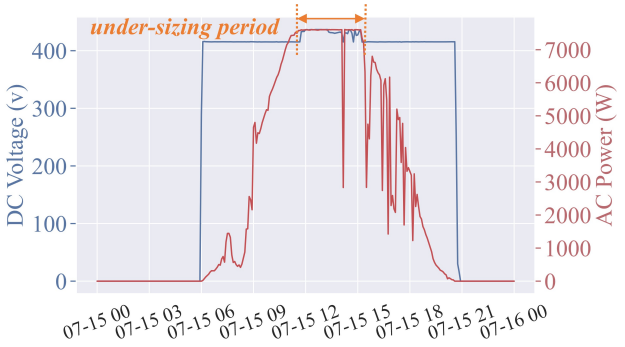

Fig. 2. The illustration of the inverter under-sizing using field measurements

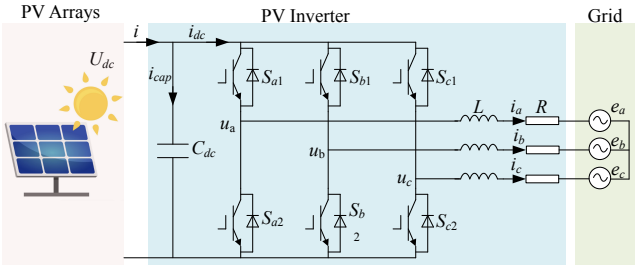

Fig. 3. Main circuit of grid-connected PV inverter

$$
\begin{cases}
i = i_{\mathrm{cap}} + i_{\mathrm{dc}} \\
i_{\mathrm{dc}} = S_{\mathrm{a}} i_{\mathrm{a}} + S_{\mathrm{b}} i_{\mathrm{b}} + S_{\mathrm{c}} i_{\mathrm{c}} \\
i_{\mathrm{cap}} = C\dfrac{dU_{\mathrm{dc}}}{dt}
\end{cases} \tag{2}
$$

where, S_x ($x = \mathrm{a}, \mathrm{b}, \mathrm{c}$) denotes the corresponding switch function; i_{dc} represents inverter DC current; i_{cap} is the capacitor current, and $i_{\mathrm{a,b,c}}$ denotes the inverter output current.

Peak sunlight conditions can lead to excess power production by PV panels, posing a challenge for undersized inverters in AC power conversion. Once the inverter's DC current limit is reached, any additional DC power charges the DC bus capacitor, as detailed in (2). This can lead to overvoltage and frequent charging/discharging cycles, jeopardizing components like the DC link capacitor and reducing their lifespan. Thus, proper sizing of inverters is key to balancing cost efficiency with maintenance demands. It is also important to note that inverter overvoltage is not solely due to under-sizing. Other factors, such as abrupt increases in PV array output, lightning strikes, and unusual system startups, can also cause voltage surges. Consequently, in designing the R_{e} indicator, we take into account not just voltage variance, but also AC power characteristics and specific time slots for analysis.

2) Abnormal Event Rate: The alert records in the dataset can help accurately locate the position of abnormal events, providing more support for monitoring the inverter's health status. Fig. 4 shows the operational features collected over three days from an inverter, with corresponding inverter-related alerts also marked on the figure. As can be seen, two types of abnormal events, DC isolation and inverter production issues, occurred within three days. When these faults happen, typical characteristics include AC output (current and power) decreasing to a low level, nearly zero, while DC voltage increases. How-

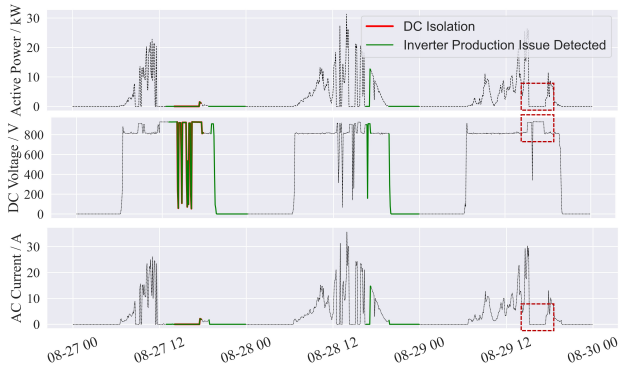

Fig. 4. The snaps of inverter operational data with alert information

ever, it is evident that not all abnormal events are captured by the alerts. Some obvious generation losses are observed in the features but not marked by the alert records, like the red rectangular area shown in the figure. To address this issue, we propose the abnormal event ratio (AER) to capture the abnormal operations and provide a general description of the operation status during any selected period. The abnormal events are defined as follows:

$$R_{\mathrm{a}} = \frac{\sum_{d=1}^{D} [\![\sum_{t=s}^{T_e} H(v_{d,t}^{\mathrm{dc}}, i_{d,t}^{\mathrm{ac}}, p_{d,t}^{\mathrm{ac}}) > \Xi]\!]}{D}$$
$$H(v_{d,t}^{\mathrm{dc}}, i_{d,t}^{\mathrm{ac}}, p_{d,t}^{\mathrm{ac}}) = \mathbf{1}_{\mathbb{R}_+}(v_{d,t}^{\mathrm{dc}} i_{d,t}^{\mathrm{ac}} p_{d,t}^{\mathrm{ac}} - \xi) \odot \mathbf{1}_{\mathbb{R}_+}(v_{d,t}^{\mathrm{dc}} - v_{\mathrm{trh}}^{\mathrm{dc}}), \tag{3}$$

where, $i_{d,t}^{\mathrm{ac}}$ denotes the AC current power of the inverter at time point t of day d. $[\![\cdot]\!]$ represents Iverson's convention; Ξ represents a tolerance for accidentally abnormal operation behaviors, which means only the abnormal status last a long time, the day will be marked as an abnormal event. ξ is a non-zero small quantity, which is set as 0.1 in our analysis. The abnormal event rate can complement alert records, providing a comprehensive view of abnormalities, which typically represent degradation symptoms.

3) Weather-related Impact Features: Extreme temperature-related factors are the significant inverter stress factors, which have been mentioned in previous works [30]. In this work, temperature factors T_{H} and T_{L} are introduced to comprehensively analyze how extreme temperature-related factors affect inverter lifetime. Specifically, T_{H} represents the average high temperature experienced by the top 1% of instances, accounting for dynamic stress conditions during periods of extremely high-temperature events. T_{L} denotes the average of the lowest 1% daily temperature over a long period.

Extreme weather events, though less impactful on inverters than solar panels due to their different installation positions, still affect inverter lifetimes. Such events include lightning, intense heat, and thunderstorms. To measure the impacts, we introduce a metric called C_{SE}, the average number of extreme weather events. It's calculated by dividing the total count of these events by the inverter's working time, which is discussed in the next section.

B. Selective Filtering of Inverter Failure Samples

Operating under diverse conditions, inverters experience varying levels of product quality and workplace stress, leading to different failure modes upon malfunctioning. These include not only the gradual wear and tear of components but also external, accidental impacts, defined here as "shock-based failures". The identification of these failures presents a challenge due to their lack of distinct symptoms. Moreover, even with a focus on degradation-based failures, the characteristics of inverters can vary significantly based on their wear levels. Distinguishing and excluding shock-based failure cases is critical to prevent data distortion in the development of the subsequent prognostic model. To effectively address this issue, the paper proposes a two-stage method for identifying shock-based failures, employing the clustering approach.

1) Clustering Analysis Based on GMMs: By clustering the failure data, we aim to identify distinct groups of failures that exhibit similar behavior or patterns. Given the feature extraction conducted earlier, we employ these extracted features for the clustering process, and the GMM clustering method is selected. GMM, widely used in pattern recognition and data mining, assigns each data point a probability of belonging to a particular cluster. It calculates Gaussian distribution parameters, including means, covariances, and cluster membership probabilities, through an iterative Expectation-Maximization algorithm. The GMM's probability density function, a weighted sum of these Gaussian distributions, represents the dataset. Given a dataset with N data points, the probability density function (PDF) of the GMM can be mathematically represented as:

$$P(X) = \sum_{k=1}^{K} \pi_k \cdot \mathcal{N}(X \mid \mu_k, \Sigma_k), \tag{4}$$

where $P(X)$is the PDF of the GMM for the dataset X; K is the number of clusters, which are the Gaussian components, in the mixture model; π_k represents the weight of the k-th Gaussian component, satisfying $\sum_{k=1}^{K} \pi_k = 1$; $\mathcal{N}(X \mid \mu_k, \Sigma_k)$ is the PDF of the k-th Gaussian component. The GMM aims to find the optimal values for the parameters π_k and μ_k that maximize the likelihood of the data given the model. Once trained, the GMM assigns data points to their most likely generating cluster. The strength of GMM lies in managing complex, overlapping or irregularly shaped data distributions. Our dataset presents specific characteristics that make the GMM particularly suitable for this task. In our case, it's impractical to rigidly separate samples of different inverter failure modes due to indistinct boundaries in the feature spaces. This issue also arises when differentiating between failing and functioning inverters. The GMM's ability to handle overlapping clusters or features makes it a promising method for analyzing inverter samples in our study. Utilizing the GMM clustering model, we designed a two-stage shock-based method for failure identification.

2) Two-stage shock-based failure identification method: The two-stage shock-based failure identification framework, illustrated in Fig. 5, focuses on identifying failure samples with operational features similar to functioning inverters. The first stage involves identifying typical failure modes within the dataset of failed inverters. Using the domain-knowledge features, these inverters are categorized into Q clusters, and

the cluster labels serve as inputs for the subsequent stage. The features for this phase are derived from previously mentioned domain-knowledge features, including the R_{e} over the inverter's lifetime, the R_{a} during specific periods (like the first three months after installation and the months preceding failure), and the continuous working time (CWT) of the inverters. The CWT is defined as the period from the installation date to the present time. Obviously, for the failed inverters, it encompasses the interval from the inverter's installation date to the moment it experiences failure, essentially representing the lifespan of the inverters. These features collectively constitute the feature set X_{F} for the failure samples.

In the first stage, the GMM clustering method is applied to segregate the inverter failures into distinct groups based on their exhibited behavior or patterns. This approach is predicated on the assumption that inverters experiencing shock-based failures do not show typical degradation symptoms. Consequently, most shock-based inverter samples are likely to be clustered into a single group, designated with the label c and the associated features denoted as $X_{\mathrm{F}}^{\mathrm{c}}$. In contrast, samples that display degradation characteristics are clustered into other groups, presumed to be labeled a and b, and the features denoted as $X_{\mathrm{F}}^{\mathrm{a}}$ and $X_{\mathrm{F}}^{\mathrm{b}}$. It is important to note that while we reference two clusters for simplicity, the actual number of clusters could be greater and will be determined based on the performance of the clustering. These clustering results then serve as inputs for the second analysis stage.

In the second stage, we amalgamate the normally working inverters, denoted as X_{N}, with each identified failure group to create distinct sub-datasets. This combination is mathematically represented as $X^g = X_{\mathrm{F}}^g \cup X_{\mathrm{N}}$, where $g \in \{\mathrm{a}, \mathrm{b}, \mathrm{c}\}$. As a result, three sub-datasets are formed: X^{a}, X^{b}, and X^{c}. Each sub-dataset undergoes a further clustering process to discern and categorize sample patterns. For X^{a}, which comprises inverters with degradation-based failures, the clustering tends to separate these failure inverters into their own group, due to their distinct characteristics from normally functioning inverters. This clear segregation is observable in the left bar chart of corresponding sub-cluster 1, where all failure samples (depicted in blue) are grouped distinctly, illustrating an obvious separation. A similar pattern is observed with X^{b}. However, the sub-dataset X^{c}, which includes shock-based failures, presents a different scenario. The underlying hypothesis is that shock-based failures, not exhibiting prolonged degradation, might display operational statuses similar to those of normally functioning inverters prior to their failure. Consequently, these inverters are not distinctly segregated by the clustering method, leading to failure samples appearing across all groups without clear separation, as depicted in the right bar chart of Fig. 5. By implementing this two-stage shock-based failure identification framework on the datasets, we can discern shock-based failures, as evidenced by the lack of isolation of these samples in the second-stage clustering.

Our primary objective in filtering the dataset is to exclude as many shock-based inverters as possible while retaining those exhibiting degradation symptoms. Indeed, there is no strict separation between shock-based inverters and others, which

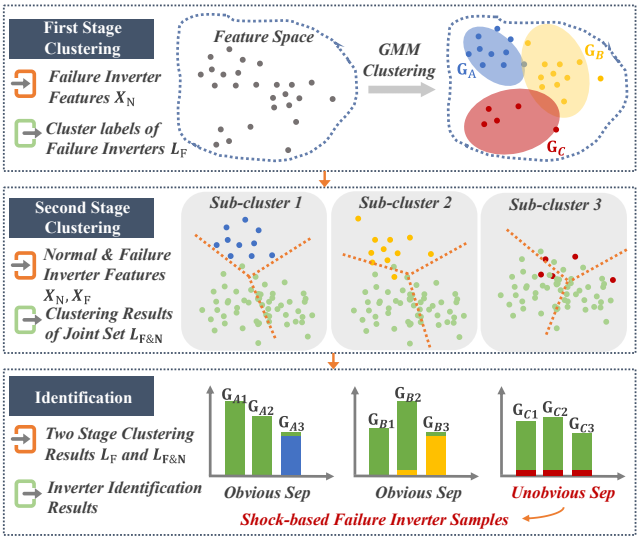

Fig. 5. The two-stage shock-based failure identification framework

complicates this task. However, this selective filtering enhances the dataset's quality and reduces the invalid information introduced by the shock-based inverters, further boosting the prognostic model's accuracy. Without this process, shock-based samples would mask the degradation signals during model training, a risk exacerbated by the limited size of the failure sample set. Although some inverters might be "mislabeled" during this process, the refined dataset provides a more reliable foundation for training our ESCNN model than an unprocessed dataset would.

IV. PV Inverter Failure Prognostic Model

In this section, the ESCNN-based PV inverter failure prognostic model is designed, as depicted in Fig. 6. The model is comprised of three primary components: the domain knowledge integrated network, the contrastive feature extraction network (i.e., basic SCNN model), and the classification network. Operational features (i.e., R_{e} and R_{a}) and weather-related impact features serve as inputs for the domain-feature integrated network. These inputs are utilized to generate long-term features since they encompass the entire lifespan of the inverter. The SCNN model plays a pivotal role in generating contrastive features, which highlight the differences between the initial operational phase and the performance after prolonged usage. By amalgamating these two types of features, namely long-term and contrastive features, the classification network is capable of deriving prognosis results.

A. Contrastive Feature Extraction

In Fig. 1, it's evident that inverters show signs of degradation after extended use. This degradation is marked by a deviation in the inverter's operational state from its initial state, eventually leading to failure. Identifying these deviations in operational features is key to predicting inverter failures. The extent and nature of these deviations are crucial indicators. The focus of the failure prognostic process is centered on the detection of differences. While certain deviations are clear and

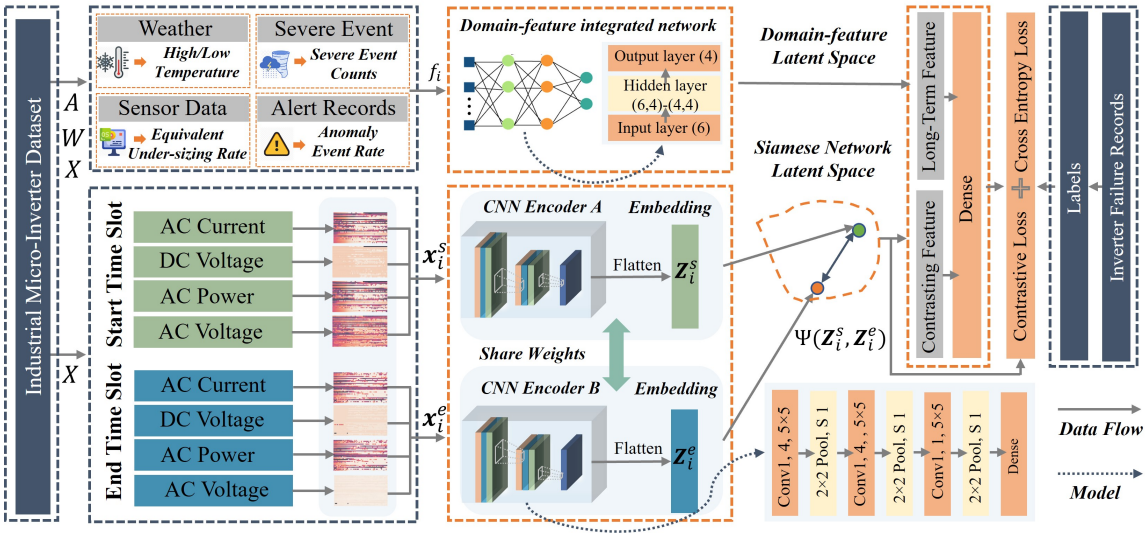

Fig. 6. The structure of the ESCNN-based PV inverter failure prognosis model

can be identified manually, subtler ones necessitate more advanced detection techniques. In this context, the deep learning model is instrumental. It serves to transform less evident features into more discernible ones, referred to as contrastive features, thereby simplifying their identification and analysis. The realization of this process is facilitated by the Siamese neural network.

B. Siamese Convolutional Neural Network

The Siamese neural network (SNN), pioneered by Bromley and LeCun [31], has significantly advanced signature verification with its unique "twin" subnetwork architecture that shares weights and processes different inputs efficiently. It excels in transforming complex, high-dimensional inputs into a simpler latent space, ideal for tasks like failure prognosis. This approach, especially effective with fewer samples, involves comparing long-term feature distributions of inverters. In our case, the SNN's subnetworks, acting as encoders, are integrated with convolutional neural network (CNN) layers [32], known for their effectiveness in image and video recognition, to form the Siamese convolutional neural network (SCNN). These CNN layers allow the SCNN to adaptively learn from varied data resolutions and historical ranges in industrial inverter operations. The CNN's local connectivity also simplifies the model by reducing the number of parameters, which is beneficial for mitigating overfitting and improving computational efficiency.

We selected the SNN model for our inverter failure prognosis task due to its superior capabilities in analyzing and comparing inputs, which perfectly aligns with our objective to detect and evaluate discrepancies between the operational statuses of inverters over time. Specifically, there are three main reasons for choosing the SNN model: firstly, the unique architecture of SNNs, featuring parallel subnetworks that share weights, is inherently proficient at assessing subtle similarities and differences between paired inputs, which is crucial for analyzing inverter performance across different periods [31]. Secondly, the robustness of SNNs in handling imbalanced and limited data sets makes them ideal for our context, where inverter failure samples are scarce, thus effectively mitigating data bias and enhancing model generalization [33]. Lastly, the transferability of learned features by SNNs facilitates effective adaptation using transfer learning techniques [34], accommodating the lack of extensive historical data and ensuring the model's efficacy under the unique conditions faced by solar companies.

Considering the twin structure of SCNN, the paired samples set building process is conducted to promise the inputs to satisfy the requirement of the model. Dividing the raw dataset into training set $\boldsymbol{\mathcal{D}}_{tr}$ and testing set $\boldsymbol{\mathcal{D}}_{te}$, the two sets can be expressed as:

$$\boldsymbol{\mathcal{D}}_{tr} = \{\boldsymbol{X}_{tr}, \boldsymbol{y}_{tr}\}$$
$$\boldsymbol{\mathcal{D}}_{te} = \{\boldsymbol{X}_{te}\}$$

where

$$\boldsymbol{X}_{tr} = \begin{bmatrix} \boldsymbol{x}_1^{tr,s} & \boldsymbol{x}_1^{tr,e} \\ \boldsymbol{x}_2^{tr,s} & \boldsymbol{x}_2^{tr,e} \\ & \vdots \\ \boldsymbol{x}_n^{tr,s} & \boldsymbol{x}_n^{tr,e} \end{bmatrix}, \boldsymbol{y}_{tr} = \begin{bmatrix} y_1^{tr} \\ y_2^{tr} \\ \vdots \\ y_n^{tr} \end{bmatrix}, \boldsymbol{X}_{te} = \begin{bmatrix} \boldsymbol{x}_1^{te,s} & \boldsymbol{x}_1^{te,e} \\ \boldsymbol{x}_2^{te,s} & \boldsymbol{x}_2^{te,e} \\ & \vdots \\ \boldsymbol{x}_n^{te,s} & \boldsymbol{x}_n^{te,e} \end{bmatrix}$$

$\boldsymbol{X}_{tr}$ is the set of sample data in the training set and $\boldsymbol{y}_{tr}$ collects the corresponding labels for the sample data, indicating whether the inverter is likely to fail in the near future (where 1 represents failure). The total number of samples in $\boldsymbol{\mathcal{D}}_{tr}$ is n and the set of their index is denoted as $\mathcal{N} = \{1, 2, ..., n\}$. $\boldsymbol{X}_{te}$ is the set of sample data in the testing set. Both $\boldsymbol{X}_{tr}$ and $\boldsymbol{X}_{te}$ are the sample pairs that contain the inverter operational data collected from their respective start and end time slots. The initial paired sample data are transformed into a lower-dimensional latent space using a CNN-based encoder layer. Subsequently, the Euclidean distance is used as the similarity function between each pair of samples. For a paired sample (x_i^s, x_i^e), the similarity can be calculated as follows:

$$\Psi(Z_i^s, Z_i^e) = \Psi(\phi(x_i^s), \phi(x_i^e))$$

$$= \|\phi(x_i^s) - \phi(x_i^e)\|_2 \tag{5}$$

where, $\phi(x_i^s)$ represents the transformation of the CNN-based encoder. $\Psi(Z_i^s, Z_i^e)$ denotes similarity distance and it will be the input of the contrastive loss, which will be introduced in the next section. The SCNN structure is designed to extract contrasting features of inverters across different time periods, specifically between the start and end time slots. These extracted features become integral components of the feature space utilized by the classification model.

C. Domain Knowledge Integrated Network

When utilizing the SCNN structure to discern differences between various periods, it's noted that the method does not fully leverage all inverter operational data; it primarily focuses on contrasting data from the start and end time slots. To enhance the model, particularly in scenarios where failure samples are limited, integrating domain knowledge into the network has proven beneficial.

To overcome these challenges, an enhanced SCNN model has been developed by integrating domain knowledge, realized by the domain knowledge integrated network, with the existing SCNN framework. The domain knowledge integrated network employs a multi-layered fully connected neural network to incorporate long-term features rich in domain knowledge. The selected features include the R_{e}, which assesses inverter stress due to prolonged overloading, and the R_{a}, which evaluates the long-term operational status and potential product defects that may cause frequent faults, even shortly after installation. Additionally, weather-related features such as peak high temperature (T_{H}), low temperature (T_{L}), and severe event data (C_{SE}) are incorporated as input variables. The output of this network is then amalgamated with the latent feature space of the SCNN, culminating in a comprehensive final feature set. Using this enriched feature set, a classification model is constructed using the multi-layer perception model, guided by label data derived from inverter failure records as shown in Fig. 6. Once trained, this integrated structure is equipped to effectively perform failure monitoring and prognosis for inverters.

D. Loss function of the framework

The proposed model is trained using a composite loss function comprising two key components: the contrastive Loss function adopted for the SCNN module and the cross-entropy loss function addressing the overarching classification task. To capture distinctive features of inverters across varying time frames, we employ the contrastive loss function [35]. This loss function aims to maximize the dissimilarity between start and end periods in case of inverter failure while minimizing this dissimilarity if the inverters are operational. Its mathematical formulation is given as:

$$L_{\mathrm{C}}\left(\boldsymbol{\mathcal{D}}_{tr}^{q}\right) = \sum_{s_j \in \boldsymbol{\mathcal{D}}_{tr}^{q}} \left[\mathbb{1}\left[y_i^t r = 0\right] \cdot \Psi(Z_i^s, Z_i^e)\right. \tag{6}$$

$$\left. + \mathbb{1}\left[y_i^t r \neq 0\right] \cdot \max\left(0, \epsilon - \Psi(Z_i^s, Z_i^e)\right)\right], \tag{7}$$

where $\boldsymbol{\mathcal{D}}_{tr}^{q}$ denotes the batch q from $\boldsymbol{\mathcal{D}}_{tr}$, and $s_j = (x_i^{tr,s},\ x_i^{tr,e},\ y_i^{tr})$ denotes a sample pair with the output label in $\boldsymbol{\mathcal{D}}_{tr}^{q}$. The hyperparameter ϵ defines the minimum separation between operational states across time periods for the target inverter.

Complementing the functionality of the contrastive loss function, the weight-regularized cross-entropy loss function is introduced to gauge the overall performance of the framework:

$$L_{\mathrm{E}}(\boldsymbol{\mathcal{D}}_{tr}^{q}) = -\sum_{s_j \in \boldsymbol{\mathcal{D}}_{tr}^{q}} \left[\log p(y_i^t r = 1 | x_i^{tr,s}, x_i^{tr,e})\right.$$

$$\left. + \log p(y_i^t r = 0 | x_i^{tr,s}, x_i^{tr,e})\right] + \lambda^T ||\mathbf{w}||^2 \tag{8}$$

where $\lambda^T ||\mathbf{w}||^2$ represents the regularization item. Then, the total loss function of the framework can be expressed as:

$$L_{\mathrm{T}} = \lambda L_{\mathrm{C}} + \beta L_{\mathrm{E}} \tag{9}$$

where, λ and β are weights of the two loss functions. This comprehensive loss formulation enables the framework to enhance classification performance while optimizing the extraction of contrasting features by the SCNN module.

V. Numerical Results

This section presents numerical case studies to validate the proposed model's performance. We specifically consider inverter working time for the sample selection, choosing only inverters that have operated for more than one year. Acknowledging the inevitable presence of noise in the data, we conduct data processing to ensure the model's stability and accuracy. First, during data processing, we discard the operational data from sunrise and sunset periods and focus on times of higher irradiance during the day, such as from 9 am to 3 pm. This helps mitigate the impact of abnormal inverter operation behaviors due to insufficient irradiance. Second, we employ interpolation to repair short-term missing data, discarding samples with prolonged missing data. These criteria ensure the reliability of our input data. Consequently, the testing dataset incorporates 44 failed inverter samples and 139 functional inverter samples.

A. Two-stage shock-based failure identification results

In order to validate the effectiveness of the proposed shock-based failure identification model, it was applied to the selected PV inverter dataset. Using the GMM clustering method, the model initially groups the failure samples into three distinct categories: Group A with 5 inverters, Group B with 9 inverters, and Group C with 21 inverters. In the second stage, the GMM clustering method, set to three clusters, was reapplied as previously described. The outcomes of this process are illustrated in Fig. 7. To enhance understanding, we have marked the proportions of failure samples in each cluster. This step highlights the percentage of failed units within each group and addresses the significant disparity between normal and failed inverter counts. In the first sub-cluster analysis, all failure samples were predominantly assigned to Group 1, comprising over 50%. Although this percentage does not reach 100%, the substantial presence of failed inverters in Group 1, which is nearly half, indicates a pronounced demarcation. In

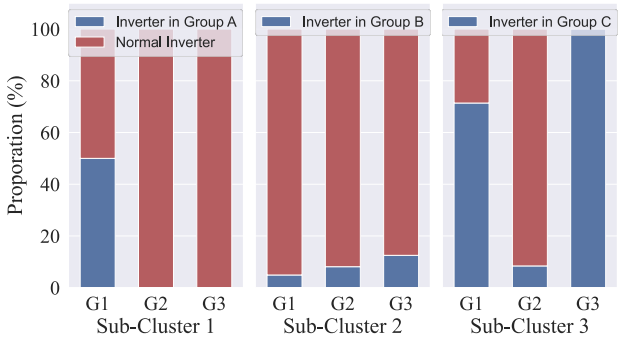

Fig. 7. The results of the two-stage shock-based failure identification method

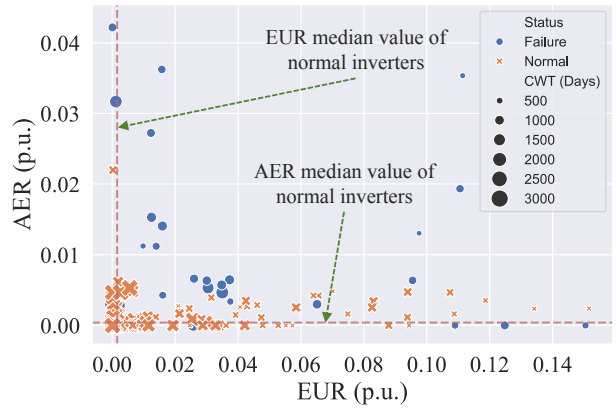

Fig. 8. The illustration of R_{a}, R_{e}, and CWT across various inverter samples

the third sub-cluster analysis, the failure samples were significantly found in Groups 1 and 3, with a minimal presence in Group 2, less than 9%, demonstrating a clear separation. However, the second sub-cluster analysis presented similar low proportions across all three groups, with average percentage less than 10%, posing a challenge in effectively identifying the failure samples. The clustering results from these three groups suggest that inverters in Group B are more likely to be shock-based failures. Consequently, these inverters will be excluded from further analyses in our study.

B. Performance of PV Inverter Failure Prognostic Model

1) Discussion on domain knowledge features: Based on the filtered dataset, the performance of the inverter failure prognostic model is discussed in this section. First, the comparison of two key features: R_{a}, R_{e}, and CWT across various inverter samples are showcased in the Fig. 8. Notably, the figure reveals that failure samples generally exhibit higher average R_{a} values compared to normal inverters, indicating a higher incidence of abnormal operation during their operational period. Furthermore, the scatter plot suggests that inverters with higher R_{e} values often have a shorter operational lifespan, a pattern observable in both failed and operational inverters. Specifically, for the failed samples, a clear correlation exists between higher R_{e} values and reduced working times. However, this trend is less definitive for normal inverters, as their CWT values are "censored" and do not represent their full lifespan. To

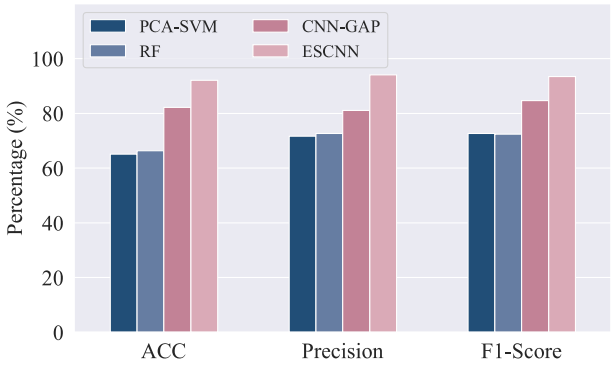

Fig. 9. The results comparison among ESCNN, CNN-GA, PCA-SV, and RF four models

further elucidate this relationship, the survival analysis can be found in our related work on the PV inverter stress factors analysis [36].

2) Simulation results comparison and analysis: To address the imbalance in the dataset, we augmented samples using the rolling horizon method for the failure inverter, generating multiple samples from start and end periods. Pair samples were then created by combining data from both timeframes, enhancing subsequent analyses and model training. Additionally, we employed a 5-fold cross-validation approach and grid searching method to ensure robustness and reliability in assessing the predictive model's performance and determining optimal parameters such as learning rate, batch size, and margin (ϵ). To showcase the effectiveness of our proposed model, we compared it with three existing solar inverter failure detection models: modified CNN with global average pooling (CNN-GAP) [21], support vector machine combined with principal component analysis (PCA-SVM) [24], and random forests (RF) [19]. In addition to these models, we explored SCNN models with long-term features, referred to as enhanced-SCNN (ESCNN), within our simulations. Comparative results for these four models are presented with common evaluation metrics, namely accuracy, precision, and F1-score, in Fig. 9.

The bar chart depicts the average outcomes from the 5-fold validation process. Notably, the ESCNN model achieves an accuracy rate exceeding 94%, with its lowest metric, the F1-score, also nearing 92%. Across all three metrics, ESCNN consistently outperforms the other three models. The contrasting features extracted from a well-trained model and transformed using t-SNE technology are depicted in Fig. 10. In comparison, the CNN-GAP model demonstrates effectiveness second only to the ESCNN model, achieving an accuracy of 88%. This may stem from its well-designed CNN-based structure, proficient in capturing features from extensive operational data. However, CNN-GAP's lack of focus on capturing differential features limits its ability to recognize failure samples with subtle degradation symptoms. Despite PCA technology excelling in feature reduction tasks, PCA-SVM struggles, likely due to inputs from large volumes of operational data. Enhancing performance would require extensive feature generation and selection, which poses challenges in terms of difficulty and time

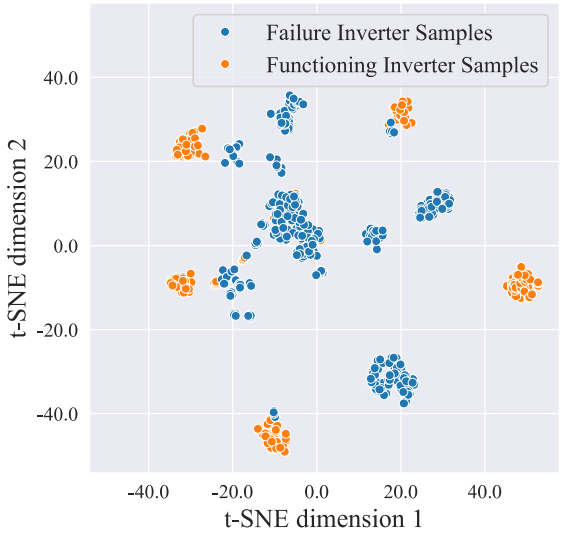

Fig. 10. The t-SNE visualization results of the contrasting features of augmented samples

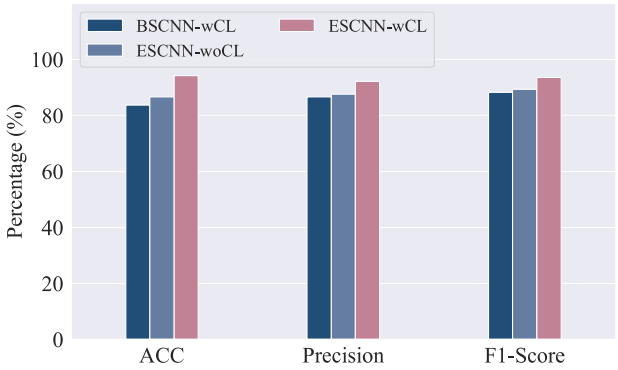

Fig. 11. The evaluation metrics of the three ablation groups

consumption.

3) Ablation study on ESCNN model: To comprehensively assess the impact of various components within the ESCNN model, an ablation study was conducted. Three distinct scenarios were meticulously chosen for analysis. These scenarios include the ESCNN model with contrastive loss (ESCNN-wCL), the ESCNN model lacking domain knowledge integration network (SCNN-wCL), and the ESCNN model without contrastive loss (ESCNN-woCL). Evaluation metrics for these scenarios are illustrated in Fig. 11. A comparison between ESCNN-woCL and ESCNN-wCL highlights the effects of removing the contrastive loss function, resulting in a 4.6% decrease in overall accuracy. It is shown that the integration of the contrastive loss function enhances the SCNN module's capacity to discern deviations in inverter operation statuses. Similarly, contrasting SCNN-wCL with ESCNN-wCL reveals a 5.6% reduction in overall accuracy following the exclusion of domain knowledge features. Detailed analysis of confusion matrices, as depicted in Fig. 12, reveals that the decline in accuracy stemming from the absence of the domain knowledge integrated network primarily arises from misclassifying failure samples as functioning ones. Domain knowledge features play a pivotal role in identifying samples lacking strong degradation symptoms, a task that the SCNN model alone struggles with. In summary, the ablation study underscores the critical importance of both the contrastive loss function and the domain knowledge integrated network within the ESCNN model.

In analyzing the prediction results of the ESCNN model, it becomes evident that a minority of inverters are misclassified. This misclassification can be attributed to the fact that not all failing inverters exhibit distinct degradation patterns. The varying intensities of failure symptoms contribute to the misidentification of certain inverters. While the proposed model aims to capture symptoms comprehensively, it is acknowledged that not every failure is predictable. Furthermore, prolonged operational stress may cause some inverters in normal conditions to display signs of degradation. Although these instances do not necessarily indicate immediate failure, they signal potential future issues. Such findings emphasize the importance of continuous monitoring and draw attention to inverters that may require future intervention.

4) Discussion on generalization of domain-feature integrated network: Domain features could vary significantly across different inverter datasets, which is a common challenge for data-driven methods. In addressing variations in domain knowledge-related features among inverters from different solar companies, a practical strategy is retraining the model separately using historical data from each company. However, for companies with limited historical data, this strategy could be infeasible. Instead, employing transfer learning by fine-tuning a pre-trained public model can be more feasible and effective. This approach helps reconcile feature discrepancies and improves model performance, adapting it to the unique conditions of each company despite data constraints. Additionally, variations in feature distributions among inverters from the same company can also be the challenge to the generalization of domain-feature integrated network. This problem can be managed by initially training a uniform model using the entire dataset and subsequently updating it with new data from specific target inverters. Over time, iterative updates incorporating new insights from target inverter data enhance the model's predictive accuracy. Furthermore, for companies seeking to incorporate additional features tailored to local weather conditions or specific inverter characteristics, our proposed method offers a flexible framework for customization. Integrating specialized features refines the model to better address unique operational scenarios while maintaining broad applicability. The solutions discussed above can mitigate the influence and strength the generalization ability of our model.

C. ESCNN-based PV Inverter Failure Monitoring

To enable real-time awareness of inverter operation status and provide operators with precise prognostic information, we develop a failure monitoring method based on well-trained ESCNN model. This method comprises two key components: firstly, the organization of inverter feature data utilizing the

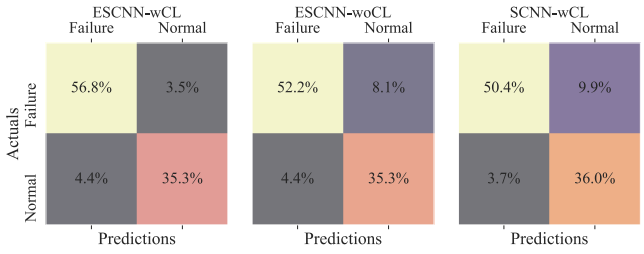

Fig. 12. The confusion matrices of the classification results from three ablation scenarios

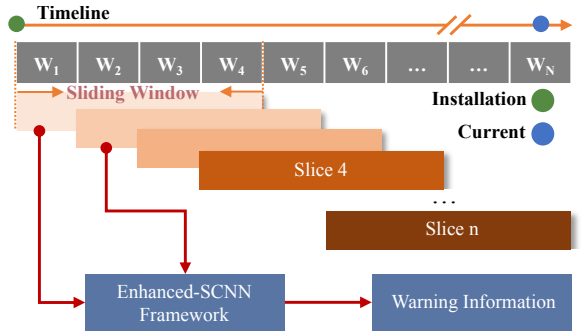

Fig. 13. Inverter measurements organization based on sliding window method

sliding window technique; secondly, the actual failure monitoring using the ESCNN model. Fig. 13 illustrates this ESCNN-based approach to realize the inverter failure monitoring.

In this method, historical operational data and extracted feature data of inverters are segmented into multiple data slices using the sliding window method. This method is configured with a window size of one month and a step size of one week. To generate warning signals for a specific data slice, the initial data slice and subsequent slices are inputted into the ESCNN model. As new inverter measurements are accumulated, they are fed into the model, enabling the continual update and retrieval of warning information. By aggregating all warning data from the initial time period to the present, a comprehensive view of the inverter's operational status throughout its lifetime is obtained.

Fig. 14 presents a case study of an inverter. In this analysis, the focus is not solely on weekly warning data. Instead, the metric of average warning counts over a selected period, such as three months, is considered. This metric calculates the mean frequency of warnings by accumulating warnings from the current time back to the specified period. The adoption of this metric is motivated by the need to reduce the influence of short-term external disturbances, which may lead to transient "false" warnings. By analyzing data over an extended period, more insightful and reliable information about the inverter's performance is obtained. However, this approach inevitably introduces a delay in the identification of inverter failures, creating a trade-off between result reliability and timeliness. For this particular case, a three-month period has been selected for analysis. Furthermore, a cumulative average warning count is calculated to reveal the general trend of warnings over time. This approach aims to strike a balance between providing timely alerts and minimizing false alarms caused by temporary external factors.

The analysis of Fig. 14 reveals that following its initial installation, the inverter exhibited a fault-free operation, indicating optimal performance according to the framework assessment. However, intermittent weak warning signals occurred approximately one year and eighteen months post-installation. Notably, these signals did not persist continuously, resulting in a return to baseline operational metrics. Such sporadic warnings may have been triggered by inadvertent disturbances attributed to the inverter's operation. Continuing its operation for nearly three additional years, the inverter demonstrated a recurrence of warning signals, coinciding with a gradual escalation in metric values. Intermittent output fluctuations, indicative of incipient degradation, characterized this phase. Ultimately, after a duration of seven weeks, the inverter suffered a critical failure.

The utilization of warning information for early prognosis of inverter failure involves setting an appropriate threshold, which can vary among solar companies, factoring in labor resource considerations. A lower threshold can facilitate earlier detection of anomalies, albeit with an increased likelihood of false positives due to temporary disturbances. Conversely, a higher threshold enhances reliability but may delay the identification of inverter failures, potentially impacting replacement timelines. The determination of this threshold is thus dependent on individual solar companies, taking into account factors such as supply chain dynamics, repair crew scheduling, and other operational considerations.

The current model, trained on a diverse range of inverter samples, aims for optimal performance across all cases. However, due to the varying characteristics of different inverters, a one-size-fits-all approach may not effectively capture the specific features of each inverter. To address this, future work will involve refining the general ESCNN model by training it specifically on data from the target inverter or inverters with similar operational features. Customizing the model in this way could enhance the accuracy of failure prognosis for specific inverters.

VI. Conclusion

This paper introduced a data-driven PV inverter failure prognostic model that leveraged real inverter measurements. The model employed the ESCNN to detect deviations in operational status that were indicative of inverter degradation. It integrated multiple operational and weather-related factors, informed by domain knowledge, to represent inverter stressors and status accurately. We proposed a monitoring scheme using the ESCNN-based model for effective inverter failure detection. Test results on real inverter data confirmed the model's effectiveness and its application in failure monitoring demonstrated practical utility. Currently, the proposed model was tested on a single private inverter dataset due to the lack of public datasets. This limitation could hinder a comprehensive evaluation of the model across diverse datasets. To solve this, we plan to convert the proposed model into a public tool intended for solar company usage and evaluate the model based

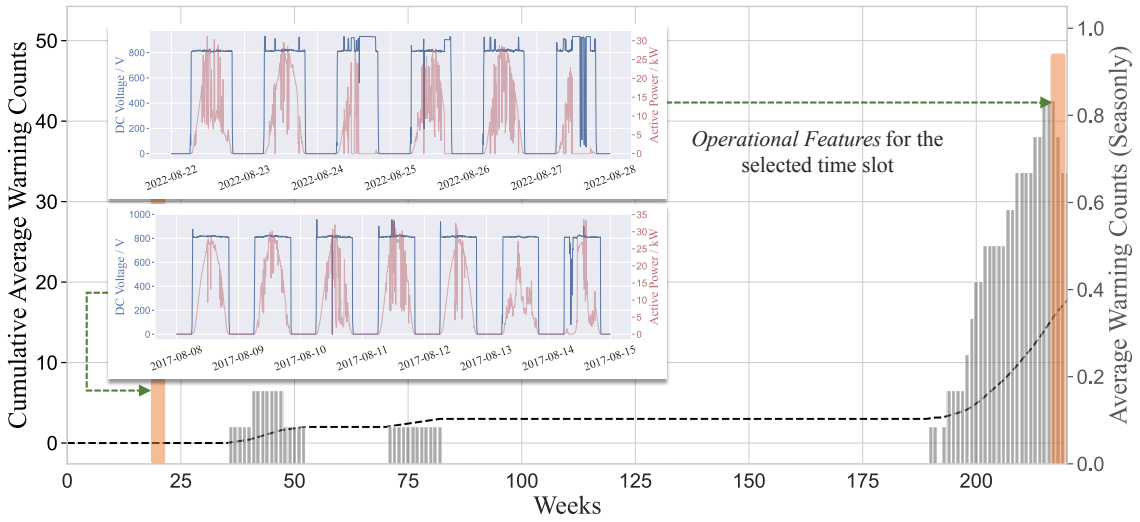

Fig. 14. The results of failure prognostic monitoring for a single inverter over the lifetime

on the feedback and performance metrics gathered from these industry users. Additionally, future research will focus on customizing the general model, trained on the overall dataset, by fine-tuning it with data from specific inverters to enhance its precision for single inverters. The combination of the ESCNN model with a transfer learning mechanism will also be explored to address scenarios commonly faced by solar companies, such as the lack of historical data needed to train a model from scratch.

REFERENCES

[1] Y. Song and B. Wang, "Survey on reliability of power electronic systems," *IEEE transactions on power electronics*, vol. 28, no. 1, pp. 591–604, 2012.

[2] A. Sangwongwanich, Y. Yang, D. Sera, and F. Blaabjerg, "Lifetime evaluation of grid-connected pv inverters considering panel degradation rates and installation sites," *IEEE transactions on Power Electronics*, vol. 33, no. 2, pp. 1225–1236, 2017.

[3] W. Qiu, A. Yadav, S. You, J. Dong, T. Kuruganti, Y. Liu, and H. Yin, "Neural networks-based inverter control: Modeling and adaptive optimization for smart distribution networks," *IEEE Transactions on Sustainable Energy*, pp. 1–11, 2023.

[4] *IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces*, IEEE Std 1547-2018, 2018.

[5] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggedakis, *et al.*, "Data-driven prediction of battery cycle life before capacity degradation," *Nature Energy*, vol. 4, no. 5, pp. 383–391, 2019.

[6] W. Mao, J. He, and M. J. Zuo, "Predicting remaining useful life of rolling bearings based on deep feature representation and transfer learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 4, pp. 1594–1608, 2019.

[7] F. Ding, Z. Tian, F. Zhao, and H. Xu, "An integrated approach for wind turbine gearbox fatigue life prediction considering instantaneously varying load conditions," *Renewable energy*, vol. 129, pp. 260–270, 2018.

[8] U. Scheuermann, R. Schmidt, and P. Newman, "Power cycling testing with different load pulse durations," in *7th IET International Conference on Power Electronics, Machines and Drives (PEMD 2014)*, pp. 1–6, Iet, 2014.

[9] A. Nagarajan, R. Thiagarajan, I. L. Repins, and P. L. Hacke, "Photovoltaic inverter reliability assessment," tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2019.

[10] A. Testa, S. De Caro, and S. Russo, "A reliability model for power mosfets working in avalanche mode based on an experimental temperature distribution analysis," *IEEE Transactions on power electronics*, vol. 27, no. 6, pp. 3093–3100, 2011.

[11] I. Kovačević, U. Drofenik, and J. W. Kolar, "New physical model for lifetime estimation of power modules," in *The 2010 International Power Electronics Conference-ECCE ASIA-*, pp. 2106–2114, IEEE, 2010.

[12] H. Berg and E. Wolfgang, "Advanced igbt modules for railway traction applications: Reliability testing," *Microelectronics Reliability*, vol. 38, no. 6-8, pp. 1319–1323, 1998.

[13] S. Bouguerra, M. R. Yaiche, O. Gassab, A. Sangwongwanich, and F. Blaabjerg, "The impact of pv panel positioning and degradation on the pv inverter lifetime and reliability," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 9, no. 3, pp. 3114–3126, 2020.

[14] H. Wang and F. Blaabjerg, "Reliability of capacitors for dc-link applications in power electronic converters—an overview," *IEEE Transactions on industry Applications*, vol. 50, no. 5, pp. 3569–3578, 2014.

[15] O. Gandhi, C. D. Rodríguez-Gallegos, N. B. Y. Gorla, M. Bieri, T. Reindl, and D. Srinivasan, "Reactive power cost from pv inverters considering inverter lifetime assessment," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 2, pp. 738–747, 2018.

[16] P. D. Reigosa, H. Wang, Y. Yang, and F. Blaabjerg, "Prediction of bond wire fatigue of igbts in a pv inverter under a long-term operation," *IEEE Transactions on Power Electronics*, vol. 31, no. 10, pp. 7171–7182, 2015.

[17] J. Fan, Y. Wang, T. Si, and S. Chang, "Health assessment of two-stage inverter based on multiple evaluation indexes feature optimization model and mahalanobis distance," in *2020 11th International Conference on Prognostics and System Health Management (PHM-2020 Jinan)*, pp. 122–125, IEEE, 2020.

[18] B. Zhang and Y. Gao, "Igbt reliability analysis of photovoltaic inverter with reactive power output capability," *Microelectronics Reliability*, vol. 147, p. 115073, 2023.

[19] L. Kou, C. Liu, G.-w. Cai, J.-n. Zhou, Q.-d. Yuan, and S.-m. Pang, "Fault diagnosis for open-circuit faults in npc inverter based on knowledge-driven and data-driven approaches," *IET Power Electronics*, vol. 13, no. 6, pp. 1236–1245, 2020.

[20] A. Malik, A. Haque, and K. S. Bharath, "Deep learning based fault diagnostic technique for grid connected inverter," in *2021 IEEE 12th Energy Conversion Congress & Exposition-Asia (ECCE-Asia)*, pp. 1390–1395, IEEE, 2021.

[21] W. Gong, H. Chen, Z. Zhang, M. Zhang, and H. Gao, "A data-driven-based fault diagnosis approach for electrical power dc-dc inverter by using modified convolutional neural network with global average pooling and 2-d feature image," *IEEE Access*, vol. 8, pp. 73677–73697, 2020.

[22] J. Fang, F. Cheng, and X. Meng, "Fault diagnosis for photovoltaic inverter based on attention recurrent neural network," in *2021 International Conference on Power System Technology (POWERCON)*, pp. 2322–2330, IEEE, 2021.

[23] D.-E. Kim and D.-C. Lee, "Fault diagnosis of three-phase pwm inverters using wavelet and svm," in *2008 IEEE International Symposium on Industrial Electronics*, pp. 329–334, IEEE, 2008.

[24] W. Yuan, T. Wang, and D. Diallo, "A secondary classification fault diagnosis strategy based on pca-svm for cascaded photovoltaic grid-connected inverter," in *IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society*, vol. 1, pp. 5986–5991, IEEE, 2019.

[25] A. Sangwongwanich, Y. Yang, D. Sera, F. Blaabjerg, and D. Zhou, "On the impacts of pv array sizing on the inverter reliability and lifetime," *IEEE Transactions on Industry Applications*, vol. 54, no. 4, pp. 3656–3667, 2018.

[26] R. S. Faranda, H. Hafezi, S. Leva, M. Mussetta, and E. Ogliari, "The optimum pv plant for a given solar dc/ac converter," *Energies*, vol. 8, no. 6, pp. 4853–4870, 2015.

[27] B. Burger and R. Rüther, "Inverter sizing of grid-connected photovoltaic systems in the light of local solar resource distribution characteristics and temperature," *Solar Energy*, vol. 80, no. 1, pp. 32–45, 2006.

[28] K. Zipp, "Why array oversizing makes financial sense," *Solar Power World*, vol. 12, 2018.

[29] Y. Luo, L. Liu, Y. Yuan, Z. Wang, F. Qiu, S. Zhao, and M. Yildirim, "Data-driven assessment of operational and harsh environmental stress factors for solar inverters using field measurements," in *2024 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, IEEE, 2024.

[30] J. He, A. Sangwongwanich, Y. Yang, and F. Iannuzzo, "Lifetime evaluation of three-level inverters for 1500-v photovoltaic systems," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 9, no. 4, pp. 4285–4298, 2020.

[31] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah, "Signature verification using a "siamese" time delay neural network," *Proc. Adv. Neural Inf. Process. Syst*, vol. 6, 1993.

[32] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.

[33] P. Bedi, N. Gupta, and V. Jindal, "Siam-ids: Handling class imbalance problem in intrusion detection systems using siamese neural network," *Procedia Computer Science*, vol. 171, pp. 780–789, 2020.

[34] M. Heidari and K. Fouladi-Ghaleh, "Using siamese networks with transfer learning for face recognition on small-samples datasets," in *2020 international conference on machine vision and image processing (MVIP)*, pp. 1–4, IEEE, 2020.

[35] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1, pp. 539–546, IEEE, 2005.

[36] Y. Luo, L. Liu, Y. Yuan, Z. Wang, F. Qiu, S. Zhao, and M. Yildirim, "Data-driven assessment of operational and harsh environmental stress factors for solar inverters using field measurements.," in *2024 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, IEEE, 2024.