On the Use of Artificial Intelligence for High Impedance Fault Detection and Electrical Safety

Shiyuan Wang, *Student Member, IEEE*, and Payman Dehghanian, *Senior Member, IEEE*

***Abstract*—Accidents caused by faults on overhead power lines have been more frequently reported under extreme weather conditions and may strongly threaten the safety and stability of the power grids, e.g., massive wildfires caused by the electrical arcs or lines getting in touch with vegetation, relay miss-operations, etc. It has been widely recognized that the electric safety concerns engendered by overhead line faults have to be timely and properly addressed to minimize the subsequent risks and damages. The existing monitoring devices and protective relays can barely detect high impedance faults (HIFs) and are unable to warn the system operators until serious abnormalities or damages are observed. Aiming at avoiding the damaging consequences of HIFs, an online monitoring system embedded with machine learning analytics is proposed that ensures a fast and accurate detection of HIFs in power systems. The performance of the proposed artificial intelligence engine is tested under a variety of simulated conditions and the numerical results demonstrate its efficacy and superiority over the state-of-the-art advancements.**

***Index Terms*—Convolutional neural network (CNN), electrical safety, event detection, feature extraction, high impedance fault (HIF), wavelet transform (WT).**

I. Introduction

ELECTRICITY has dramatically changed our daily lives and fuels our modern society. The electricity grid is constantly exposed, and yet vulnerable, to a wide range of threats, some foreseeable and some unpredictable and random in nature. One safety-threatening disruption in power systems is recognized as high impedance faults (HIF), the detection of which has long remained a challenging concern in the electric industry. HIFs can cause "arcs" or "flashover" from the wires, through the air, to the neighboring trees, other vegetation or equipment, where it can cause fires, injuries, or even fatalities [1]. A life-threatening example is the constant exposure and contact of a power line with a tree branch during high-wind conditions, which can threaten homes in residential neighborhoods and spark wildfires in rural areas. Such types of faults are commonly caused by undesired contacts with bare energized electrical conductive parts, the high-impedance nature of which significantly restricts the flow of fault current to a level hard to be detected by the overcurrent protective relays [2]. In particular, the existing commercial microprocessor-based protective relays activate a tripping decision when the electrical measurements are observed well beyond the detection threshold; however, an unsatisfactory performance is observed in their detection logic when facing the HIF events. Electrical safety studies have shown that conventional protection schemes detected and cleared only 17.5% of staged HIFs [3]–[5]. Therefore, HIF detection and localization in electrical power systems yet remain a safety-threatening challenge for power system protection engineers, a fast and accurate solution to which is urgently needed to limit the safety risks, prevent power grid operation violations, and save human lives [6], [7].

Many research works have been conducted in diagnosing HIFs, each with some advantages and disadvantages; reference [8] harnessed the high-frequency content in real vegetation fault signatures and proposed a method for the detection of distinct and very small-current HIFs. In [9], an HIF detection approach for power distribution networks is suggested using fuzzy logic control that evaluates the 3 rd and 5th harmonics in the electrical current signals. However, due to the different HIFs characteristics compounded by the existence of harmonics and noises in the power waveforms, their accuracy and speed performance may be compromised in real-world scenarios. A scheme to detect HIFs using the Time–Time (TT) transform that analyzes and determines the fault wave patterns is introduced in [10]; this scheme can handle low signal to noise ratio (SNR) in power waveforms through a threshold selection procedure using unscented transformation. While it can be applied to microgrids with different ratings and structures, the detection threshold must be tailor-made and adjusted appropriately. Cui *et al.* [11] presented a variable-importance-based feature selection method to identify HIFs from a large pool of signal signatures; this feature selection scheme utilizes the discrete Fourier transform (DFT) and Kalman filter (KF) for harmonics coefficient estimation and HIF duration and magnitude measurement.

With the rapid advancements in artificial intelligence (AI), many researchers have implemented various HIF detection techniques through a variety of machine learning technologies. Sarwar *et al.* [1], [12], [13] utilized waveform pattern analysis and support vector machine to classify and ultimately detect HIFs in power grids. In [4], semi-supervised learning and probabilistic learning are used for HIF detection and localization, revealing

Manuscript received June 30, 2020; accepted August 13, 2020. Date of publication August 18, 2020; date of current version November 19, 2020. Paper 2020-ESafC-0765, approved for publication in the IEEE Transaction on Industry Applications by the Electrical Safety Committee of the IEEE Industry Application Society. This work was supported in part by the National Science Foundation (NSF) under Grant ICER-2022505. *(Corresponding author: Shiyuan Wang.)*

The authors are with the Department of Electrical, and Computer Engineering, The George Washington University Washington, DC 20052 USA (e-mail: shiyuan1225@gwu.edu; payman@gwu.edu).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIA.2020.3017698

promising detection accuracy, but with a compromised response time of half a second. In [14], an HIF detection approach using empirical mode decomposition combined with an artificial neural network (ANN) is proposed, where the HIF detection and classification are achieved through predominant harmonic signatures caused by HIFs in the electrical signals. Discrete wavelet transform (WT) is applied in [15] to monitor the high-frequency components and long short term memory to detect HIFs, revealing detection accuracy of 90% under scenarios with clean (not noisy) measurements. An HIF detection approach using discrete wavelet transformations and back propagation artificial neural network (BP-ANN) is introduced in [16] and tested using often-noisy real measurement data from a substation, where the detection accuracy is reported 76%. While the accurate and swift HIF detection is a yet to be solved challenge in the electric industry, the state-of-the-art literature has demonstrated that AI technologies offer a yet-untapped potential in detecting HIFs and improving electrical safety by enabling a timely warming notification to the system operator and activating trip signals if needed.

Inspired by the promising feature extraction and pattern recognition capabilities through WTs and the wide adoption of machine learning analytics in electric power grids [17]–[22], we propose an AI solution that unlocks the full potential in computing at the edge for HIF detection and classification, thereby improving the electrical safety. An improved HIF model is first proposed for generating a large pool of HIF waveforms. A modified Gabor WT is next proposed for feature extraction and pattern recognition from the electrical waveforms during HIF scenarios. An efficient convolutional neural network (CNN) engine is eventually developed to process the extracted features and make the detection and classification decisions. The proposed solution offers a promising accuracy and response time in detecting and classifying low-intensity HIFs under noise inferences. The suggested functionality can be installed within the existing sensors, protective relays and other intelligent electronic devices (IEDs) with online monitoring capabilities, thereby rendering a cost-effective solution technology.

This article is organized as follows. Section II introduces the background information on HIF modeling, WT, and the pattern classification and recognition through CNNs. Section III presents the improved HIF model and the proposed HIF detection and classification scheme, which is consisted of, first, power waveform feature extraction through a pseudo-continuous quadrature wavelet transform (PCQ-WT), and, second, event detection and classification via CNN. Case studies and numerical results are analyzed in Section IV. Section V eventually concludes this article and provides directions for future developments.

II. Background and Motivation

A. HIF Modeling

A single-line diagram of an HIF event in a radial distribution system is illustrated in Fig. 1(a). The sending node is modeled by an ideal ac source; Z_1 and Z_2 are known impedance values and can be estimated according to the system topology and

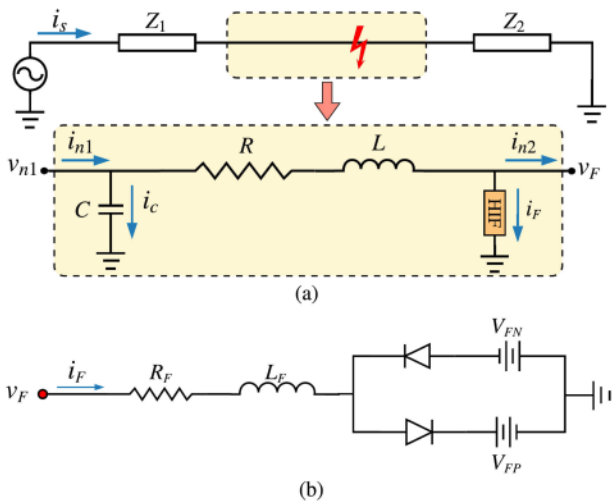


Fig. 1. Single-phase diagram for the HIF model in [4] and [5]. (a) Single-phase diagram for HIF modeling. (b) Two antidiode HIF model.

operating conditions; R, L, and C stand for the per unit length resistance, inductance, and capacitance of the line. By applying Kirchhoff's voltage law to the dashed circle in Fig. 1(a), the following equation is derived:

$$v_{n1} = \delta\left(R(i_{n1} - i_c) + L\frac{\mathrm{d}(i_{n1} - i_c)}{\mathrm{d}t}\right) + v_F \tag{1}$$

where v_{n1} and i_{n1} are the sending-terminal voltage and current; v_f stands for the voltage at the location of the fault; and δ is the fault distance from the sending-terminal. However, during an HIF event, v_F and δ are the values to be estimated. The current flow can be found through

$$i_{n1} = i_c + i_{n2} + i_F, \quad i_c = \delta\frac{\mathrm{d}v_{n1}}{\mathrm{d}t} \tag{2}$$

where i_c, i_{n2}, and i_F are, respectively, the currents flowing through the shunt capacitor, received at the end-terminal, and observed at the HIF branch.

Under some HIF scenarios, arcs can be observed when the air gap between the power line conductor and the high impedance object is energized. Once the imposed voltage magnitude is higher than the voltage (breakdown voltage), there would be arc ignitions across the air gap. On the other hand, an arc extinguishes when the fault voltage is lower than the breakdown voltage. Therefore, the HIF current changes during each cycle, making its magnitude follow a nonlinear characteristic [23]. Based on the above properties of an HIF event, an HIF model is developed and has been widely used in [1], [4], [5], [9], [11], [14], as shown in Fig. 1(b). The fault voltage in this model can be written in the following format:

$$v_F = \begin{cases} R_F i_F + L_F\frac{\mathrm{d}i_F}{\mathrm{d}t} + V_{FP}, & i_F \geq 0 \\ R_F i_F + L_F\frac{\mathrm{d}i_F}{\mathrm{d}t} - V_{FN}, & i_F < 0 \end{cases} \tag{3}$$

where R_F and L_F are the HIF's resistance and inductance in series; V_{FP} and V_{FN} are the positive and negative arc voltages during HIF events, which archives the nonlinearity of HIFs. The simulated voltage and current waveforms are demonstrated in

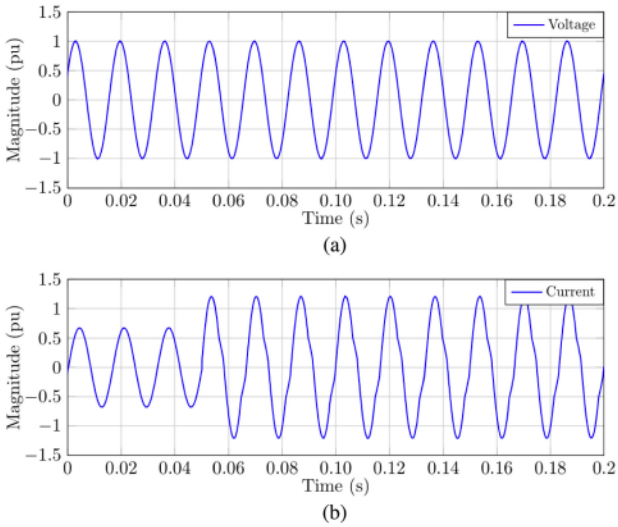


Fig. 2. Simulated voltage and current waveforms (with per unit values) captured at the sending-terminal based on [4]. (a) HIF voltage. (b) HIF current.

Fig. 2. It can be found that although there is almost no influence on the voltage waveform, the HIF event slightly distorts the current signal while the current magnitude has remained around the rated range; in other words, the HIF event cannot be detected easily by the conventional protective relays. One should note that, at the beginning of an HIF event, this HIF model can barely match the characteristics of the first period due to the build-up phenomenon [24]. However, following a few fundamental cycles, the HIF stabilizes and the fault parameters—R_F, L_F, V_{FP}, and V_{FN}—are approximated as constant values. Therefore, the HIF current waveform and the related shoulder phenomenon can still be characterized using the model presented in Fig. 1(b).

With the wide installation of phasor measurement units (PMUs) in the power grid, micro-PMUs (μ-PMUs) in power distribution systems [25] and many other IEDs with PMU functionalities, HIF detection can be achieved through such devices with high-precision and high-resolution measurements; that is, the case presented in Fig. 1(a) can be easily solved as long as both sending and receiving terminals are equipped with such devices that can ensure the availability of high-precision high-resolution electrical measurements. However, a full observation is very hard to achieve in every segment of the power network due to the cost limitations [26] and only the power waveforms from the upstream (sending-end) terminal are usually measured; meanwhile, the performance of the PMU surveillance and monitoring network can be compromised by communication failures, latencies, and cyber-attacks [27]. Therefore, it makes it very challenging for the HIF detection and classification scheme to operate as desired since the HIF current is viewed incremental by the upstream terminal, not large enough to violate the tripping thresholds in the protective relays, and are often mistakenly corresponded to the common load increments in the network.

B. Power Waveform Modeling

With the current waveforms carrying valuable information on the underlying phenomenon, one can evaluate the current waveform in each phase to examine the existence of HIFs. First, the current waveform can be generalized by

$$x(t) = A\cos(\omega t + \phi) \tag{4}$$

where A, ω, and ϕ are the magnitude, frequency, and phase angle in each phase. Although HIF phenomena are hard to detect through current amplitudes, such events often cause waveform abnormalities and distortions. One should note that the orders, magnitudes, and phase angles corresponding to the HIF-caused harmonics would be totally different under different combinations of HIF parameters. Thus, detecting HIF through analyzing certain orders of harmonics will be extremely challenging; this is even further exacerbated by the existence of noise or other harmonics, altogether could compromise the performance of the HIF detection schemes. To deeply investigate the waveform features, we expand the sampled current by the Fourier series as follows:

$$x(t) = A_1\cos(\omega_1 t + \phi_1) + \underbrace{\sum_{h=2}^{H} A_h\cos(\omega_h t + \phi_h)}_{\text{Harmonic Components}} \tag{5}$$

where h is the order of harmonics; H is the maximum order of harmonics limited by the sampling rate; A_h, ω_h, and ϕ_h are the hth order harmonic component's instantaneous magnitude, frequency, and phase angle, respectively. Also, $h = 1$ stands for the fundamental frequency component. One can see, in Fig. 2, that the fundamental magnitude is affected by both HIF and load variation events. Therefore, it is clear that the second term in (5) carries valuable information in assessing the HIF impacts, thus the main focus in the proposed data mining and pattern extraction process.

C. WT and Pseudo-Continuous WT

As the combinations of harmonic components contain valuable information on HIFs, it would be preferred to implement time-frequency analysis on the current waveforms. Short time Fourier transform (STFT) [28] or DFT through a sliding window [11] can be approached assisting the HIF detection process; this is because the STFT offers a high frequency resolution; however, this is achieved at the cost of a high computational complexity [29], [30] and it is hard to provide accurate detection performance utilizing the spectrum alone. On the other hand, WT, first, offers a promising computing efficiency compared to the STFT in time-frequency analysis, and, second, has been widely applied for feature extractions in many real-world applications [17], [31], [32], including research on HIF detection [16], [33], [34].

Fig. 3 illustrates a performance comparison of the STFT and WT. It can be observed, from the spectrum in Fig. 3(b) and (f), that the HIF current waveform is featured with more harmonic components at larger than 5th orders, where the performance of the HIF detection method in [9] will be easily compromised by such harmonic pollution or noise interferences. Meanwhile, comparing the scalograms in Fig. 3(c) and (d) with Fig. 3(g) and (h), one can note that the results from STFT provide very

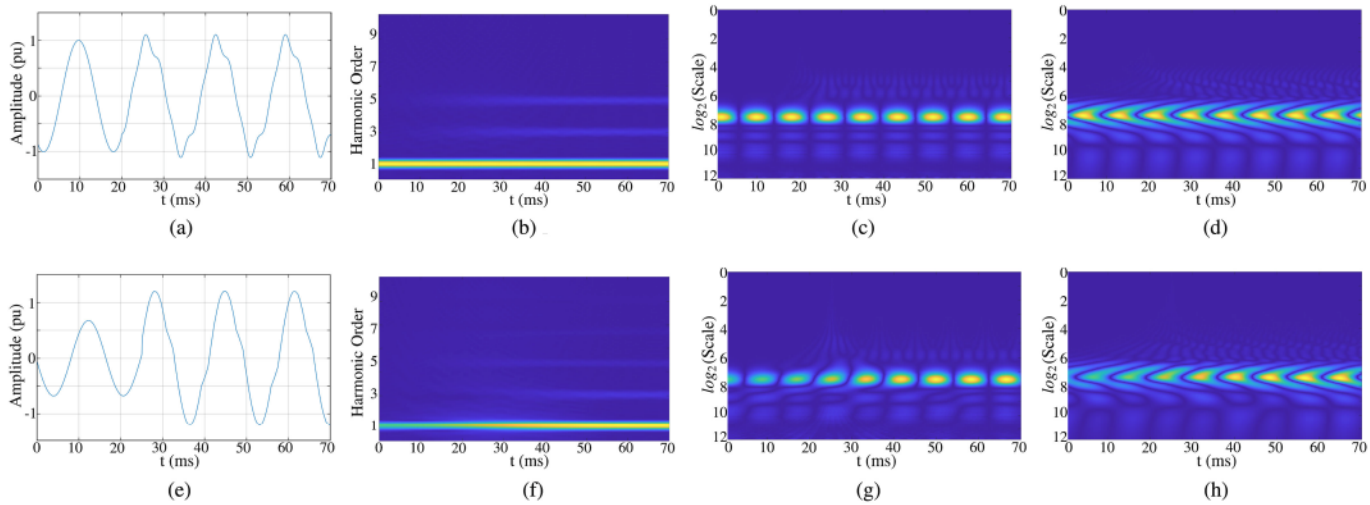


Fig. 3. Comparison of the STFT versus WT with: HIF-affected waveform and harmonic-injected waveform starting at $t = 20$ ms with harmonic orders $h = 3$ and 5 and magnitudes of 0.08 p.u. and 0.08 p.u., respectively. (a) Harmonic waveform. (b) STFT. (c) Morlet. (d) Daubechies 20. (e) HIF waveform. (f) STFT. (g) Morlet. (h) Daubechies 20.

detailed frequency information, and the scalograms from HIF-affected waveform reveal very obvious features compared to those from the harmonic-polluted waveforms; nevertheless, the accuracy in frequency measurements is compromised in WT. Therefore, a combination of WT and STFT is generally much desired. However, this could highly increase the computational burden of the detection mechanism, and in particular, challenge the online applicability of the HIF detection schemes. In order to archive a low-computing complexity and yet accurate design, we select WT alone in this article for power waveform feature extraction and HIF detection.

The WT can be seen as the computation process of the similarities between the signal of interest $x(t)$ and the selected wavelet. $x(t)$ is here the measured single-phase waveform. The WT calculation is defined as follows:

$$X(\omega|a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \Psi^* \left(\frac{t-b}{a} \right) dt \tag{6}$$

where a, b are the scaling factor and time shift; $\Psi(t)$ is the selected wavelet (called mother wavelet when $a = 1$ and $b = 0$), and "*" denotes the complex conjugate operator. With different values of a and b selected, $\Psi(\frac{t-b}{a})$ becomes the "daughter wavelets" of $\Psi(t)$ [35]–[37], and each value of a corresponds to a pseudo frequency. Typically, the WT uses discrete scaling factors $a_k = 2^i$, where i is integer. Thus, the extracted feature from each pseudo frequency is finite. To have more information extracted and waveform feature redundancy in the scalogram, a set of linearly increasing real numbers can be assigned to i, and the WT then becomes *pseudo-continuous* (PCWT). Here, the PCWT is defined as follows:

$$X[\omega|a_k,b_k] = \frac{1}{\sqrt{a}} \sum_{n=0}^{W-1} x[n] \Psi^* \left[\frac{nT_s - b_k}{a_k} \right] \tag{7}$$

where T_s is the sampling time, and W is the number of data sampled in the buffer. Choosing a proper type of wavelet and adjusting the parameters correctly are very critical as they will affect the covered range for the pseudo frequency and the PCWT performance. Besides, the computing complexity can be reduced and the response time efficiency can be improved by using well-designed parameters since the online feature extraction process requires uncongested real-time data streams.

D. Convolutional Neural Networks

CNNs are among deep machine learning techniques that have been proven very effective in processing image-related tasks [38]–[40]. Therefore, once the scalograms from WT are generated, CNNs can be used as the HIF event detection engine through classifying the scalograms. The CNN training process mainly involves feature/representation learning for feature detection and classification. In general, the execution of CNN for HIF detection and classification is achieved through a set of cross-correlation assessments as follows:

$$s^p(m,n) = \sum_u \sum_v \sum_w \boldsymbol{I}^u(m+v, n+w) \boldsymbol{K}^p(v,w) \tag{8}$$

where $s^p(m,n)$ stands for the convolutional layer's output at position (m,n) and pth channel; the uth convolutional kernel is marked as $\boldsymbol{K}^u$; and $\boldsymbol{I}^u$ denotes the image/data volume in the uth channel. A complex convolutional layer is comprised of a set of simple layers [41], as expressed in the following:

$$\boldsymbol{I}_l = \text{pool}\left(\sigma(s)\right) \tag{9}$$

where $\boldsymbol{I}_l$ stands for the lth layer's output volume; $\sigma(\cdot)$ stands for the nonlinear operation of the active function; and pool$(\cdot)$ is a pooling (down-sample) operation in the pooling layer. The abstraction ability of the network generally increases with the number of stacked convolutional layers [42]. The final representations in a CNN's last layer are usually reshaped to vectors and fed into the fully-connected (FC) layers.

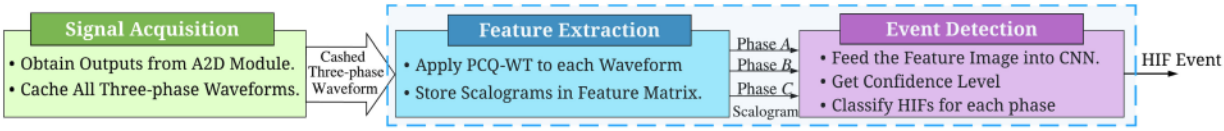


Fig. 4. Series of integrated functions in the proposed HIF detection system.

III. Proposed Feature Extraction and HIF Detection By CNNs

As each single-phase current waveform captured from the upstream terminal in a radial power distribution system carries information of the downstream terminal, we mainly focus on detecting low-intensity HIFs which is one most challenging task in electric industry. The workflow of the proposed HIF detection technology is demonstrated in Fig. 4; it functionally consists of the following four modules.

1) *Signal Acquisition:* The proposed framework shares the same input waveforms of a typical PMU (or protective relay with PMU functionality) from the Analog to Digital (A2D) processing module. It avoids any additional A2D modules in the front-end, and makes the proposed architecture an economically viable sensor solution. The waveform data is stored in the cache for the subsequent analytical processes.
2) *Feature Extraction:* This module applies pseudo-continuous quadrature wavelet transform (PCQ-WT) to the cached waveform data and generates scalograms. The scalograms are matrices that contain signal signatures corresponding to the HIF events in the power grid. The scalograms are then quantized to digital images in order to compress the data size.
3) *Event Detection and Classification:* The images obtained in the previous stage are fed into a CNN that, with a detection confidence, classifies whether there is an HIF event. Finally, the detected event will be reported to the local protection device or the control center.

As one can see, the proposed HIF detection system only requires software-level modifications to the existing sensors (e.g., PMUs, protective relays, etc.), with no additional hardware investments. Here, we utilize a physics-guided machine learning technique, as the overall detection system only requires offline training. The training process is under the guidance of the pre-recorded and/or simulated HIF waveforms as the training dataset.

A. Improved HIF Model

The waveform assessment based on the HIF model presented in Fig. 1(b) is adequate to some extent; however, to enrich the proposed solution's knowledge on a variety of HIF waveforms, we propose a comprehensive HIF model shown in Fig. 5, where the fault resistance and inductance are assigned to the positive and negative branch. For simplicity, we name Z_{FP} and Z_{FN} as positive and negative arc impedances, respectively. One advantage of this proposed model is its flexibility to approximate different HIF conditions including those studied in [1], [4], [5], [9], [11], [14].

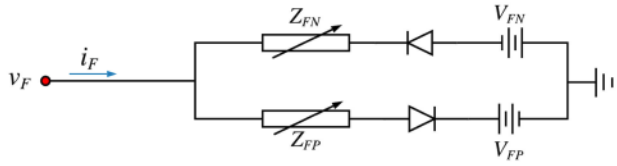


Fig. 5. Proposed improved HIF model.

B. Proposed PCQ-WT Based Feature Extraction

In order to achieve a low computing complexity solution with high-fidelity feature extraction performance, a modified complex Gabor wavelet from [43] is adopted in this article. Gabor wavelets have been widely used in two-dimension (2-D) pattern recognition processes [43]–[45]. The modified Gabor wavelet in this article is expressed as follows:

$$\Psi(t) = \underbrace{\exp\left(j\omega_c(t-b)\right)}_{\text{Periodic Component}} \cdot \underbrace{\exp\left(-\frac{(t-b)^2}{\alpha_0^2}\right)}_{\text{Gaussian Envelope}} \tag{10}$$

where w_c is the central frequency which determines the pseudo frequency of the feature extraction in the waveform. This can be proved by the Fourier transform of this Gabor wavelet as

$$\mathbf{F}_{\Psi}(\omega) = \alpha_0\sqrt{\pi} \cdot \exp(-j\omega b) \cdot \exp\left(-\frac{\alpha_0^2}{4}(\omega-\omega_c)^2\right) \tag{11}$$

when $\omega = \omega_c$, the magnitude of $\mathbf{F}_{\Psi}(\omega)$ is at its maximum; in other words, the component with frequency ω_c in the sampled waveform will have the maximum correlation with this wavelet. Thus, the feature is extracted and highlighted in the scalogram. Another advantage of this modified Gabor wavelet is that it has a predictable narrow-bandwidth (11) due to the exponential operation, which makes it easier to determine its pseudo frequency. By properly selecting α_0 in PCQ-WT, a desired bandwidth of the pseudo frequencies is covered at ω_c, while the time shift b plays no magnitude impact on (11). By substituting $t-b$ with $\frac{t-b}{a}$, the Gabor WT of each harmonic component in (5) turns into

$$X(\omega_h|a,b) = \int_{-\infty}^{\infty} \frac{A_h}{2}\left(e^{j(\omega_h t+\phi_h)} + e^{-j(\omega_h t+\phi_h)}\right) \times \Psi^*\left(\frac{t-b}{a}\right) \mathrm{d}t. \tag{12}$$

According to the Hubbard–Stratonovich transformation [46]

$$\exp\left(-\frac{\alpha}{2}x^2\right) = \sqrt{\frac{1}{2\pi\alpha}}\int_{-\infty}^{\infty} exp\left(-\frac{y^2}{2\alpha} - jxy\right)\mathrm{d}y. \quad (13)$$

The WT of Gabor wavelet in (12) becomes

$$X(\omega_h|a,b) = \frac{A_h}{2}e^{j\omega_h(\phi_h+b)}\cdot a\alpha_0\sqrt{\pi}e^{-\frac{\alpha_0^2}{4}(a\omega_h-\omega_c)^2} + \underbrace{\frac{A_h}{2}e^{-j\omega_h(\phi_h+b)}\cdot a\alpha_0\sqrt{\pi}e^{-\frac{\alpha_0^2}{4}(a\omega_h+\omega_c)^2}}_{\approx 0}. \quad (14)$$

It can be seen that the second term on the right side of (14) can be neglected as the exponent in the last exponential operation is a large negative. When $\omega_0 = \omega_c/a$, $X(\omega_h|a,b)$ reaches its maximum value, indicating that the dominant feature of the selected frequency is extracted. Here, we make

$$a\alpha_0 = \omega_c/\gamma \quad (15)$$

where γ is a constant. Accordingly, and based on (10), the length of the Gaussian window in Gabor wavelet also adapts different frequencies.

The proposed Gabor wavelet in the discrete time domain is

$$\Psi[n|a_k,b_k] = \exp\left(j\frac{\omega_c T_s(n-b_k)}{a_k}\right)\exp\left(-\frac{T_s^2(n-b_k)^2}{a_k^2\alpha_0^2}\right). \quad (16)$$

Applying different discrete scaling factors a_k and time shift b_k, we achieve the proposed PCQ-WT as follows:

$$X(\omega_k|a_k,b_k) = \sum_{n=0}^{W-1} x[n]\Psi^*\left[\frac{T_s(n-b_k)}{a_k}\right] = \sum_{n=0}^{W-1} x[n] \times \exp\left(-j\frac{\omega_c}{a_k}T_s(n-b_k) - \frac{T_s^2(n-b_k)^2}{a_k^2\alpha_0^2}\right). \quad (17)$$

If the pseudo frequencies of interest and the Gabor wavelet bank are designed properly, a set of PCQ-WTs in form of a vector $\mathbf{X}_{\underline{\omega}}$ can be generated conveying waveform features in a certain frequency range. During both transient and steady state operating modes, WT time-frequency analysis is conducted along time, and a scalogram stream can be then achieved.

C. Proposed CNN Configuration for HIF Detection

With the PCQ-WT extracted features in form of scalograms available, the HIF detection problem is converted to a supervised scalograms classification problem. However, the classification of the high-dimensional 2-D scalograms is challenging. specifically, every frame of the obtained scalogram has hundreds by hundreds ($scales \times time$) pixels; it is very challenging to process such high dimensional data through the conventional pattern classification approaches. Here, we convert the PCQ-WT scalograms into 2-D images and propose a compact CNN architecture to classify the HIFs concealed in the scalograms by the PCQ-WT. The proposed CNN has a simple architecture for HIF detection, yet achieving a very fast processing time for online applications.

D. PCWT and CNN Parameter Setting

The sampling rate Fs for the signal and the feature extraction is 7680 Hz which provides 128 samples per nominal fundamental cycle. The observation window for the PCQ-WT is set to 308 samples (40 ms). The time shift b for the modified Gabor daughter wavelets is 10 ms (77 samples) for simplicity. The scaling factor a for the proposed PCQ-WT is chosen as 2^i, where i is sampled uniformly from 0 to 8. The central frequency $\omega_c = 1152$ and the pseudo frequency will roughly reach up to the 19th order of harmonics. All the Gabor wavelets have 20 ms duration. By this design, the computational complexity in computing the scalograms is reduced, while the pseudo frequency bandwidth coverage for feature extraction will not be compromised. The scalograms fed into the CNN are cropped from 10 to 50 ms (of the total 60 ms WT output)—i.e., observation+ wavelet length—which has 40 ms (308 samples) duration.

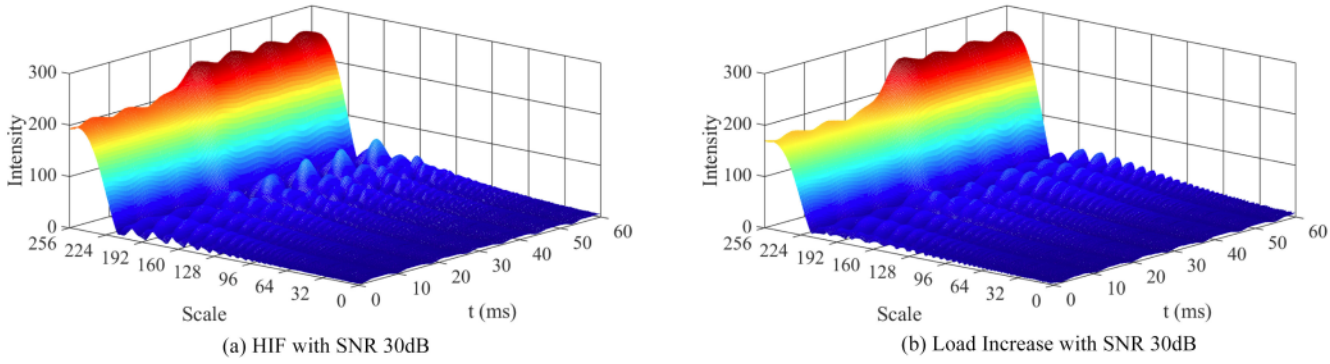
The proposed compact CNN configuration for scalogram classification is as follows: Input layer (256 × 308); Convolution (Conv.) layer (32 ×5× 11); Max-pooling layer (3 × 3); Conv. layer (32×5×5); Max-pooling layer (3 × 3); Conv. layer (32×5×5); FC layer (200 × 1); FC layer (3 × 1). Conventional images have homogeneous units on the horizontal and vertical axes, while the scalograms axes carry different information regarding the HIF events on either time or frequency. Therefore, a wide kernel in the first Conv. layer that can extract more information from the transitions along the time axis is used. The stride of the first layer is (2, 3), and the remaining Conv. layers use strides with a size of (1, 1). Besides, batch normalization [47] is used in the last FC layer. In the last Conv. layer and the first FC layer, Dropout [48] is used to prevent over-fitting. All activation functions in the CNN are rectified linear unit (ReLU). We choose cross-entropy as the loss function. One can see that our proposed CNN is not that "deep" compared to the regular image classification CNNs; the suggested compact CNN architecture further reduces the computing complexity in HIF detection.

IV. Case Study and Numerical Experiments

A. Test Scenarios Configuration

The parameter specifications used for generating the test waveforms are listed in Table I. We, in particular, focus on HIF event detection since conventional faults can typically be detected by existing protective devices. Three test scenarios (HIF event, load change event, and normal operation event) are simulated. For each HIF simulation, all parameters are *randomly* selected in the designated ranges. In each simulated waveform, only one event occurs at a random point in time. The waveform generation system is developed according to Fig. 1(a) and the improved HIF model in Fig. 5. Gaussian noises with SNR of 40 dB are added to the ac voltage source to approximate the thermal and measurement noises in different conditions.

A total of 20 000 samples from the test waveforms are simulated for each event; therefore, a total of 60 000 samples (wavelet scalogram) are simulated in the MATLAB/Simulink environment. A total of 48 000 samples are randomly selected as the training dataset, 6000 samples for validation, and 6000



(a) HIF with SNR 30dB (b) Load Increase with SNR 30dB

Fig. 6. Test waveform simulation results. (a) HIF event. (b) Load increase event; both simulated events start at $t = 20$ ms.

TABLE I
PARAMETER SPECIFICATIONS FOR GENERATING THE TEST WAVEFORMS

System Setting			
V_{base} *	13.8kV	S_{base}	0.5 MW
Fs	7.68kHz	Line Length	2km
S_{Load}	0.5 - 1.5 pu	pf_{load}	0.8 - 1
HIF Model Setting			
Parameter	Range	Parameter	Range
Z_{FP}	0.02 pu - 1.5 pu	Z_{FN}	0.02 pu - 1.5 pu
pf_{FP}	0 - 1	pf_{FN}	0 - 1
V_{FP}	0.08 pu - 0.65 pu	V_{FN}	0.08 pu - 0.65 pu
Fault Location	1% - 99%		

* :An SNR of 40 dB is added to the ac voltage source.

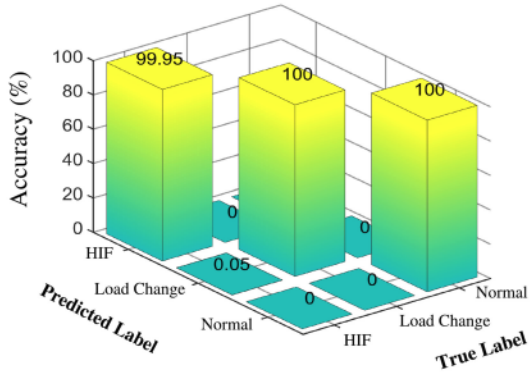


Fig. 7. Test results of the proposed CNN framework for online HIF detection and classification.

samples for testing datasets. For training the neural network and increasing the versatility of the CNN, the generated waveforms are manually imposed by white Gaussian noise with 30 dB SNR. We use Adam [49] as the optimizer, which has the initial learning rate of 1×10^{-4}, and weight decay of 1×10^{-5}. We trained the proposed CNN 120 epochs. The best validated model was recorded and tested.

B. Experimental Results and Analysis

1) Feature Extraction: The extracted features from an HIF and a load change events are demonstrated in Fig. 6, where the main energy concentration with high intensity is marked by red standing for the fundamental frequency component extracted from the waveforms. In Fig. 6(a), the HIF at $t = 20$ ms increases the main energy concentration at $t = 30$ ms, which matches the corresponding magnitude increase in Fig. 3(e). This phenomenon can also be observed in Fig. 6(b), as a load increase will indeed increase the magnitude of the current waveform. When evaluating high frequency areas in both scalograms, one can see that the higher frequency range (scale from 0 to 96) has less discontinuous patterns compared to the lower frequency range (scale from 96 to 192). However, these frequency portions (96 to 192) are discontinuous with ripple shape carrying significant differences in Fig. 6(a) and (b).

For the HIF scalogram, the ripple pattern is formed at $t = 35$ ms and becomes stable afterward; thus, one can see that it will take less than 20 ms (PCQ-WT window length) to reveal the corresponding HIF patterns—some pyramid shape ripples appear in a group. When there is a load increase event, the load change feature takes 10 ms to emerge [see the input waveform in Fig. 6(b)] and the impact remains consistently present. However, for scale from 96 to 192, there is no pyramid shape ripple found as those in Fig. 6(a). Also, only the intensity of the existing ripples slightly increases. One should notice that the patterns extracted from this simulated HIF event approximately covers the scale from 100 to 200, which roughly corresponds to the 12th to 5th order harmonics in the frequency span.

2) Event Detection: The HIF detection test results are summarized in the confusion matrix in Fig. 7, where the true label stands for the actual tested events, and the predicted label corresponds to the classification outcomes of the CNN module. The average accuracy of the proposed HIF event detection scheme is found 99.95%. To further examine the online event detection and classification performance, we used a workstation with a stock eight-core AMD Ryzen 3800X CPU as the computational platform. We transfer the PCQ-WT and CNN modules into MATLAB 2020a to record the computational time on one single core of the CPU. The overall time for processing the PCQ-WT and CNN is recorded as 6.3 ± 0.6 ms, confirming a promising solution to be used in real-time HIF detection applications.

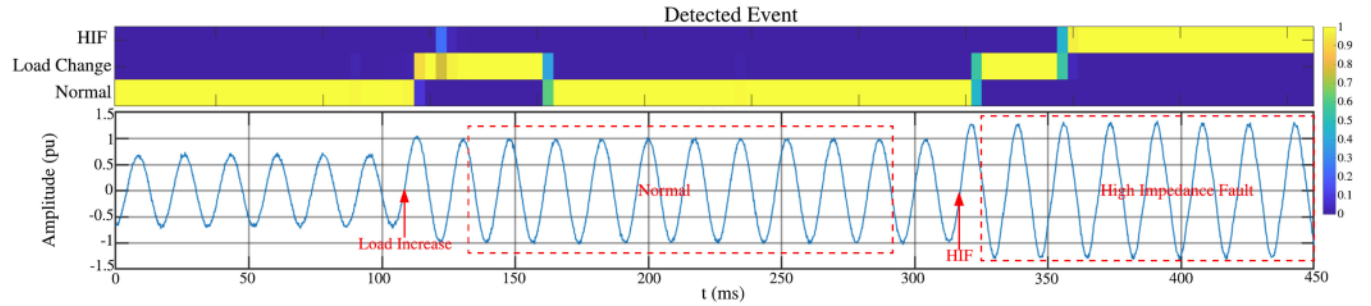


Fig. 8. Online HIF detection on a simulated single-phase current waveform: Detected result (top) and original waveform (bottom).

We also test the proposed framework on a recorded waveform of 0.45 s duration, the result of which is shown in Fig. 8. The top heat-map is the event detection results over time, where the classification confidence rate is marked with the color bar. One should note that, during the *Normal* operating event, the confidence rate is very high and the classification result is accurate even though the waveform is polluted with 30 dB Gaussian noise and the distortion is very obvious. Relatively lower confidence rates always exist during the transitions between two different events.

The *Load Change* event is detected accurately within 10 ms and it takes another 40 ms for the detection scheme to report a normal condition event. Meanwhile, at the moment of HIF occurrence, the CNN module classified the first two cycles of HIF-contained waveform (from $t = 325$ ms to $t = 355$ ms) mistakenly as the *load change* event. The reason lies in the fact that a *load change* event [see Fig. 6(b)] reveals similar patterns at the very beginning when compared to the *HIF* event [see Fig. 6(a)]; therefore, it is very hard to make a correct detection and classification decision at the exact time when the HIF event happens. However, when the waveform magnitude stabilizes, the CNN module shows a high confidence rate in classifying the HIF event correctly. Also, one can see that the maximum peak value of the current waveform affected by an HIF event is observed 1.25 p.u. which will barely trip the protective relays in the absence of the proposed HIF detection scheme.

In summary, the proposed solution provides satisfactory results in detecting low-intensity HIFs under noisy measurements and can distinguish it with load change events. Moreover, the overall detection accuracy is desirable (99.98%) at a promising speed of within 40 ms (33 ms delay plus 6.3 ± 0.6 ms processing time).

V. Conclusion

This article proposed an effective scheme that leverages AI advancements for detecting HIFs in power grids and improving electrical safety. The proposed solution functionally integrates a PCQ-WT feature extraction tool using a modified Gabor wavelet and a compact CNN-based event detection technique. Experiment results demonstrated that the proposed analytics successfully achieved an ultra-fast (within 40 ms) and accurate (99.95% accuracy) HIF detection performance even under noisy measurements.

The proposed function could be embedded within the existing PMUs and/or other IEDs that are capable of capturing and processing the power waveforms. Future work can be focused on first enriching the proposed approach with a variety of HIF models and a further strong pool of HIF waveforms, second, investigating HIF intensity measurements through CNNs or other machine learning algorithms, and third, hardware-in-the-loop performance analysis when the proposed HIF detection and localization function is embedded into the PMUs, protective relays, and/or other IEDs.

References

[1] M. Sarwar, F. Mehmood, M. Abid, A. Q. Khan, S. T. Gul, and A. S. Khan, "High impedance fault detection and isolation in power distribution networks using support vector machines," *J. King Saud Univ.-Eng. Sci.*, to be published, doi: 10.1016/j.jksues.2019.07.001.

[2] C. G. Wester, "High impedance fault detection on distribution systems," *presented at 42nd Annu. Conf. Rural Electric Power Conf.*, 1998, pp. c5–1.

[3] J. Tengdin *et al.*, "High impedance fault detection technology," Rep. of IEEE PSRC Working Group D15, Mar. 1996. [Online] Available: http://grouper.ieee.org/groups/td/dist/documents/highz.pdf

[4] Q. Cui and Y. Weng, "Enhance high impedance fault detection and location accuracy via μ-PMUs," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 797–809, Jan. 2020.

[5] L. U. Iurinic, A. R. Herrera-Orozco, R. G. Ferraz, and A. S. Bretas, "Distribution systems high-impedance fault location: A parameter estimation approach," *IEEE Trans. Power Del.*, vol. 31, no. 4, pp. 1806–1814, Aug. 2016.

[6] S. Wang, "On the use of artificial intelligence for electrical safety," presented at the Elect. Saf. Workshop, Mar., 2020, p. 1.

[7] S. Maximov, V. Torres, H. F. Ruiz, and J. L. Guardado, "Analytical model for high impedance fault analysis in transmission lines," *Math. Probl. Eng.*, vol. 2014, 2014, Art. no. 837496.

[8] D. P. S. Gomes, C. Ozansoy, and A. Ulhaq, "High-sensitivity vegetation high-impedance fault detection based on signal's high-frequency contents," *IEEE Trans. Power Del.*, vol. 33, no. 3, pp. 1398–1407, Jun. 2018.

[9] A. P. Kujur and T. Biswal, "Detection of high impedance fault in distribution system considering distributed generation," in *Proc. Int. Conf. Innov. Mechanisms Ind. Appl.*, 2017, pp. 406–410.

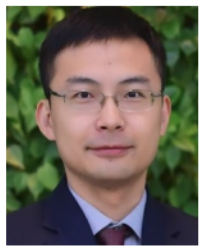
[10] O. Gashteroodkhani, M. Majidi, M. Fadali, M. Etezadi-Amoli, and E. Maali-Amiri, "A protection scheme for microgrids using time-time matrix z-score vector," *Int. J. Elect. Power Energy Syst.*, vol. 110, pp. 400–410, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0142061518322270

[11] Q. Cui, K. El-Arroudi, and Y. Weng, "A feature selection method for high impedance fault detection," *IEEE Trans. Power Del.*, vol. 34, no. 3, pp. 1203–1215, Jun. 2019.

[12] S. Bhongade and S. Golhani, "HIF detection using wavelet transform, travelling wave and support vector machine," in *Proc. Int. Conf. Elect. Power Energy Syst.*, 2016, pp. 151–156.

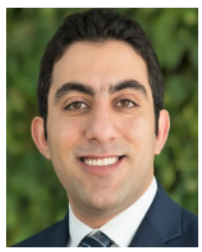
[13] K. Moloi, J. A. Jordaan, and Y. Hamam, "High impedance fault detection technique based on discrete wavelet transform and support vector machine in power distribution networks," in *Proc. IEEE AFRICON*, 2017, pp. 9–14.

[14] H. Lala and S. Karmakar, "Detection and experimental validation of high impedance arc fault in distribution system using empirical mode decomposition," *IEEE Syst. J.*, to be published, doi: 10.1109/JSYST.2020.2969966.
[15] K. V. Shihabudheen *et al.*, "Detection of high impedance fault using machine learning techniques," in *Proc. TENCON IEEE Region 10 Conf.*, 2019, pp. 2117–2122.
[16] M.-T. Yang, J.-C. Gu, J.-L. Guan, and C.-Y. Cheng, "Detection of high impedance faults in distribution system," in *Proc. IEEE/PES Transmiss. Distrib. Conf. Expo., Asia Pacific*, 2005, pp. 1–5.
[17] S. Wang, P. Dehghanian, L. Li, and B. Wang, "A machine learning approach to detection of geomagnetically induced currents in power grids," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1098–1106, Mar./Apr. 2019.
[18] S. Wang, P. Dehghanian, and L. Li, "Power grid online surveillance through PMU-embedded convolutional neural networks," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1146–1155, Mar./Apr. 2019.
[19] M. H. Rezaeian Koochi, P. Dehghanian, S. Esmaeili, P. Dehghanian, and S. Wang, "A synchrophasor-based decision tree approach for identification of most coherent generating units," in *Proc. Annu. Conf. IEEE Ind. Electron. Soc.*, Oct. 2018, pp. 71–76.
[20] S. Wang, P. Dehghanian, L. Li, and B. Wang, "A machine learning approach to detection of geomagnetically induced currents in power grids," in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting*, 2019, pp. 1–7.
[21] M. Babakmehr, F. Harirchi, M. Nazir, S. Wang, P. Dehghanian, and J. Enslin, "Sparse representation-based classification of geomagnetically induced currents," in *Proc. Clemson Univ. Power Syst. Conf.*, Mar. 2020, pp. 1–7.
[22] B. Shinde, S. Wang, P. Dehghanian, and M. Babakmehr, "Real-time detection of critical generators in power systems: A deep learning HCP approach," in *Proc. IEEE Texas Power Energy Conf.*, 2020, pp. 1–6.
[23] A. Ghaderi, H. L. Ginn III, and H. A. Mohammadpour, "High impedance fault detection: A review," *Elect. Power Syst. Res.*, vol. 143, pp. 376–388, 2017.
[24] W. C. dos Santos, B. A. de Souza, N. S. D. Brito, F. B. Costa, and M. R. C. Paes, "High impedance faults: From field tests to modeling," *J. Control, Autom. Elect. Syst.*, vol. 24, no. 6, pp. 885–896, 2013.
[25] A. von Meier, E. Stewart, A. McEachern, M. Andersen, and L. Mehrmanesh, "Precision micro-synchrophasors for distribution systems: A summary of applications," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2926–2936, Nov. 2017.
[26] R. F. Nuqui and A. G. Phadke, "Phasor measurement unit placement techniques for complete and incomplete observability," *IEEE Trans. Power Del.*, vol. 20, no. 4, pp. 2381–2388, Oct. 2005.
[27] T. Nguyen, S. Wang, M. Alhazmi, M. Nazemi, A. Estebsari, and P. Dehghanian, "Electric power grid resilience to cyber adversaries: State of the art," *IEEE Access*, vol. 8, pp. 87 592–87 608, 2020.
[28] E. M. Lima *et al.*, "High impedance fault detection method based on the short-time Fourier transform," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 11, pp. 2577–2584, 2018.
[29] Y.-C. Su, K.-L. Lian, and H.-H. Chang, "Feature selection of non-intrusive load monitoring system using STFT and wavelet transform," in *Proc. IEEE 8th Int. Conf. e-Bus. Eng.*, 2011, pp. 293–298.
[30] S.-H. Ni, K.-F. Lo, L. Lehmann, and Y.-H. Huang, "Time–frequency analyses of pile-integrity testing using wavelet transform," *Comput. Geotechnics*, vol. 35, no. 4, pp. 600–607, 2008.
[31] S. Wang, P. Dehghanian, and B. Zhang, "A data-driven algorithm for online power grid topology change identification with PMUs," in *Proc. IEEE Power Energy Soc. General Meeting*, 2019, pp. 1–5.
[32] S. Wang, P. Dehghanian, and Y. Gu, "A novel multi-resolution wavelet transform for online power grid waveform classification," in *Proc. 1st IEEE Int. Conf. Smart Grid Synchronized Meas. Analytics*, May 2019, pp. 1–8.
[33] J. C. Chen, B. T. Phung, H. W. Wu, D. M. Zhang, and T. Blackburn, "Detection of high impedance faults using wavelet transform," in *Proc. Australas. Universities Power Eng. Conf.*, 2014, pp. 1–6.
[34] S.-J. Huang and C.-T. Hsieh, "High-impedance fault detection utilizing a morlet wavelet transform approach," *IEEE Trans. Power Del.*, vol. 14, no. 4, pp. 1401–1410, Oct. 1999.
[35] D. P. Mishra, S. R. Samantaray, and G. Joos, "A combined wavelet and data-mining based intelligent protection scheme for microgrid," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2295–2304, Sep. 2016.
[36] K. Thirumala, M. S. Prasad, T. Jain, and A. C. Umarikar, "Tunable-q wavelet transform and dual multiclass SVM for online automatic detection of power quality disturbances," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3018–3028, Jul. 2018.
[37] S. Wang, L. Li, and P. Dehghanian, "Power grid online surveillance through PMU-embedded convolutional neural networks," in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting*, 2019, pp. 1–7.
[38] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
[39] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015, *arXiv:1409.1556v6*.
[40] L. Li, M. Doroslovački, and M. H. Loew, "Discriminant analysis deep neural networks," in *Proc. 53rd Annu. Conf. Inf. Sci. Syst.*, Mar. 2019, pp. 1–6.
[41] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org.
[42] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.
[43] J. V. B. Soares, J. J. G. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification," *IEEE Trans. Med. Imag.*, vol. 25, no. 9, pp. 1214–1222, Sep. 2006.
[44] Z. Zhang, M. Lyons, M. Schuster, and S. Akamatsu, "Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron," in *Proc. 3rd IEEE Int. Conf. Autom. Face Gesture Recognit.*, Apr. 1998, pp. 454–459.
[45] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 8, pp. 837–842, Aug. 1996.
[46] J. Hubbard, "Calculation of partition functions," *Phys. Rev. Lett.*, vol. 3, no. 2, 1959, Art. no. 77.
[47] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167v3*.
[48] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
[49] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017, *arXiv:1412.6980v9*.



Shiyuan Wang (Student Member, IEEE) received the B.Eng. degree in mechanical engineering from the University of Science and Technology Beijing, China, in 2012, and the M.Sc. and Ph.D. degrees in electrical engineering from The George Washington University, Washington, DC, USA, in 2014 and 2020, respectively.

His research interests include power system reliability and resiliency, smart grid and renewable energy, power grid harmonic analysis, and application of signal processing in energy analytics.



Payman Dehghanian (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 2009, the M.Sc. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2011, and the Ph.D. degree in electrical engineering from Texas A&M University, Texas, USA, in 2017.

He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, The George Washington University, Washington, DC, USA. His research interests include power system protection and control, power system reliability and resiliency, asset management, and smart electricity grid applications.

Dr. Dehghanian was a recipient of the 2013 IEEE Iran Section Best M.Sc. Thesis Award in Electrical Engineering, the 2014 and 2015 IEEE Region 5 Outstanding Professional Achievement Award, and the 2015 IEEE-HKN Outstanding Young Professional Award.