A Machine Learning Approach to Detection of Geomagnetically Induced Currents in Power Grids

Shiyuan Wang, *Student Member, IEEE*, Payman Dehghanian, *Member, IEEE*, Li Li, *Student Member, IEEE*, and Bo Wang, *Student Member, IEEE*

***Abstract*—Geomagnetically induced currents (GICs) in power grids are mainly caused by geomagnetic disturbances especially during solar storms. Such currents can potentially cause negative impacts on power grid equipment and even damage the power transformers resulting in a significant risk of blackouts. Therefore, monitoring GICs in power systems and developing solutions to mitigate their impacts before rising to a certain threatening level is urgently in need. Monitoring GICs is, however, quite a challenge and costly, as they usually appear in forms of DC components in the high voltage transmission lines, which are barely accessible through transformers. By examining the measured currents from the current transformers (CTs), this paper proposes a framework to detect GICs in power transmission systems through a hybrid time-frequency analysis combined with machining learning technology. Simulated results verify that the proposed approach can promisingly estimate GICs in power systems during a variety of grid operating conditions.**

***Index Terms*—Convolutional Neural Network (CNN); feature extraction; geomagnetic disturbance (GMD); geomagnetically induced current (GIC); harmonics; wavelet transform (WT).**

I. Introduction

Geomagnetic disturbances (GMDs) are mainly caused by solar storms, during which charged particles erupt from solar flares resulting in the associated coronal mass ejections into space during the intensity peak of the sun's cycle. Consequently, geomagnetically induced currents (GICs) will appear in the conductor surface of Earth. The flow of these currents into power transmission lines can potentially cause "half-cycle saturation" of high-voltage bulk power transformers. This phenomenon can lead to relay miss-operations, voltage dips, elevated reactive power demand, transformer overheating, disruptive harmonics, aging or malfunction of the electric power devices, and even a total collapse of the grid in the worst case scenarios [1]–[5].

Northern North America is particularly susceptible to problems resulting from GICs. On March 13, 1989, an exceptionally strong GMD caused major damages to electrical power equipment in Canada, Scandinavia, and the United States. Hydro-Quebec extra high voltage (EHV) transmission system experienced instability and tripping of lines carrying power to Montreal resulting in the total blackout of the Hydro-Quebec system [6], [7]. In the United States, a voltage fluctuation of up to 4 percent was recorded on the EHV systems in Pennsylvania, New Jersey, and Maryland. On September 19, 1989, at the Salem Unit 2 nuclear power plant, a second solar storm damaged the step-up transformers [8].

To limit the potential GICs-caused damages in power grids calls for developing advanced tools and mechanisms to monitor and detect the GICs as their consequences unfold and also solutions to mitigate the impacts before they rise to a certain threatening level. On one hand, the GMD phenomena may not always result in GICs in power systems. On the other hand, directly accessing GICs—which represents itself as a DC component in the high voltage transmission lines—is costly and a challenge. Conventional techniques to monitoring high voltage transmission lines rely solely on the AC measurements through voltage transformers (VTs) and current transformers (CTs), simply neglecting the DC components flowing on transmission lines. Additionally, there are other sources of harmonics in power grids, generated by nonlinear loads or overloading transformers, that may flow in the grid [9] and can interfere with those harmonics generated by GICs, particularly when GMD level is low. Such interference will make the GICs detection a challenge, as the measured waveforms from CTs or VTs capture all harmonic signals together with the fundamental component—50 Hz or 60 Hz.

Several studies have focused on modeling harmonic-embedded power flows and mitigation solutions when facing GICs in power systems [2], [10]–[12]. Different mitigation strategies should be taken depending on the severity of GICs impacts on the grid. However, research and development efforts on GICs impact detection mechanisms are scarce so far. The proposed techniques in [13], [14] are centered on monitoring GICs mainly based on analyzing the distorted waveforms; however, none has considered the appearance of other grid harmonics and the conflicting interactions with those of GICs. Furthermore, thermal noise would be higher in the transformer secondary winding due to the accumulated heat during transformer saturation, which adds another layer of interference, making GICs detection harder than usual. Additionally, studies in [3] demonstrated that the excited harmonic current magnitudes of different orders vary when GICs intensity increases. The harmonic components generated by GICs behave in different ways and the existing detection mechanisms do not consider such interference.

Studies in [9] show the promising performance of transformer overloading detection by applying wavelet transforms. Machine learning mechanisms have been widely utilized in solving electric power system problems [15]–[17] and are being frequently approached to revolutionize the solution techniques and emerging technologies in power grids. Inspired

S. Wang, P. Dehghanian, L. Li, and B. Wang are with the Department of Electrical and Computer Engineering, The George Washington University, Washington, DC 20052, USA (e-mails: shiyuan1225@gwu.edu; payman@gwu.edu; lili1986@gwu.edu; wangbo@gwu.edu).

978-1-5386-4539-0/19/$31.00 © 2019 IEEE

by the principle concepts of feature extraction and event detection based on waveform analysis, this paper proposes a GIC detection solution in the high voltage transmission systems. Two major time-frequency analysis techniques, namely the wavelet transform (WT) and short-time Fourier transform (STFT), are applied and their performance are evaluated. We further propose a GIC detection algorithm centered on a hybrid WT and STFT combined with a machine learning approach, Convolutional Neural Networks (CNN). We demonstrate the promising performance of the suggested analytics in detecting the GIC impacts in power grids under a variety of grid operating conditions. The proposed framework is featuring: (i) resistive to harmonic distortion and background noise, (ii) capable of detecting low-intensity GICs, and (iii) can be algorithmically embedded within Phasor Measurement Units (PMUs) and other intelligent electronic devices (IEDs) currently in place for online monitoring.

The rest of the paper is organized as follows: Section II introduces the GICs impact modeling and the motivation to feature extraction based on STFT and continuous WT. Section III describes the proposed online GICs detection framework, consisted of (i) feature extraction from the current signals via pseudo-continuous wavelet transform (PCWT) and (ii) event classification via CNN. Case studies and experimental results are analyzed in Section IV to verify the performance of the proposed GICs detection mechanism. And finally come the conclusions in Section V.

II. Background and Motivation

A. Overview of GICs Modeling

The GMD phenomenon introduces an earth magnetic field change rate of usually below 1Hz [2]. Typically, the GICs can be determined by assessing the DC network power flow as:

$$\mathbf{I} = \mathbf{G}\mathbf{V} \tag{1}$$

where $\mathbf{G}$ is the network bus admittance matrix, determined by taking into account the three-phase signals in parallel, the substation neutral buses, as well as the ground resistance. Utilizing the approach presented in [1], the GMD introduces voltages that can be approximated by

$$V = E_N L_N + E_E L_E \tag{2}$$

where E_N and E_E are the Northward and Eastward electric field (V/km), respectively, and L_N and L_E are the Northward and Eastward distance, respectively.

In transmission systems, Y-Y configuration transformers are vulnerable to half-cycle saturation the most during the GMD event. As graphically demonstrated in Fig. 1, this is because the DC flux has the lowest reluctance path in such scenarios and a semi-saturation can occur [18]. However, GICs alone may not be able to cause a transformer half-cycle saturation as (i) GICs intensity may be low and (ii) when the transformer load level is small. According to [3], when a transformer is injected with different levels of GICs, the harmonic magnitude curves (in frequency domain) will differ significantly (see Fig. 2), thus the waveforms will change dramatically too. Such behaviours in the waveforms will make the GICs detection process extremely difficult, calling for a well-thought and accurate approach.

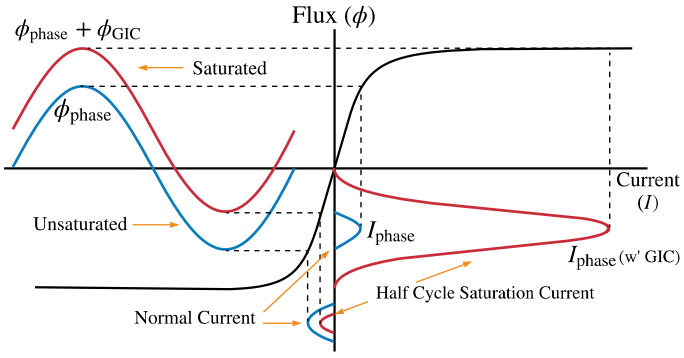


Fig. 1. Half-cycle saturation of a single-phase transformer due to GICs.

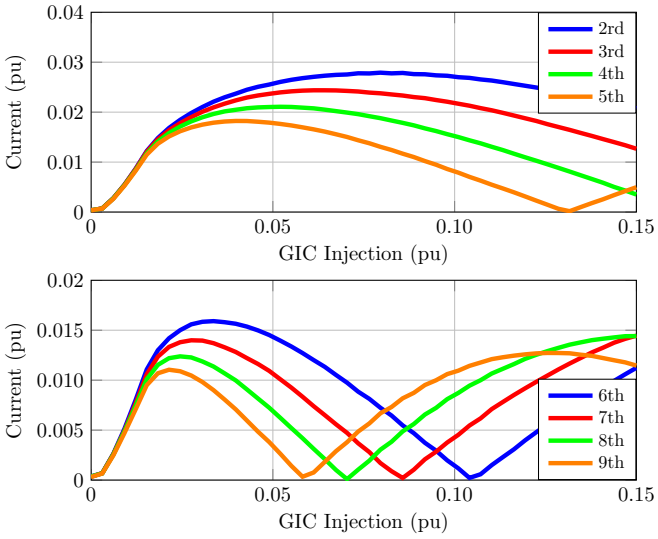


Fig. 2. Excited harmonic current components in different levels of GICs.

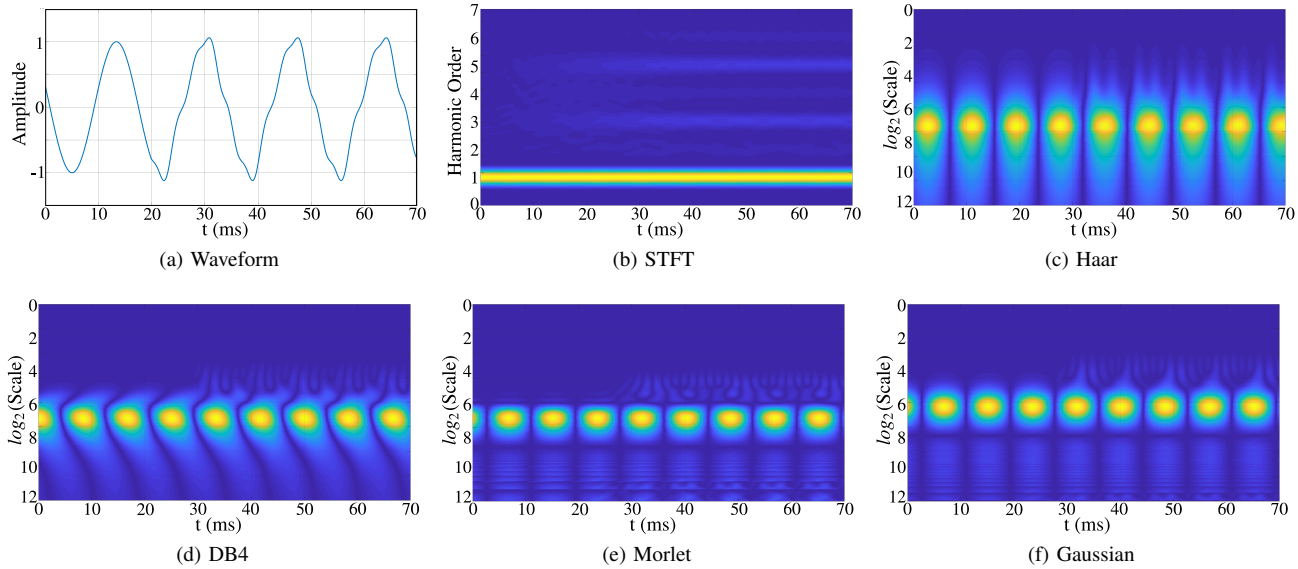
B. Power Waveform Modeling

As the current in the secondary winding of CTs can be captured and assuming a balance three-phase system, the power waveform in each phase can be represented as follows:

$$x(t) = A cos\left(\omega t + \phi\right) \tag{3}$$

where $x(t)$ is the measured one-dimension (1-D) waveform; A, ω, and ϕ are the instantaneous magnitude, fundamental frequency, and phase angle in each phase, respectively. Although the GICs cannot be measured directly, their impact can be assessed on the DC saturation level of transformers. Because GICs impacts are a set of harmonic components generated only during the transformer half-cycle saturation. Thus, the waveform can be expanded by Fourier series as

$$x(t) = A_1 cos\left(\omega_1 t + \phi_1\right) + \underbrace{\sum_{h=2}^{H} A_h cos(\omega_h t + \phi_h)}_{\text{Harmonic Components}} \tag{4}$$



(a) Waveform (b) STFT (c) Haar

(d) DB4 (e) Morlet (f) Gaussian

Fig. 3. Comparison of the STFT vs. CWT: harmonic injection starts at t=20ms, with harmonic orders h=2, 3, 4, 5, 6 and magnitude of 0.02pu, 0.08pu, 0.02pu, 0.08pu, and 0.02pu, respectively.

where h is the order of harmonics; H is the maximum order of harmonic of interest. Under a particular level of GICs, different values of A_h and f_h can be detected according to [3], resulting in a unique set of patterns. A certain combination of A_h for ω_h (for h = 2, 3, 4, ..., H) will possibly be resulted, even if some GICs-caused harmonics are influenced by those generated from other sources. The second term in (4) will still contain valuable information on the GICs impacts and, thus, could be the main target for data fusion and pattern extraction.

For time-frequency analysis of the waveforms, short time Fourier transform (STFT) [19], [20] would be one promising approach facilitating the GICs impact detection. The STFT offers a high measurement accuracy at the cost of high computational complexity [21], [22]. Meanwhile, according to [23]–[27], wavelet transform has shown speed advantages over the STFT in time-frequency analysis, especially when applied for feature extractions. The performance of different time-frequency analysis techniques are demonstrated and compared in Fig. 3. It can be seen from the spectrum in Fig. 3(b) and scalograms in Fig. 3(c)-(f), that the feature extraction outcomes through CWT are much more conspicuous than those of STFT, while compromising the accuracy in frequency measurements. Therefore, a joint hybrid application of WT and STFT is pursued in this paper for waveform feature extraction and the corresponding analysis.

C. CWT and Pseudo-CWT

The wavelet transform of a 1-D waveform is achieved by computing the cross-correlation between the signal of interest $x(t)$ and designated wavelets. This process is defined by the following equation:

$$X(\omega|a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\Psi(\frac{t-b}{a})dt \tag{5}$$

where $\Psi(t)$ is the mother wavelet when the scaling factors $a = 1$ and the time shift $b = 0$ are set. $\Psi(\frac{t-b}{a})$ is the "daughter wavelets" of $\Psi(t)$ with different selection of a and b [24], [25]. In real applications where the discrete signal processing is applied, both the integration interval and the number of scaling factors are finite; therefore, the CWT becomes *pseudo* continuous with a set of discrete scaling factors. Here, the pseudo-CWT (PCWT) is defined as follows:

$$X[\omega|a_k,b_k] = \frac{1}{\sqrt{a}} \sum_{n=0}^{W-1} x[n]\Psi[\frac{nT_s - b_k}{a_k}] \tag{6}$$

where T_s denotes the sampling interval, and W stands for the window (buffer) length. When applying PCWT, choosing a proper wavelet and tuning the parameters correctly are crucial as they will significantly affect the PCWT performance. Furthermore, a set of proper parameters would reduce the computational burden and improve the time efficiency, since the online feature extraction mandates real-time considerations when designing PCWT.

D. Convolutional Neural Networks

Once the features from waveforms are extracted, Convolutional Neural Networks (CNNs) can be used as an event detector. CNNs are artificial neural networks that are primarily used to classify images based on their contents. The process mainly focuses on representation learning, meanwhile this process adapts the feature extractor automatically and has proven successful in a wide range of image-related tasks [28]–[30]. In general, the implementation of the convolutional networks is achieved by calculating cross-correlations as defined in the following equation

$$\boldsymbol{s}^p(m,n) = \sum_u \sum_v \sum_w \boldsymbol{I}^u(m+v, n+w)\boldsymbol{K}^p(v,w), \tag{7}$$

where $\boldsymbol{s}^p(m,n)$ denotes the output of the convolutional layer at position (m,n) and p-th channel, $\boldsymbol{K}^u$ is the u-th convolutional kernel, and $\boldsymbol{I}^u$ represents the u-th channel of the

978-1-5386-4539-0/19/$31.00 © 2019 IEEE

image/data volume. A complex convolutional layer consists of small number of basic layers [31] and can be expressed by the following functions:

$$\boldsymbol{I}_l = pool\left(\sigma(\boldsymbol{s})\right), \tag{8}$$

where $\boldsymbol{I}_l$ represents the output volume of the l-th layer, $\sigma(\cdot)$ denotes a non-linearity of the neurons, and $pool(\cdot)$ is a pooling (down sampling) procedure. By stacking the convolutional layers, the abstraction level of the network generally increases [32]. The representations in the last layer in a CNN are usually expanded to vectors and fed into the general fully-connected layers. Cross-entropy is then used as the loss function.

III. Proposed Feature Extraction and GIC Detection by CNNs

As single-phase CTs can access the current waveforms from the transmission lines, the GIC impacts can be evaluated through CTs located in different locations. In this Section, the assessment focuses on one CT modeled based on Fig. 1 and Fig. 2. The approach is, however, generic enough to be applied to different models and number of CTs across the system.

A. PCWT-based Feature Extraction during Transformer Half-Cycle Saturation

Based on the half-cycle saturation waveform in Fig. 1, which is a Gaussian-like curve, the Gaussian wavelet is a natural candidate for the mother wavelet. The Gaussian wavelet can be expressed as follows

$$g_n(x) = (-1)^n \frac{d^n}{dx^n} e^{-\frac{x^2}{2}} \tag{9}$$

where n is the order of the Gaussian wavelet [33]. According to (5), the CWT with Gaussian mother wavelet can be mathematically expressed as follows

$$X_g(\omega|a,b) = \frac{1}{\sqrt{C_{g_n}}} \int_{-\infty}^{\infty} x(t) g_n(\frac{t-b}{a}) dt \tag{10}$$

$$C_{g_n} = 2\pi(n-1)! \tag{11}$$

When conducting Gaussian wavelet transform of order N with scaling and shifting factors, (10) would be modified as

$$X_g(\omega|a,b) = \underbrace{\frac{1}{\sqrt{aC_{g_N}}} \sum_{n=0}^{N-1} [\frac{d^n}{dt^n} x(t) \frac{d^{(N-n)}}{dt^{(N-n)}} g_0(\frac{t-b}{a})]\ {}^{+\infty}_{-\infty}}_{\text{zero}} + \frac{1}{\sqrt{aC_{g_N}}} \int_{-\infty}^{\infty} \frac{d^N}{dt^N} x(t) g_0(\frac{t-b}{a}) dt \tag{12}$$

As $x(t)$ only consists of sinusoidal waveforms, for a given frequency ω_h with A_h and θ_h, (12) can be rewritten as

$$X_g(\omega_h|a,b) = \frac{A_h}{\sqrt{aC_{g_N}}} \int_{-\infty}^{\infty} \frac{d^N}{dt^N} cos(\omega_h t + \theta_h) \cdot e^{\frac{-(t-b)^2}{2a^2}} dt = \frac{A_h}{2\sqrt{aC_{g_N}}} \int_{-\infty}^{\infty} \frac{d^N}{dt^N} [e^{j(\omega_h t+\theta_h)} + e^{-j(\omega_h t+\theta_h)}] e^{\frac{-(t-b)^2}{2a^2}} dt \tag{13}$$

By applying Hubbard–Stratonovich transformation [34] and using the rule of changing signs of integration limits, (13) can be simplified as

$$\begin{aligned} X_g(\omega_h|a,b) &= \\ &= \frac{A_h \omega_h^N}{\sqrt{aC_{g_N}}} cos(\omega_h b + \theta_h + \frac{N\pi}{2}) \int_{-\infty}^{\infty} e^{-\frac{y^2}{2a^2} - j\omega_h y} dy \\ &= \frac{A_h \omega_h^N}{\sqrt{(N-1)!}} cos(\omega_h b + \theta_h + \frac{N\pi}{2}) e^{-\frac{a}{2} w_h^2} \end{aligned} \tag{14}$$

ω_h, A_h, θ_h are constants for one harmonic component, and N is also a constant when the order of the Gaussian wavelet is selected. Only b controls the moment when $|X_g(\omega_h|a,b)|$ reaches maximum. The value of a controls the attenuation of the selected frequency components, i.e., a large value of a makes the amplitudes at higher frequencies attenuated rapidly and those of the low frequencies slightly, since the exponent is quadratic to ω_h. This property would help tracking a narrow bandwidth for the low-frequency components. On the other hand, when a is small, a wide bandwidth of frequency components will be extracted, and the high-frequency component will be amplified by ω_h^N. These characteristics are very suitable for feature extraction corresponding to a transformer saturation event caused by GICs. As the GICs' frequency is extremely low, daughter wavelets with large a can detect it promisingly while eliminating the influence of high frequencies. With small value of a, however, the saturation features (harmonic components) can be extracted all together and highlighted, which provide the information on the DC saturation intensity. Therefore, we chose the Gaussian wavelet to be used in PCWT for the application of interest. Finally, the PCWT for feature extraction to detect the GICs impacts can be easily obtained by applying (6) to (12). All the parameter settings will be introduced in Section III-C.

B. GICs Detection by CNNs

The overall framework for the proposed detection mechanism is demonstrated in Fig. 4. First, PCWT with K number of a and STFT in L time instants are applied sequentially. Then, the scalogram and spectrogram will both be of $K \times L$ size and carry the valuable information on the intensity of GICs. The detection process could be converted as a supervised classification problem on the scalograms. However, the classification for the 2-D scalograms is challenging due to their high dimensionality. Specifically, every frame of the obtained scalogram and spectrogram has hundreds by hundreds pixels; such high dimensional data is restrictive in most of the conventional pattern classification approaches. We cast the event detection (saturation caused by GICs) to an image classification problem based on the scalogram and spectrogram; we propose a CNN-based architecture to classify the images. The proposed CNN offers a simple architecture that ensures an accurate detection, yet fast and computationally effective. Our proposed CNN consists of five layers: three convolutional (Conv.) layers and two fully-connected (FC) layers; the specifications of the CNN will be introduced in Section III-C. This framework can work as a standalone event detector or classification tool in a PMU to detect the GICs.

978-1-5386-4539-0/19/$31.00 © 2019 IEEE

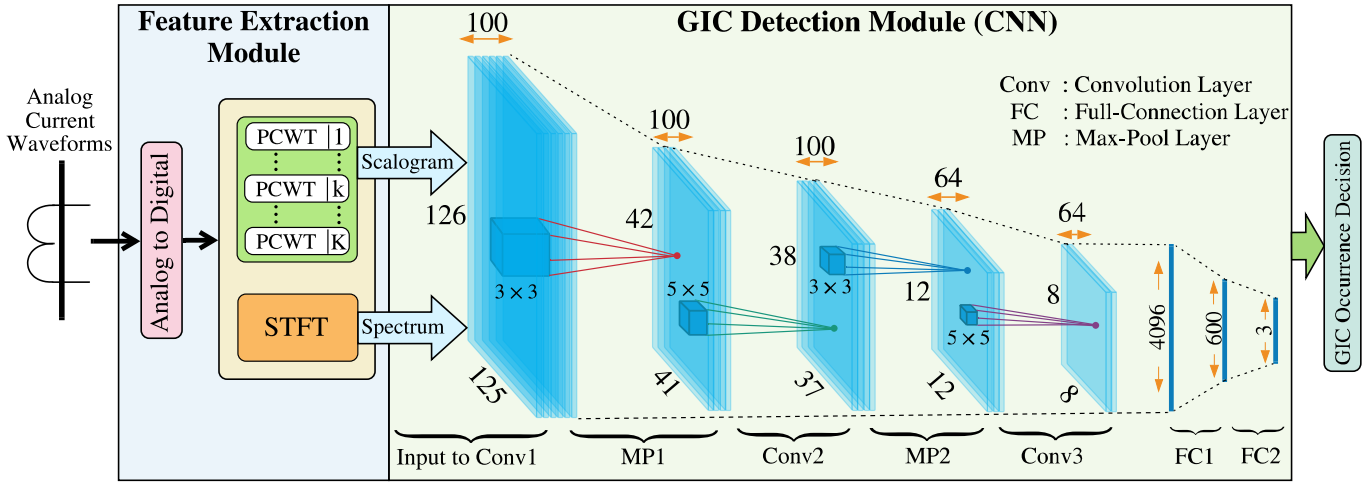


Fig. 4. The general architecture of the proposed framework for GIC detection in power grids.

C. PCWT and CNN Parameter Setting

The sampling frequency used in this paper is $Fs = 9600$Hz, which is sufficiently high to cover up to 50^{th} order of harmonics. The buffer size for the PCWT and STFT are both set to 192 samples (20ms). A Gaussian wavelet with order of 8 is employed. The time shifting b for all Gaussian daughter wavelets is set to be 10ms (96 samples) for the sake of simplicity. The proposed scaling factor a is chosen to be a set of scaling factors dyadic with 256 exponents ranging from 4 to 12. This is accomplished aiming to reduce the computational burden associate with the scalogram generating while not sacrificing the bandwidth coverage for feature extraction. The scalograms have a duration of 40ms (385 samples).

The proposed CNN for the scalogram classification is illustrated in Fig. 4 and has the following architecture: Input(256×385)–Conv(100, 5×11)–Max-pool(3×3)–Conv(100, 5×5)–Max-pool(3×3)–Conv(64, 5×5)–FC(600)–FC(3). A wide-shape kernel is chosen in the first convolutional layer aiming to extract more information of the scalogram and spectrogram along the time axis. The stride of the convolution operation in the first layer is (2, 3), while that of the other convolutional layers are (1, 1). Batch normalization [35] is used in each Conv and FC layers except the last FC layer. To prevent overfitting, dropout [36] is adopted in the third convolutional layer and the first FC layer. Rectified Linear Unit (ReLU) were chosen as nonlinearities in the neural network. Cross-entropy are used as the loss function.

IV. Case Study and Experiments

A. Test Scenarios Configuration

Three test cases are chosen from [37] to generate harmonics, aiming to simulate interferences in real-life power grid operating conditions and facilitate the performance evaluations: harmonic distortions, out-of-band interferences, and harmonics from nonlinear loads. The parameter specifications of the test power waveforms for CNN training are shown in Table. I. Each saturation scenario (AC saturation, DC saturation, and non-saturation) is associated with the three test cases plus a normal waveform. Therefore, 12 types of test waveforms are generated in total. Each type has 1000 samples. All test waveforms are polluted by Gaussian noises with a signal to noise ratio (SNR) of 30dB to approximate the thermal and measurement noises. All the parameters are uniformly located in the designated ranges. Each type of events occurs within a 20ms simulation run-time window randomly and individually. The transformer saturation model for waveform generation is obtained according to Fig. 2.

TABLE I
Simulated Test Waveform Parameter Specification

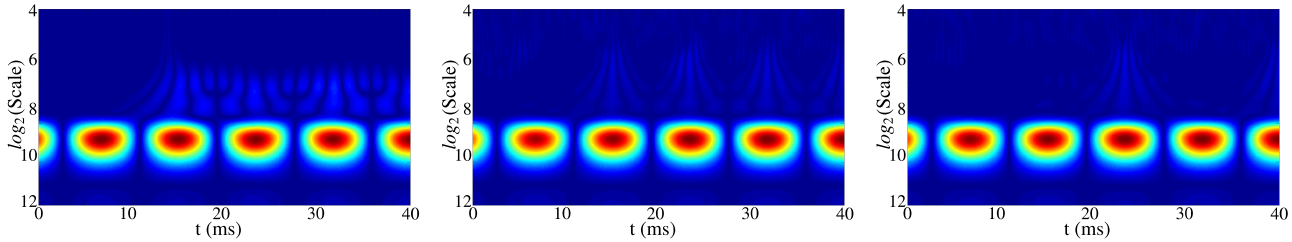
Test Case	Saturation Type		
	AC	DC	NO
Saturation level	0.001pu-0.15pu	0.001pu-0.15pu	0
Harmonic Distortion	0.5 %-10 % THD; random choose up to 50^{th} order		
Out-of-Band	10Hz to 120Hz; level 0.01pu-0.1pu		
Nonlinear Load	1% to 20% of total load		
Signal to Noise Ratio	30dB applied to all generated signal		

12,000 samples of the wavelet scalogram and spectrogram for four types of events are simulated in the MATLAB environment, wherein 10,800 samples are used as the training dataset, 1200 samples for validation, and 6,000 extra samples are generated for testing. Adam [38] was employed as the optimizer, which has initial learning rate of 1×10^{-3}, $\beta_1 = 0.9$, and $\beta_2 = 0.999$. The CNN was trained 120 epochs; in every 30 epochs, the learning rate decayed 1/10. The best validated model was recorded and tested.

B. Experimental Results and Analysis

Three patterns generated by PCWT for AC and DC saturation scenarios plus non-saturation condition are demonstrated in Fig. 5. One can see that the scalograms generated by the proposed PCWT successfully reveal unique features in such scenarios: Fig. 5(b) shows consistent spikes indicating a full-cycle saturation. Fig. 5(c) shows one spike only that stands for

978-1-5386-4539-0/19/$31.00 © 2019 IEEE



(a) Non-saturation with Harmonics (b) AC saturation (c) DC saturation

Fig. 5. Test waveform simulation results: (a) polluted with random harmonics; the AC (b) and DC (c) saturation level is 0.01pu; all events start at t=10ms.

TABLE II
ACCURACY PERFORMANCE OF THE TEST RESULTS

GIC (pu)	0-0.03	0.03-0.06	0.06-0.09	0.09-0.12	0.12-0.15	Overall
Hybrid*	73.93%	82.70%	85.15%	86.85%	87.02%	82.50%

*: PCWT + STFT

a half-cycle saturation caused by GICs. In real-world operating conditions, the patterns in Fig. 5(a) can definitely affect the classification results during AC and DC saturation. The reason lies in the fact that the patterns in Fig. 5(a) would overlap those in Fig. 5(b)(c), if harmonics and saturation occur at the same time. Therefore, the performance of CNN needs to be verified in such circumstances, in which the patterns are overlapped.

To verify the accuracy in different GIC levels, a quantization test was conducted. The GIC severity is quantized to five intervals as shown in Table II. The classification accuracy is tested for each quantization interval separately. One can see from Table II that the higher the GICs intensity, the higher the detection accuracy. In total, the proposed framework successfully achieves a desired performance even under low-intensity GICs, high harmonics, and elevated noises. Moreover, the framework can obtain the detection results within 60ms (20ms window size plus 40ms observation duration); therefore, the framework can be applied to online monitoring platforms in power transmission systems.

V. CONCLUSION

This paper aims to effectively detect GICs in power transmission systems during the GMD events. Our proposed approach consists of a hybrid feature extraction using Gaussian PCWT and STFT, and a CNN-based detection mechanism. Experiments demonstrated that the proposed analytics achieved a high accuracy for online detection of GICs under different operating conditions. This framework would be installed within PMUs and/or other IEDs that can capture the power grid waveforms. Future work will focus on (i) applying the proposed framework to a variety of CT saturation models, and (ii) investigating the potential use of CNN or other machine learning algorithms for GICs measurements.

REFERENCES

[1] D. H. Boteler and R. J. Pirjola, "Modelling geomagnetically induced currents produced by realistic and uniform electric fields," *IEEE Transactions on Power Delivery*, vol. 13, no. 4, pp. 1303–1308, Oct 1998.

[2] T. J. Overbye, T. R. Hutchins, K. Shetye, J. Weber, and S. Dahman, "Integration of geomagnetic disturbance modeling into the power flow: A methodology for large-scale system studies," in *2012 North American Power Symposium (NAPS)*, Sep. 2012, pp. 1–7.

[3] R. A. Walling, "Potential impacts of harmonics on bulk system integrity during geomagnetic disturbances," in *2013 IEEE Power Energy Society General Meeting*, July 2013, pp. 1–5.

[4] V. D. Albertson, J. M. Thorson, R. E. Clayton, and S. C. Tripathy, "Solar-induced-currents in power systems: Cause and effects," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-92, no. 2, pp. 471–477, March 1973.

[5] and and J. G. Kappenman, "Comparative analysis of exciting current harmonics and reactive power consumption from gic saturated transformers," in *2001 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No.01CH37194)*, vol. 1, Jan 2001, pp. 318–322 vol.1.

[6] A. Pulkkinen, S. Lindahl, A. Viljanen, and R. Pirjola, "Geomagnetic storm of 29–31 october 2003: Geomagnetically induced currents and their relation to problems in the swedish high-voltage power transmission system," *Space Weather*, vol. 3, no. 8, pp. 1–19, Aug 2005.

[7] S. Guillon, P. Toner, L. Gibson, and D. Boteler, "A colorful blackout: The havoc caused by auroral electrojet generated magnetic field variations in 1989," *IEEE Power and Energy Magazine*, vol. 14, no. 6, pp. 59–71, Nov 2016.

[8] J. Kappenman, *Geomagnetic storms and their impacts on the US power grid*. Citeseer, 2010.

[9] R. P. Medeiros and F. B. Costa, "A wavelet-based transformer differential protection with differential current transformer saturation and cross-country fault detection," *IEEE Transactions on Power Delivery*, vol. 33, no. 2, pp. 789–799, April 2018.

[10] A. H. Etemadi and A. Rezaei-Zare, "Optimal placement of gic blocking devices for geomagnetic disturbance mitigation," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2753–2762, Nov 2014.

[11] H. Zhu and T. J. Overbye, "Blocking device placement for mitigating the effects of geomagnetically induced currents," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 2081–2089, July 2015.

[12] M. Kazerooni, H. Zhu, and T. J. Overbye, "Mitigation of geomagnetically induced currents using corrective line switching," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2563–2571, May 2018.

[13] P. Ripka, K. Draxler, and R. Styblikova, "Measurement of dc currents in the power grid by current transformer," *IEEE Transactions on Magnetics*, vol. 49, no. 1, pp. 73–76, Jan 2013.

[14] L. Lian-guang, Z. Hao, L. Chun-ming, G. Jian-hui, and G. Qing-xiong, "Technology of detecting gic in power grids & its monitoring device," in *2005 IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific*. IEEE, 2005, pp. 1–5.

[15] C. Rudin, D. Waltz, R. N. Anderson, A. Boulanger, A. Salleb-Aouissi, M. Chow, H. Dutta, P. N. Gross, B. Huang, S. Ierome *et al.*, "Machine learning for the new york city power grid," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 2, pp. 328–345, 2012.

[16] S. Fan, L. Chen, and W.-J. Lee, "Machine learning based switching model for electricity load forecasting," *Energy Conversion and Management*, vol. 49, no. 6, pp. 1331–1344, 2008.

[17] 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 *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, Oct 2018, pp. 71–76.

[18] N. A. E. R. Corporation, "Effects of geomagnetic disturbances on the bulk power system," 2012.

[19] J. A. de la O Serna and J. Rodríguez-Maldonado, "Taylor–kalman–fourier filters for instantaneous oscillating phasor and harmonic estimates," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 4, pp. 941–951, April 2012.

[20] M. Bertocco, G. Frigo, C. Narduzzi, C. Muscas, and P. A. Pegoraro, "Compressive sensing of a taylor-fourier multifrequency model for synchrophasor estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 12, pp. 3274–3283, Dec 2015.

[21] Y.-C. Su, K.-L. Lian, and H.-H. Chang, "Feature selection of non-intrusive load monitoring system using stft and wavelet transform," in *e-Business Engineering (ICEBE), 2011 IEEE 8th International Conference on*. IEEE, 2011, pp. 293–298.

[22] S.-H. Ni, K.-F. Lo, L. Lehmann, and Y.-H. Huang, "Time–frequency analyses of pile-integrity testing using wavelet transform," *Computers and Geotechnics*, vol. 35, no. 4, pp. 600–607, 2008.

[23] S. Wang, P. Dehghanian, and B. Zhang, "A data-driven algorithm for online power grid topology change identification with PMUs," in *2019 IEEE Power Energy Society General Meeting (PESGM)*, Aug 2019, pp. 1–5.

[24] D. P. Mishra, S. R. Samantaray, and G. Joos, "A combined wavelet and data-mining based intelligent protection scheme for microgrid," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2295–2304, Sept 2016.

[25] 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 Transactions on Smart Grid*, vol. 9, no. 4, pp. 3018–3028, July 2018.

[26] P. Rajaraman, N. A. Sundaravaradan, R. Meyur, M. J. B. Reddy, and D. K. Mohanta, "Fault classification in transmission lines using wavelet multiresolution analysis," *IEEE Potentials*, vol. 35, no. 1, pp. 38–44, Jan 2016.

[27] S. Wang, P. Dehghanian, and Y. Gu, "A novel multi-resolution wavelet transform for online power grid waveform classification," in *The 1st IEEE International Conference on Smart Grid Synchronized Measurements and Analytics (SGSMA)*, May 2019, pp. 1–8.

[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[30] L. Li, M. Doroslovački, and M. H. Loew, "Discriminant analysis deep neural networks," in *53rd Annual Conference on Information Sciences and Systems*, March 2019, pp. 1–6.

[31] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[32] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[33] G. Ososkov and A. Shitov, "Gaussian wavelet features and their applications for analysis of discretized signals," *Computer physics communications*, vol. 126, no. 1-2, pp. 149–157, 2000.

[34] J. Hubbard, "Calculation of partition functions," *Physical Review Letters*, vol. 3, no. 2, p. 77, 1959.

[35] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[36] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[37] T. Becejac, P. Dehghanian, and M. Kezunovic, "Probabilistic assessment of pmu integrity for planning of periodic maintenance and testing," in *IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 2016, pp. 1–6.

[38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.