Deep complex neural network learning for high-voltage insulation fault classification from complex bispectrum representation

Mitiche, Imene; Jenkins, Mark David; Boreham, Philip ; Nesbitt, Alan; Morison, Gordon

Deep Complex Neural Network Learning for High-Voltage Insulation Fault Classification from Complex Bispectrum Representation

Imene Mitiche*, Mark David Jenkins†, Philip Boreham‡
Alan Nesbitt* and Gordon Morison*
* School of Engineering and Built Environment
Glasgow Caledonian University,
Glasgow, United Kingdom
† Geckotech Solutions Ltd.
Edinburgh, United Kingdom
‡ Innovation Centre for Online Systems
Bere Regis,
United Kingdom

Abstract—**Bispectrum representations previously achieved a successful classification of insulation fault signals in High-Voltage (HV) power plant. The magnitude information of the Bispectrum was implemented as a feature for a Deep Neural Network. This preliminary research brought interest in evaluating the performance of Bispectrum as complex input features that are implemented into a Deep Complex Valued Convolutional Neural Network (CV-CNN). This paper presents the application of this novel method to condition monitoring of High Voltage (HV) power plant equipment. Discharge signals related to HV insulation faults are measured in a real-world power plant using the Electromagnetic Interference (EMI) method and processed using third order Higher-Order Statistics (HOS) to obtain a Bispectrum representation. By mapping the time-domain signal to Bispectrum representations the problem can be approached as a complex-valued classification task. This allows for the novel combination of complex Bispectrum and CV-CNN applied to the classification of HV discharge signals. The network is trained on signals from 9 classes and achieves high classification accuracy in each category, improving upon the performance of a Real Valued CNN (RV-CNN).**

I. INTRODUCTION

Power generating plants exploit High-Voltage (HV) electrical and mechanical assets including motors, transformers, generators, cables and auxiliary equipment. The health and proper operation of these assets is crucial as any failure will primarily lead to system downtime which affects electricity production. The consequences include high replacement or maintenance costs. Furthermore, employees safety within the site could be put at risk if the failure engenders significant damage, where for instance an asset is on fire. Early detection and identification of the asset failure would avoid such consequences through condition monitoring. The latter avoids high costs, losses, fines and civil complaints [1]. Condition monitoring not only avoids failure occurrence, but it also prevents over maintenance and waste of resources when maintenance is not required. This maximises business profit and return of investment for the owning companies.

Faults within HV systems are commonly mechanical, which generally occur in gearboxes and bearings, or they are related to electrical insulation such as Arcing and Partial Discharges (PD). Identification of such faults is usually performed through an investigation of vibration [2], acoustic [3], audio or video recordings during the asset's operation. Another traditional method of fault diagnosis is Electromagnetic Interference (EMI) analysis approach [4], [5]. The exertion of EMI method in condition monitoring of HV equipment within the power production industry has been useful in the past decades. This approach is able to detect insulation degradation and conduction faults as well as mechanical defects [6]. In the current industry, EMI method is exploited by field experts who collect and analyse EMI data in operating HV sites. This approach has been successful throughout the years based on previous forensic confirmation on assets shut-down or maintenance decisions that were made due to faults occurrence. However, utility companies are suffering from a lack in expert knowledge and staff for condition monitoring, due to retirement of an ageing workforce, as well as sickness and periods of annual leave. This can have an impact on tasks that depend on expert knowledge, specifically the analysis of data measured on site in order to identify potential faults in a particular asset. Companies can be left with no other option but to hire or train more staff, however this requires a considerable amount of investment. Furthermore, their process is time consuming and not practical for continuous monitoring as the move to online condition assessment requires automatic fault classification and alerting. The implementation of continuous monitoring within the power industry reduces both cost and production losses by early detection of the potential risks.

The precise and early faults detection stage is critical in HV power systems in order to avoid catastrophic failures. Thus, a continuous and highly accurate condition monitoring system is desired. The aim here is to approach this problem by utilising Deep Learning technique. This implies

a supervised learning for fault classification which reduces the dependence on staff while possibly helping experts and non-experts in gaining more confidence and information on data interpretations. Furthermore, this brings the benefit of minimising the consequences associated with asset failure, such as maintenance costs, due to poor servicing.

The approach involves two main stages: data pre-processing and Deep Learning classification. The first stage is achieved using Bispectrum calculation on the captured EMI time series signals. This provides both magnitude and phase information contained within the time signals. The complex Bispectrum is implemented as a two channels input to a Deep Complex Valued Convolutional Neural Network (CV-CNN) inspired by [7].

Bispectrum, has recently received the attention of biomedical [8]–[10] and audio [11], [12] applications as a feature extraction technique. The authors in [8] utilised Bispectrum entropies to retrieve information from lung sound signals with the aim of identifying lung health related condition (normal, asthma or pneumonia). This feature extraction method was introduced in [9] and [10] to Machine Learning and neural network classifiers, on electrocardiogram signals for classification of five different types of cardiac arrhythmia and emotion recognition purposes, respectively. Mel frequency Cepstral coefficients and Higher-Order Statistics (HOS) features were extracted in [11] and [12] from non-linear and non-stationary music audio signals, with the aim of musical instruments and instrument family classification, using different neural networks (e.g., counter propagation, feed forward neural network). Bispectrum analysis was also exploited on vibration signals, with non-stationarity characteristics, to identify gearbox fault types [13]. More importantly, the magnitude of Bispectrum was employed as image classification task for EMI-based faults using Deep Neural Network [14]. The previously cited papers demonstrated the ability of Bispectrum to retrieve robust features for the analysis of complex, non-linear and non-stationary signals. This is the main motivation to employ Bispectrum, as EMI signals are believed to be of similar nature. However, the phase of the Bispectrum remained unexploited. It is interesting to investigate the performance of employing the complex Bispectrum, which represents both magnitude and phase information, instead of using the magnitude only.

The idea to adopt a CV-CNN was practised by few researchers. Authors in [15] and [16] exploited CV-CNN for image classification tasks and obtained the state-of-the-art results. Authors in [17] applied CV-CNN in the biomedical field for classification of sleep stage, and it was found that CV-CNN achieve an improved classification performance and considerably reduce the convergence speed when compared to CNN performance. The motif behind using a CV-CNN in this paper is due to the following advantages. The model is computationally stable and efficient as it employs complex numbers. Furthermore, the imaginary part of the complex value may offer additional information regarding the data that aids in classification task, since the imaginary and real

parts are treated independently as two channels in this work. To the authors knowledge, the combination of the Complex Bispectrum image and CV-CNN is novel to the literature. In this paper, the newly developed model is tested on real-world HV data for insulation fault classification.

The remaining of this paper is structured as follows. Section 2 introduces the Bispectrum and CV-CNN algorithms. Section 3 provides a description on the experimental set-up and data, followed by results. Finally, Section 4 draws conclusions from this research with future research and recommendations.

II. FEATURE EXTRACTION AND CLASSIFICATION ALGORITHMS DESCRIPTION

This section describes the mathematical framework of the proposed analysis. First, let's denote scalars by lower case, vectors by bold lower case and matrices by bold upper case.

A. Bispectrum Analysis

In time-domain HOS analysis, the first-order moment of a discrete signal $\mathbf{x}[n]; n = 1, ..., N$ is the mean μ and its second-order moment is the variance σ^2, both are defined in Equation 1 and 2 respectively.

$$\mu = E(\mathbf{x}[n]) \tag{1}$$

$$\sigma^2 = E(\mathbf{x}[n]\mathbf{x}[n + \tau_1]) \tag{2}$$

The third-order moment of the signal is defined as:

$$\gamma = E(\mathbf{x}[n]\mathbf{x}[n + \tau_1]\mathbf{x}[n + \tau_2]) \tag{3}$$

where τ_1 and τ_2 are time lags. Assuming that the signal $\mathbf{x}[n]$ is zero mean, then the second- and third- order cumulants are equal to the respective moments [18]. In frequency-domain definition, the third-order cumulant spectrum, known as the Bispectrum, is the Fourier Transform (FT) of the third-order moment which can be defined as:

$$\mathbf{B}(f_1, f_2) = \sum_{\tau_1=1}^{N} \sum_{\tau_2=1}^{N} \sum_{n=1}^{N} \mathbf{x}[n]\mathbf{x}[n + \tau_1]\mathbf{x}[n + \tau_2]$$
$$\cdot e^{-j2\pi \cdot (f_1\tau_1 + f_2\tau_2)} \tag{4}$$

Let $n + \tau_1 = m$ and $n + \tau_2 = k$, by rewriting Equation 4 and splitting the exponent, Bispectrum can be reformulated as:

$$\mathbf{B}(f_1, f_2) = \{\sum_{m=1}^{N} \mathbf{x}[m]e^{-j2\pi f_1 m}\} \cdot \{\sum_{k=1}^{N} \mathbf{x}[k]e^{-j2\pi f_2 k}\}$$
$$\cdot \{\sum_{n=1}^{N} \mathbf{x}[n]e^{+j2\pi(f_1+f_2)n}\} = X(f_1)X(f_2)X^*(f_1 + f_2) \tag{5}$$

where $X(f)$ is the FT and $X^*(f)$ is the complex conjugate of $X(f)$. This equation allows a faster computation of the Bispectrum, which is a complex measure that contains both magnitude and phase information, unlike the power spectrum. The Bispectrum reveals its content in two independent frequencies f_1 and f_2. The outcome of Bispectrum analysis is implemented as a complex input to the CV-CNN

B. Deep Complex Neural Network

In this section, the mathematical definition of the CV-CNN architecture and the main complex blocks are described.

1) Complex Convolution: The 2D complex convolution is performed by convolving a complex filter matrix $\mathbf{W} = \mathbf{A} + j\mathbf{B}$ by a complex vector $\mathbf{z} = \mathbf{x} + j\mathbf{y}$, where the matrices $\mathbf{A}, \mathbf{B}$, and vectors $\mathbf{x}, \mathbf{y}$ consist of real-valued entities. The convolution in frequency domain can be formulated as follows:

$$\mathbf{W} * \mathbf{z} = (\mathbf{A} * \mathbf{x} - \mathbf{B} * \mathbf{y}) + (j(\mathbf{B} * \mathbf{x} + \mathbf{A} * \mathbf{y})) \quad (6)$$

The resulting convolution has also real and imaginary parts which can be written as:

$$\begin{bmatrix} \mathscr{R}(\mathbf{W} * \mathbf{z}) \\ \mathscr{I}(\mathbf{W} * \mathbf{z}) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{-B} \\ \mathbf{B} & \mathbf{A} \end{bmatrix} * \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \quad (7)$$

2) Complex Batch Normalization: Batch Normalization (BN) is an important step in Deep Networks for various advantages such as accelerating the training and providing a higher learning rate to avoid vanishing gradient. A BN method for CV-CNN is utilised to fit within the complex model, which is calculated for a data input $\mathbf{x}$ in the following steps. First, the centered data is multiplied by the inverse square root of the 2×2 covariance matrix defined in Equation 9 as:

$$\hat{\mathbf{x}} = (\mathbf{V}^{-1/2}(\mathbf{x} - E[\mathbf{x}])) \quad (8)$$

$$\mathbf{V} = \begin{bmatrix} Cov(\mathscr{R}(\mathbf{x}), \mathscr{R}(\mathbf{x})) & Cov(\mathscr{R}(\mathbf{x}), \mathscr{I}(\mathbf{x})) \\ Cov(\mathscr{I}(\mathbf{x}), \mathscr{R}(\mathbf{x})) & Cov(\mathscr{I}(\mathbf{x}), \mathscr{I}(\mathbf{x})) \end{bmatrix} \quad (9)$$

The BN is then calculated for $\hat{\mathbf{x}}$ as:

$$BN(\hat{\mathbf{x}}) = \mathbf{\Gamma}\hat{\mathbf{x}} + \beta \quad (10)$$

where β is a complex parameter with real and imaginary learnable components, $\mathbf{\Gamma}$ is a 2×2 scaling parameter defined in Equation 11.

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_{rr} & \gamma_{ri} \\ \gamma_{ri} & \gamma_{ii} \end{bmatrix} \quad (11)$$

γ_{rr}, γ_{ii} and γ_{ri} are set to $1/\sqrt{2}$, $1/\sqrt{2}$ and 0 respectively at the start of the operation. Moving averages are implemented to estimate the complex BN parameters, where they are initialised to zero for β, $Cov(\mathscr{R}(\mathbf{x}), \mathscr{I}(\mathbf{x}))$, and to $1/\sqrt{2}$ for $Cov(\mathscr{R}(\mathbf{x}), \mathscr{R}(\mathbf{x}))$ and $Cov(\mathscr{I}(\mathbf{x}), \mathscr{I}(\mathbf{x}))$.

3) Complex ReLU: In this work the activation is applied to the real and imaginary parts of the complex value independently. Here, a complex ReLU (cReLU) is utilised as follows:

$$cReLU(\mathbf{z}) = ReLU(\mathscr{R}(\mathbf{z})) + jReLU(\mathscr{I}(\mathbf{z})) \quad (12)$$

This results in preserved phase and magnitude information when the angle $\theta \in]0, \pi/2[$. The real part is discarded when $\theta \in [\pi/2, \pi]$, the imaginary part when $\theta \in [3\pi/2, 2\pi]$ and both parts are discarded when $\theta \in]\pi, 3\pi/2[$.

TABLE I
STRUCTURE OF THE CV-CNN ARCHITECTURE. THE COMPLEX CONVOLUTION BLOCK CONSISTS OF 2D COMPLEX CONV., COMPLEX BN AND cReLU.

Layer	Kernel	No. filters	Stride	Operation
1	6x6	16	2	Complex Conv. block
2	2x2	-	2	Average 2D Pool.
3	3x3	32	2	Complex Conv. Block
4	2x2	-	2	Average 2D Pool.
5	3x3	64	1	Complex Conv. Block
6	2x2	-	2	Average 2D Pool.
7	3x3	64	1	Complex Conv. Block
8	2x2	-	2	Average 2D Pool.
9	3x3	128	1	Complex 2D Conv.+ReLU
10	3x3	128	1	Complex Conv. Block
11	2x2	-	2	Average 2D Pool.
12	Flatten, Dense(cReLU), SoftMax			

TABLE II
DATA SET INFORMATION.

	No. Training	No. Validation	No. Testing
Per class	95	12	12
Total	855	108	108

4) CV-CNN architecture: The CV-CNN model was constructed using the complex 2D convolution, cReLU and complex BN described above, in addition to the standard Keras average pooling. The overall architecture is summarised in Table I and it is implemented in this work with a learning rate starting at 0.001 then it is set to 0.0001 after 10 epochs. The model is trained over 100 epochs with a batch size of 32.

III. EXPERIMENTAL SET-UP

The data utilised for classification in this paper was measured by means of EMI method from real-world operating assets within HV power plant. It consists of time signals captured at a sampling rate of 24kHz, where each signal was examined by an expert and labelled according to the event. The interested reader in further details on EMI measurement can be referred to [4], [5], [19]. A total of 9 different events were identified and labelled as Arcing, Corona, Data Modulation, PD, Process Noise, Random Noise, Exciter pulses, minor PD and micro Sparking. Some of these events represent an insulation fault (e.g. Arcing, micro Sparking, Corona etc.) and some other are common interference (e.g. Data Modulation, Process and Random Noise). Figure 2 illustrates time signals example from each class label. A total of 1071 signals are utilised, where 80% of the data is used for training the model with 10% for validation, and the remaining 10% were used for testing. This is performed over 10 folds to achieve a ten-fold cross validation. Information on the data splitting for each fold is summarised in Table II.

The overall signal analysis and classification process is summarised in Figure 1. First, Bispectrum of each signal was calculated to obtain complex valued information as a 64x64 complex matrix. The latter was reshaped to a 64x64x2 tensor, where the 2 channels represent the real and imaginary
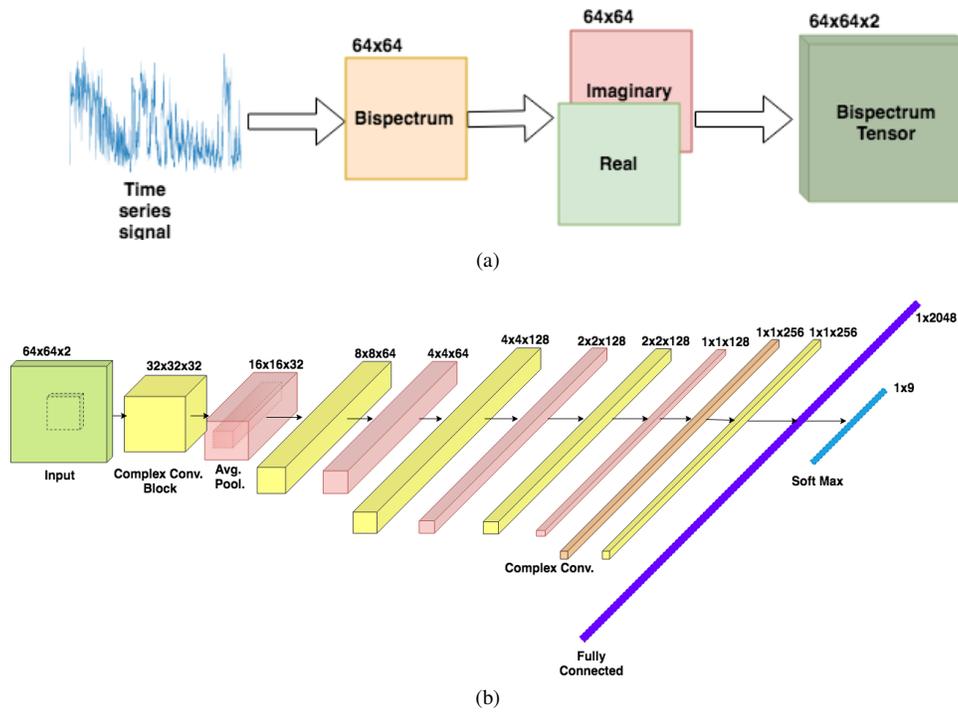
Fig. 1. (a) Signal pre-processing (b) CV-CNN model architecture, where complex conv. block includes complex convolution, complex BN and cReLU.
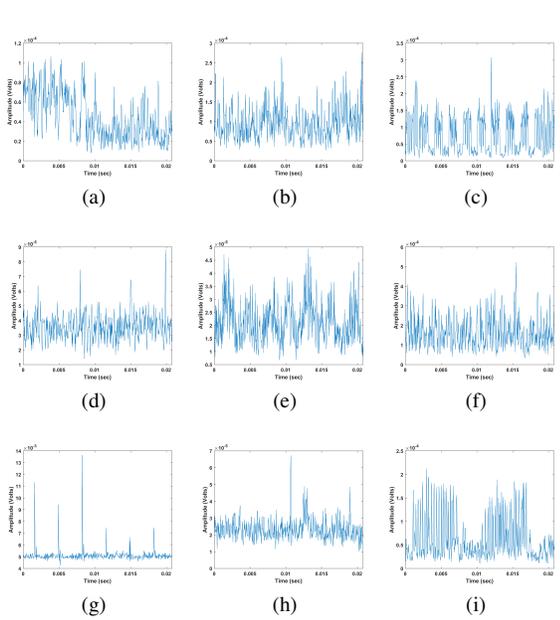

Fig. 2. Time signal examples (a) Arcing (b) Corona (c) Data Modulation (d) Partial Discharge (PD) (e) Process Noise (f) Random Noise (g) Exciter (h) minor PD (i) micro Sparking.

Fig. 3. Average confusion matrix calculated over the 10 validation folds for RV-CNN. Labels denoted as Arcing (A), Corona (C), Data Modulation (DM),Partial Discharge (PD), Process Noise (PN), Random Noise (RN), Exciter (E), minor PD (mPD), micro Sparking(mS). Classification accuracy is shown in the bottom right corner.

parts. The resulting tensor is then implemented into the CV-CNN. For comparison purposes, the same experiment was performed on a Real Valued CNN (RV-CNN) which follows the same architecture as CV-CNN's with the following exceptions. Normal convolution and BN are employed instead of the complex ones, max pooling function is applied instead of

average pooling. Here, the input to the model is the magnitude of the Bispectrum. The average results over the ten folds are presented in this section in terms of confusion matrices for RV-CNN in Figure 3 and for CV-CNN in Figure 4. The precision, recall and average accuracy are also presented in the confusion matrices. It is observed that the RV-CNN is unable to classify Corona and has lower overall accuracy. HV condition monitoring results are critical, therefore it is important for the model to have a minimum level of confidentiality and accuracy in classifying each class. Employing the CV-CNN improves
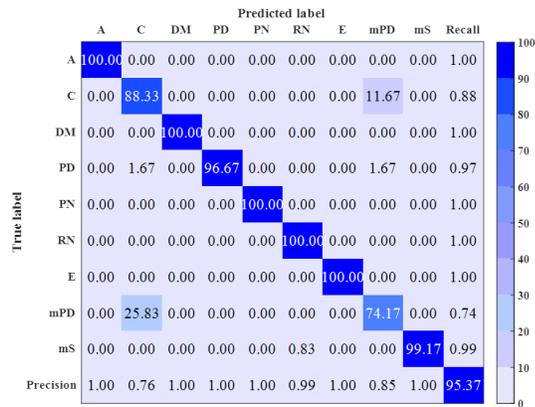
True label \ Predicted label	A	C	DM	PD	PN	RN	E	mPD	mS	Recall
A	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
C	0.00	88.33	0.00	0.00	0.00	0.00	0.00	11.67	0.00	0.88
DM	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
PD	0.00	1.67	0.00	96.67	0.00	0.00	0.00	1.67	0.00	0.97
PN	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00	1.00
RN	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	1.00
E	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	1.00
mPD	0.00	25.83	0.00	0.00	0.00	0.00	0.00	74.17	0.00	0.74
mS	0.00	0.00	0.00	0.00	0.83	0.00	0.00	0.00	99.17	0.99
Precision	1.00	0.76	1.00	1.00	1.00	0.99	1.00	0.85	1.00	95.37

Fig. 4. Average confusion matrix calculated over the 10 validation folds for CV-CNN. Labels denoted as Arcing (A), Corona (C), Data Modulation (DM),Partial Discharge (PD), Process Noise (PN), Random Noise (RN), Exciter (E), minor PD (mPD), micro Sparking(mS). Classification accuracy is shown in the bottom right corner.

the classification of this particular class, as well as PD class. In addition, the overall accuracy is improved. These results could justify the advantage of utilising the imaginary part of the Bispectrum and hence the phase information to distinguish between the classes more efficiently. By referring to Figure 2, the different signal classes can be easily confused to the naked eye, for instance, Corona, Data Modulation, Process Noise, Random Noise and minor PD. This hinders the classification task. Furthermore, the presence of hidden features may not be retrieved in the real space which explains the obtained results.

IV. Conclusion

A CV-CNN combined with complex Bispectrum was proposed in this paper in order to exploit both phase and magnitude information contained within EMI signals to effectively classify between faults within HV power plants. The preliminary experiments on real-world data demonstrate that the CV-CNN performs better than RV-CNN, which employs the magnitude information of the Bispectrum only. Therefore it can be concluded that the phase information is potentially a useful feature and hence the proposed combination was successful. The promising findings open various opportunities to develop different complex Deep Networks for an enhanced performance as there is still room for improvement. Future studies could also aim to replicate results in a larger data set.

References

[1] G. Robles, E. Parrado-Hernández, J. Ardila-Rey, and J. M. Martínez-Tarifa, "Multiple Partial Discharge Source Discrimination with Multiclass Support Vector Machines," *Expert Systems with Applications*, vol. 55, pp. 417 – 428, 2016.

[2] J. Zarei, M. A. Tajeddini, and H. R. Karimi, "Vibration analysis for bearing fault detection and classification using an intelligent filter," *Mechatronics*, vol. 24, no. 2, pp. 151 – 157, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S095741581400004X

[3] P. Casals-Torrens, A. González-Parada, and R. Bosch-Tous, "Online PD Detection on High Voltage Underground Power Cables by Acoustic Emission," *Procedia Engineering*, vol. 35, pp. 22 – 30, 2012.

[4] J. E. Timperley and J. M. Vallejo, "Condition Assessment of Electrical Apparatus With EMI Diagnostics," *IEEE Transactions on Industry Applications*, vol. 53, pp. 693–699, 2017.

[5] J. E. Timperley, J. M. Vallejo, and A. Nesbitt, "Trending of EMI Data Over Years and Overnight," in *2014 IEEE Electrical Insulation Conference (EIC)*, Philadelphia, USA, 2014, pp. 176–179.

[6] J. Timperley, "Application of EMI Diagnostics to Hydro Generators," Technical Workshop, 2012, the Northwest Hydroelectric Association Annual Conference.

[7] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal, "Deep Complex Networks," in *International Conference on Machine Learning*, Vancouver, Canada, 2018, pp. 1–19.

[8] S. Li and Y. Liu, "Feature Extraction of Lung Sounds Based on Bispectrum Analysis," in *2010 Third International Symposium on Information Processing*, 2010, pp. 393–397.

[9] A. Lanatá, G. Valenza, C. Mancuso, and E. Scilingo, "Robust Multiple Cardiac Arrhythmia Detection Through Bispectrum Analysis," *Expert Systems with Applications*, vol. 38, pp. 6798 – 6804, 2011.

[10] N. Kumar, K. Khaund, and S. M. Hazarika, "Bispectral Analysis of EEG for Emotion Recognition," *Procedia Computer Science*, vol. 84, pp. 31 – 35, 2016.

[11] D. Bhalke, C. R. Rao, and D. Bormane, "Hybridisation of Mel Frequency Cepstral Coefficient and Higher Order Spectral Features for Musical Instruments Classification," *Archives of Acoustics*, vol. 41, pp. 427 – 436, 2016.

[12] D. G. Bhalke, C. B. R. Rao, and D. S. Bormane, "Musical Instrument Classification Using Higher Order Spectra," in *2014 International Conference on Signal Processing and Integrated Networks (SPIN)*, 2014, pp. 40–45.

[13] F. A. H. Jinying, S. B. P. Hongxia, and T. C. B. Shihua, "Bispectrum Entropy Feature Extraction and Its Application for Fault Diagnosis of Gearbox," in *International Conference on Fuzzy Systems*, 2010, pp. 1–6.

[14] I. Mitiche, M. D. Jenkins, P. Boreham, A. Nesbitt, B. G. Stewart, and G. Morison, "Deep Residual Neural Network for EMI Event Classification Using Bispectrum Representations," in *26th European Signal Processing Conference, EUSIPCO 2018*, Rome, Italy, 2018, pp. 186–190.

[15] Z. Zhang, H. Wang, F. Xu, and Y. Jin, "Complex-Valued Convolutional Neural Network and Its Application in Polarimetric SAR Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 12, pp. 7177–7188, Dec 2017.

[16] J. Gao, B. Deng, Y. Qin, H. Wang, and X. Li, "Enhanced Radar Imaging Using a Complex-Valued Convolutional Neural Network," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 1, Jan 2019.

[17] J. Zhang and Y. Wu, "Automatic Sleep Stage Classification of Single-Channel EEG by Using Complex-Valued Convolutional Neural Network," *Biomedical Engineering*, vol. 63, no. 2, Mar 2018.

[18] K. C. Chua, V. Chandran, U. R. Acharya, and C. M. Lim, "Application of Higher Order Statistics/Spectra in Biomedical Signals—A Review," *Medical Engineering and Physics*, vol. 32, pp. 679 – 689, 2010.

[19] J. E. Timperley, "Comparison of PDA and EMI Diagnostic Measurements [for Machine Insulation]," in *Conference Record of the the 2002 IEEE International Symposium on Electrical Insulation*, Boston,USA, 2002, pp. 575–578.