

Article

Not peer-reviewed version

A Deep Learning Approach for Automated Classification of Geomagnetically Induced Current Scalograms

[Tatyana Aksenovich](#) * and [Vasilij Selivanov](#)

Posted Date: 12 December 2023

doi: 10.20944/preprints202312.0747.v1

Keywords: geomagnetically induced currents; autotransformer; continuous wavelet transform; convolutional neural network; binary classification



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

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

A Deep Learning Approach for Automated Classification of Geomagnetically Induced Current Scalograms

Tatyana Aksenovich * and Vasiliy Selivanov

Northern Energetics Research Centre – Branch of the Federal Research Centre “Kola Science Centre of the Russian Academy of Sciences” (NERC KSC RAS), 184209 Apatity, Russia; v.selivanov@ksc.ru (V.S.)

* Correspondence: t.aksenovich@ksc.ru

Abstract: During geomagnetic storms, which are a result of solar wind interaction with the Earth's magnetosphere, geomagnetically induced currents (GICs) begin to flow in the long-distance high-voltage power grids on the Earth's surface. It causes a number of negative phenomena that affect the normal operation of the entire electric power system. To investigate the nature of the phenomenon and its possible effects on transformers, a GIC monitoring system was created in 2011, the devices of which were installed at five substations of the Kola-Karelian power transit in northwestern Russia. Over 12 years of operating the system a large amount of data has been accumulated, which cannot be analyzed manually within a reasonable amount of time. To analyze the constantly growing volume of recorded data effectively, a method for automatic classification of GIC in autotransformer neutral was proposed. The method is based on a continuous wavelet transform of the neutral current data combined with a convolutional neural network (CNN) to classify the obtained scalogram images. As the result of comparing four CNNs with different architectures, a model that showed excellent GIC classification performance on the validation set (100.00% accuracy and loss of 0.0115) was chosen.

Keywords: geomagnetically induced currents; autotransformer; continuous wavelet transform; convolutional neural network; binary classification

1. Introduction

During geomagnetic storms, which are a result of solar wind interaction with the Earth's magnetosphere, geomagnetically induced currents (GICs) begin to flow in the long-distance high-voltage power grids on the Earth's surface [1–5]. One of the paths for their flow is the grounded neutrals of power transformers and autotransformers. Even a slight deviation of the neutral current from zero (tens of amperes) leads to half-cycle saturation of the electrical device core. It, in turn, leads to a number of disruptions of electrical network operation. Examples of detrimental effects of GICs are generation of even and odd harmonics, increase of reactive power demand, appearance of local overheating of transformer windings and elements of its construction, generation of vibrations of transformer core and windings, and unwanted tripping of protective devices [6–8]. Normally the duration of GICs flowing doesn't exceed a few minutes. But, due to repeated GICs impact on transformers, they may fail. There have been several cases in the past where GICs flowing during strong geomagnetic storms have resulted in a blackout of high-voltage power systems in Canada and Sweden [9,10].

More than 12 years ago, a system for monitoring GICs in power lines in northwestern Russia was created [11]. This system was developed by the Polar Geophysical Institute (Russian Academy of Sciences) and the Northern Energetics Research Centre (Kola Science Centre of the Russian Academy of Sciences) in the frameworks of the European Risk from Geomagnetically Induced Currents (EURISGIC) project [12]. It was deployed on the Kola Peninsula and in Karelia and consists of devices at five substations, located in the meridional direction. Figure 1 shows the location of 330

kV transmission lines and part of 110 kV transmission lines in the Murmansk region and Karelia, as well as substations where continuous recording of GICs is carried out.

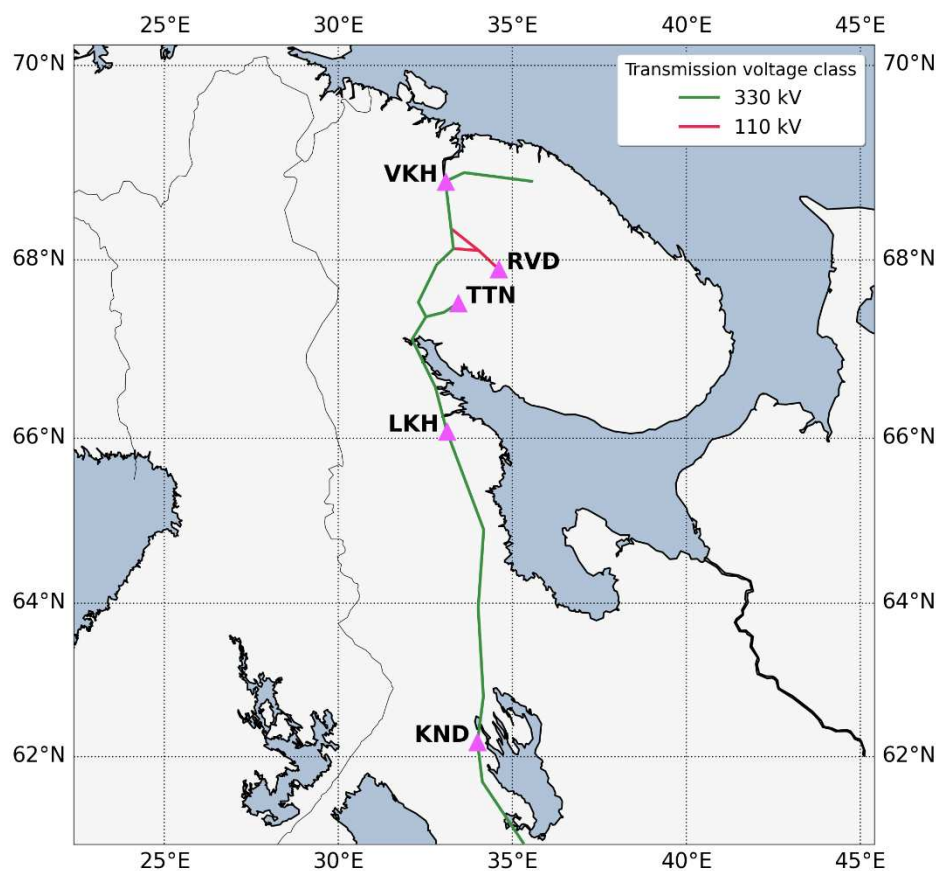


Figure 1. A map of substations equipped with GIC monitoring devices (magenta triangles) in northwestern Russia and schematically drawn transmission lines of the Kola-Karelian power transit (green lines). Red color denotes 110 kV transmission lines, in particular lines, which are directed to RVD substation.

The devices at VKH, TTN, LKH and KND substations installed in the dead-grounded neutral of 330kV autotransformers, and in the neutral of 110kV transformer at the RVD substation. The total current in the network node is twice the measured value, because probes are located on one of the two step-down autotransformers at a substation, with the exception of the KND and RVD substations.

The GIC monitoring system was created to solve the following problems:

- Assessment of GIC impact on power autotransformers of 330 kV power lines;
- Investigation of correlation dependencies between GICs values and the level of geomagnetic activity;
- Identification of lines sections that are most affected by geomagnetic disturbances;
- Development of measures to protect equipment from the GICs influence.

Table 1 provides information about the substations where the devices of the GIC monitoring system are installed, information about the duration of these devices operation and the total size of files recorded at each substation. As can be seen from the numbers in the last column, the size of the data varies depending on the point of study. This is explained by the fact that part of the data is missing. The reasons for the missing files are: hard drive failure; shutdown of the device due to external interference; problems with data transfer to the server associated with weak cell signal reception [13]. Since the start of the GIC monitoring system, more than 16,000 files have been created containing data of currents in the autotransformers neutrals. The total size of these files is about 27

GB at the beginning of 2023. Constantly growing volume of recorded neutral current data cannot be analyzed manually within a reasonable amount of time.

Table 1. Basic information about the GIC monitoring system in power lines in northwestern Russia.

Substation		Geographical coordinates		The period of registration	Volume of recorded data (GB)
Code	Name	Latitude (°N)	Longitude (°E)		
VKH	Vykhodnoy	68.83	33.08	October 2011 – until now	6.49
RVD	Revda	67.90	34.61	May 2011 – until now	5.65
TTN	Titan	67.53	33.44	June 2010 – December 2014	2.28
LKH	Loukhi	66.08	33.12	September 2011 – until now	5.67
KND	Kondopoga	62.22	34.36	September 2011 – until now	6.67

The purpose of this study is to develop a convolutional neural network (CNN) model for automatic analysis of 12 years of data accumulated by the GIC monitoring system in northwestern Russia. To train the CNN model, a dataset containing 800 scalogram images was created. The dataset includes images of two classes: GIC and geomagnetically quiet hours. To select the most suitable neural network for solving the problem, the performance of four models with different configurations was compared.

The structure of the paper is as follows. Section 2 provides an overview of existing research that applies continuous wavelet transform (CWT) and convolutional neural networks to analyze nonstationary signals in power systems, including geomagnetically induced currents. Section 3 is dedicated to the methodology and the created dataset. Subsection 3.1 introduces the basic concepts of CNNs, describes the layers required to build ones, and explains the choice of performance metrics for binary classification. Subsection 3.2 presents a description of the continuous wavelet transform method. Subsection 3.3 details the dataset specially created for this study, consisting of two-dimensional images of wavelet spectra (scalograms) of the current in autotransformer neutral. Section 4 presents the results of a performance comparison of four CNN models and a detailed description of the architecture of the chosen model. Conclusions are presented in Section 5.

2. Related Work

This section contains a review of recent studies that employed continuous wavelet transform and convolutional neural networks to analyze nonstationary signals in power systems, including geomagnetically induced currents.

Wang et al. [14] proposed a machine learning approach to detection of geomagnetically induced currents in power grids. The authors used a CNN-based architecture to classify scalogram and spectrogram images obtained by continuous wavelet transform and short-time Fourier transform (STFT) of currents of single-phase current transformers, respectively. The proposed framework, consisting of a hybrid feature extractor using Gaussian CWT and STFT and CNN, successfully achieved good performance even under low-intensity GICs. The framework works very quickly (detection is achieved within 30 ms). Thus, it can be applied for online monitoring applications in power systems.

A study conducted by Li et al. [15] proposed a deep learning method for fault diagnosis of dry-type transformers using vibration signals. To convert the transformer vibration signals to scalogram images, a CWT method was used, which allows for the extraction of fault features from different conditions. The proposed diagnosis method achieved the mean accuracy of 99.95%. The feature extraction and classification process took less than 7 s, which potentially provides fast and accurate online diagnosis of transformer faults.

One of the studies that used deep learning algorithms for power quality disturbances classification was conducted by Salles and Ribeiro [16]. This study has examined the application of several CNN models for voltage disturbances classification. The results obtained showed high accuracy for a CNN created from scratch and ResNet-50 (transfer learning). This study confirmed the suitability of using CNN for voltage disturbances classification and the advantage of using scalograms to characterize these disturbances.

3. Data and Methods

3.1. Convolutional Neural Networks

To analyze data from devices of the GICs monitoring system, we use a convolutional neural network for binary classification of scalograms obtained after applying a CWT to the neutral currents. The choice of CNN is justified by the fact that this deep learning algorithm is the most popular, including for solving image classification tasks [17,18]. Classification is a concept that categorizes a dataset, in our case images, into classes. Convolutional neural networks are a class of feedforward learning algorithms. They are designed to automatically and adaptively learn spatial hierarchies of features, from minor features to the global ones. This aspect improves the generalization ability of such neural networks and makes them perfect for image classification.

3.1.1. Architecture of Convolutional Neural Networks

The typical CNN structure consists of three types of layers, which are the convolutional layer, pooling layer, and fully connected layer. In addition to the above main layers of CNN, it may include layers such as non-linear activation layer, classification layer, dropout layer, and preprocessing layer. The convolutional and non-linear layers are very commonly combined into one. Brief information about these layers is presented below.

The convolutional layer, as its name suggests, is the most significant component in CNN architecture and plays an important role in its operation. The main purpose of these layers is feature extraction: they learn feature representations of their input images. The convolutional layer consists of a set of convolutional filters called kernels. The kernel is a grid of discrete numbers or values called kernel weights. At the beginning of the CNN training process, random numbers are assigned as the weights of the kernel. The input image is convolved with filters to generate the feature map of the output. The kernel slides over the whole image horizontally and vertically. A dot product is determined between the input image and the kernel, where their corresponding values are simultaneously multiplied and then summed up to calculate a single scalar value. The whole process is repeated until no further sliding is possible. The calculated dot product values create the output feature map.

By optimizing the output of convolutional layers, they are able to significantly reduce model complexity. The output is optimized through three hyperparameters: the depth, the stride and setting zero-padding.

Non-linear activation functions allow for the extraction of non-linear features. The main purpose of all types of activation functions is to map the input to the output. The input value is determined by computing the weighted summation of the neuron input. Thus, the activation function decides which neuron to fire with reference to a particular input by creating the corresponding output. These layers allow CNN to learn more complex things. Non-linear activation layers are applied after all convolutional and fully connected layers, i.e. after layers with weights. The following types of

activation functions are commonly used in CNN: sigmoid, tanh, ReLU (Rectified Linear Units), leaky and noisy ReLU [19]. In this study, ReLU is chosen as the activation function [20].

The purpose of the pooling layer is to reduce the spatial resolution of the feature maps, due to which spatial invariance to input distortions and shifts is achieved. At this layer, part of data is discarded, decreasing the complexity of the model. This, in turn, reduces the chance of it overfitting. There are several types of pooling methods. These methods include average pooling, min pooling, max pooling, tree pooling, gated pooling and global average pooling [21]. This study uses max pooling because it provides unique performance features. Max pooling calculates the maximum value in each patch of the feature map covered by the kernel, thereby highlighting the most present feature in the patch.

The fully connected layers are located at the end of each CNN architecture. These layers got their name because of the approach of the same name, which is that each neuron in the layer is connected to all neurons of the preceding layer. The input of these layers is in the form of a vector, which is created from the feature maps after flattening. The fully connected layer is used as a CNN classifier: it takes input from feature extraction stages and globally analyzes the output of all previous layers. The result is a non-linear combination of selected features that are used to classify the data.

The final classification is performed on the last layer, which represents the output layer of the CNN. To calculate the class predicted error, loss functions are used in the output layer, the error value reveals the difference between the actual output and the predicted one. The predicted error is optimized through the CNN learning process. In the case of binary classification, the following loss functions are used: binary cross-entropy loss, Hinge loss and squared Hinge loss. The proposed CNN uses the binary cross-entropy loss function, a detailed description of which is given in subsection 3.1.2.

One of the problems encountered with training neural networks on small datasets is overfitting [22]. Along with underfitting, such an undesirable machine learning behavior is the main cause for poor performance of its algorithms. Overfitting is a phenomenon when a model performs well on training data, but shows not so good performance on test data. Thus, the poor ability of the model to generalize new, previously unused examples makes it useless. In order to prevent overfitting, a regularization technique called dropout is often used in neural networks [23]. The key idea of the dropout layer is that randomly selected units from the neural network are ignored during training of a model, i.e. "dropped out" randomly. This allows for a reduction in the interdependent learning of neurons: since a neuron can no longer rely on the presence of other neurons, it is forced to learn more robust features. The dropout layer is placed before the layer to which we want to apply the dropout.

The preprocessing layers are used in the early stages of a CNN, and include data augmentation. Data augmentation is a set of methods used to artificially expand the size of the training dataset [24]. This technique is often used on small datasets, as it allows the model to avoid overfitting and improve its performance. The essence of this method is to perform random (but realistic) transformations to existing images to create a set of new variants without altering their natures. Typical data augmentation approaches in image classification are cropping, reflection, translation, rotation, zoom and contrast adjustment. The preprocessing layers are only active during model training.

3.1.2. Performance Metrics for Binary Classification

The choice of neural network performance metrics is an integral part of creating any model. The main purpose of metrics is to monitor and measure the performance of the model during its training and testing. They indicate weaknesses in the model and dataset, the elimination of which leads to an increase in the generalization ability of the final model. Depending on the objectives and specifics of the research, different metrics are used that are characteristic of a particular field of science. Frequently the choice of methods of evaluating model quality based on the experience of researchers in the same field, as well as on the experience of the researchers themselves. In the case of binary classification, metrics such as accuracy, precision, recall, F1 score, binary cross-entropy loss, Receiver Operator Characteristic and Area Under Curve (ROC-AUC), Matthews Correlation Coefficient (MCC) are used to evaluate the performance of the neural network [25]. In this paper, the following

metrics were chosen to evaluate the performance of the created CNN models: accuracy and binary cross-entropy loss function. Both metrics are often used when debugging classifiers with two classes [26]. One of the necessary conditions for the correct use of accuracy as an evaluation metric is the presence of a balanced dataset. In this study, this condition is completely fulfilled.

Accuracy is the most widely spread measure of performance, which describes the overall prediction accuracy of a model across all classes. It is calculated as the ratio between the number of correct class predictions to the total number of predictions. In other words, the metric is the fraction of predictions trained model got right. Accuracy is measured as a percentage, ranging from 0% to 100%. A higher value of the metric corresponds to a better-performing classifier, while a lower one indicates more errors in class predictions.

For binary classification, accuracy can be defined in terms of positives and negatives according to the following equation [27]:

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} , \quad (1)$$

where TP (true positive) is a number of correctly classified positive samples; TN (true negative) is a number of correctly classified negative samples; FP (false positive) is a number of misclassified positive samples; and FN (false negative) is a number of misclassified negative samples.

Binary cross-entropy loss function is a metric in machine learning, used to evaluate how well a classification model with two classes performs. This function takes two values, actual and predicted, and returns the comparative evaluation that shows how close the predicted probability is to the actual value (0 or 1). Thus, the loss function rewards the model for giving a correct prediction with a low loss and penalizes for a wrong outcome with a high error value. In the case of an ideal model, the value of binary cross-entropy loss is zero. The main goal throughout training of CNN is to minimize the loss value, i.e. to minimize the difference between the predicted probability and true label in a final model.

Binary cross-entropy is a special case of cross-entropy [28], and is determined by the formula:

$$BCE_{loss} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)], \quad (2)$$

where N is a number of final layer outputs (1 for sigmoid); y_i is the binary indicator (0 or 1) denoting the class for the sample i ; and p_i is the classifier's predicted probability distribution.

3.2. Continuous Wavelet Transform

The continuous wavelet transform is widely spread spectral analysis method of nonstationary signals, in particular geomagnetically induced currents [29–32]. In the case of intellectual fault diagnosis of power systems, the CWT is often used as a feature extraction method from different transformer signals such as vibrations, currents and voltages [14–16]. The CWT provides an accurate representation about changes in frequency characteristics of a signal over time. The essence of the method is to convolve the analyzed signal with a scaled and translated version of the analyzing function, which is called the mother wavelet. The wavelet is a mathematical function that most frequently has a wave-like form and its amplitude tends to zero with distance from the origin of coordinates. An important property of wavelets is time-frequency localization, which makes it possible to obtain frequency spectrum features of a signal with respect to time. As the result of the CWT, the wavelet coefficients are calculated, the values of which are directly proportional to the degree of similarity between the analyzed signal and the selected wavelet. The continuous wavelet transform of discrete time series is defined by the following equation [33]:

$$W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \psi^* \left[\frac{(n'-n)\delta t}{s} \right], \quad (3)$$

where n is the localized time index, s is the wavelet scaling factor, ψ is the mother wavelet, the (*) symbolizes a complex conjugation, and δt is the sampling interval.

The graphical representation of a three-dimensional array of wavelet coefficients $W_n(s)$ is a scalogram: the x axis indicates time, the y axis represents a frequency, the value of which is inversely proportional to wavelet scale (width), and colors indicate the wavelet coefficients values that characterize the similarity degree between the analyzed signal and the wavelet. The scalogram is calculated as the square of the modulus of the wavelet coefficients values $|W_n(s)|^2$. Two-dimensional images of wavelet spectra provide a visual representation of various phenomena in the researched signal. They are used in such areas of machine learning as classification and pattern recognition [16].

To compute scalograms for the dataset, this paper uses open source library PyWavelets [34], which contains a number of CWT-compatible mother wavelets.

3.3. The Geomagnetically Induced Current Scalograms Dataset

The quality of a dataset on which a CNN is trained has a strong impact on its classification performance. In order to be considered high quality, a labeled dataset must satisfy two basic conditions. Firstly, the images must be directly related to the task. Secondly, the images in each class must be uniform.

To train the CNN, a dataset containing two-dimensional images of wavelet spectra of the current in autotransformer neutral was created. The produced dataset consists of 800 RGB scalogram images with a size of 224x224 pixels. Since the goal of this study is to detect the presence or absence of GIC in the neutral current (binary classification), the dataset includes images of two classes: GIC and geomagnetically quiet hours (Figure 2). Quiet hours mean the absence, in addition to GICs, of any other phenomena reflected in the neutral current, such as thunderstorms and commutations. The number of examples in the "GIC" class (400) equals the number of examples in "quiet hours" class (400), which makes the dataset balanced. This characteristic of the dataset makes training a classifier easier by helping to prevent the model from becoming biased towards the class with more examples, and potentially improves the accuracy of the future model.

Each image in the dataset is a scalogram of the neutral current with duration of 60 min and a frequency range from 1.36 mHz to 5.42 mHz. The choice of such time interval for the study is based on several reasons. Firstly, files that contain neutral current data are transferred from GIC monitoring devices to the server every hour. In the future, the proposed CNN model will be able to analyze new measurements in real time. Secondly, GIC events have a duration of up to several hours. The CWT was applied to the currents data with a sampling interval of 1 s. As previous study has shown [32], a sampling rate of 1 Hz is sufficient for reliable estimate of GICs values, including peak values, as well as for determining their time localization. The non-orthogonal Morlet wavelet was chosen as the mother wavelet. The benefits of using this wavelet function as the most suitable function for the analysis of GICs in autotransformers neutrals have also been considered and proven in practice in a previous study [32]. In the scalograms colors denote the values of wavelet coefficients (Figure 2) and characterize the degree of matching between the researched signal and the wavelet. Red color corresponds to the maximum value of the energy and blue color corresponds to the minimum one.

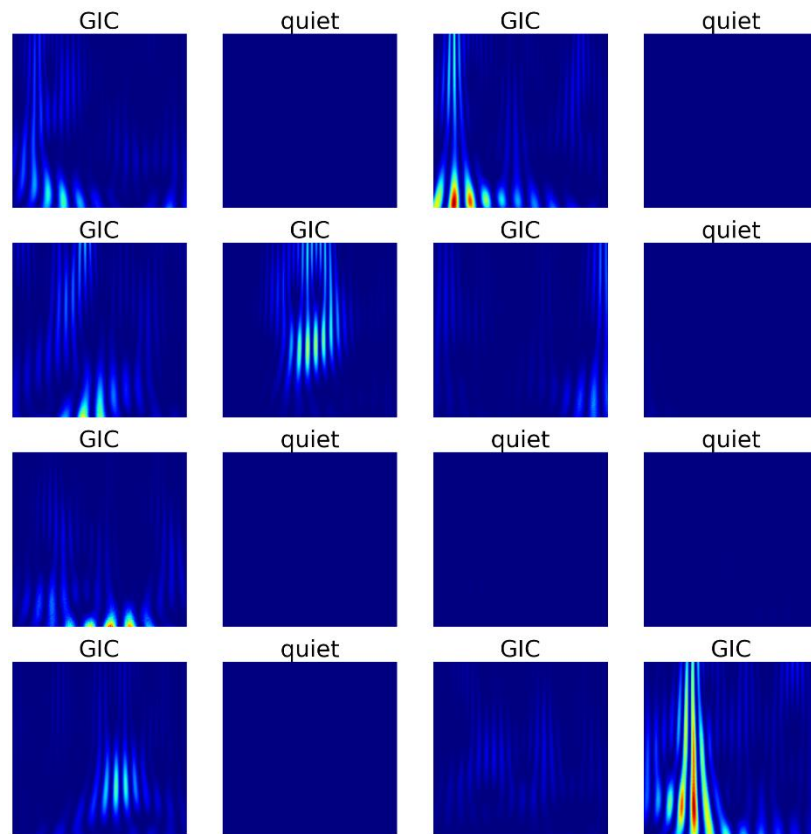


Figure 2. Sample scalogram images taken from different days at different substations from the created dataset.

4. Geomagnetically Induced Currents Classification by Convolutional Neural Network

The main goal of this study is to create a binary classifier of scalograms of currents in an autotransformer neutral for analyzing data from the GICs monitoring system in northwestern Russia. For this purpose, a comparative analysis of four CNN models with different configurations was carried out. All models were built from scratch. The first model has the following architecture: four convolutional layers, four max pooling layers, flatten layer, and two fully connected layers. The architectures of the second and third models differ from the first one by the presence of a dropout layer after the fourth max pooling layer and preprocessing layers before the first convolutional layer, respectively. The fourth model contains all of the above layers in its architecture.

The models training and testing were conducted using Intel(R) Core(TM) i5-11400 CPU (Central Processing Unit) at 2.6 GHz. Over computer specifications used in the experiment are 16 GB RAM (Random Access Memory) and Microsoft Windows 10 Pro 64-bit OS (Operating System). A well-known Keras library was chosen as a deep learning framework [35]. Keras is the open-source high-level API (Application Programming Interface), running on top of the TensorFlow platform for machine learning.

CNN models were trained by employing 80% (640) of the scalogram images and validated using the rest 20% (160). Accuracy and binary cross-entropy loss function were used as evaluation metrics. Training all models was done for 15 epochs. Table 2 shows the results of models performance validation.

Table 2. Comparison of CNN models performance with different configurations on the validation set.

Number	CNN models	Time, min	Binary cross-entropy	Accuracy, %
1	Simple CNN	03:06	0.0098	99.37
2	CNN with dropout	03:08	0.0092	99.37
3	CNN with augmentation	03:19	0.0344	98.75
4	CNN with augmentation and dropout	03:22	0.0115	100.00

In general, it can be stated that all CNN models showed good results. An accuracy of 99.37% with 0.0098 binary cross-entropy loss was obtained for the first CNN model. The second model of binary classifier successfully classified the scalogram images with 99.37% accuracy and the lowest binary cross-entropy loss of 0.0092. For the third CNN model, the lowest accuracy of 98.75% with the highest value of binary cross-entropy loss of 0.0344 was obtained. The best accuracy of 100.00% with binary cross-entropy loss of 0.0115 was estimated for the fourth CNN model. The difference in the total training time (15 epochs) between the fastest and longest CNN model was only 16 s. As the result of comparing the performance of CNN models with different structures, the last neural network model containing image augmentation layers and a dropout layer was selected, since this one achieved maximum accuracy with a relatively low loss value. Figure 3 shows the learning curves for this model on the training and validation sets.

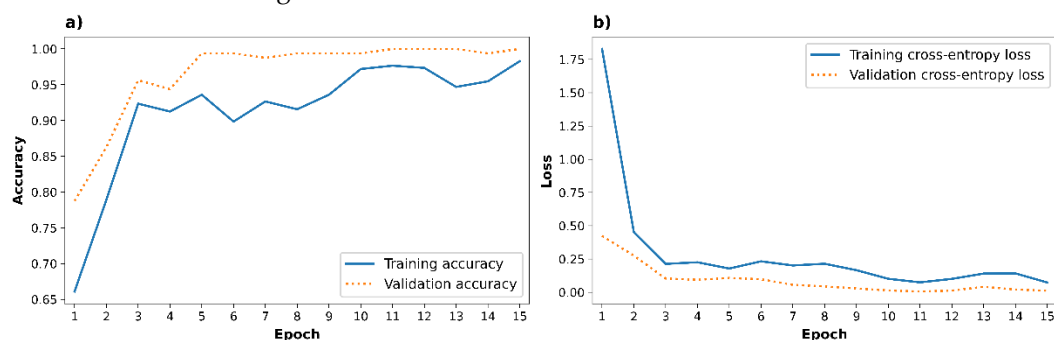


Figure 3. The change of accuracy (a) and cross-entropy loss (b) versus epoch values during network training and validation.

The architecture of the proposed CNN model for scalogram classification of currents in autotransformer neutral is illustrated in Figure 4. It consists of four convolutional layers, four max pooling layers, dropout layer, flatten layer, and two fully connected layers. Each convolutional layer and the first fully connected layer are accompanied by a ReLU activation function, which is used to provide an efficient training performance. The kernel size of each subsequent convolutional layer is reduced, making it possible to extract local features from images at different filter scales. Since our goal is binary classification, the number of neurons in the last layer is equal to 1. The sigmoid function was chosen as the activation function for the output layer because its use reduces the output to a value from 0 to 1 representing the probability. A binary cross-entropy loss is used to calculate the discrepancy between the predicted and actual value. The loss function is purpose-built for binary classifiers. The rate for the dropout layer is set to 0.5 which means that half of the units are randomly ignored during model training. The layers listed above are preceded by preprocessing layers, which are used to generate additional training images by augmenting existing examples. The random augmentation transforms that were applied are reflection ('horizontal'), rotation (set to 0.01) and zoom (height factor set to (-0.001, -0.1), width factor set to (-0.01, -0.6)). Data augmentation was

employed to improve the generalization of the proposed CNN model. The Adam (Adaptive Moment Estimation) learning algorithm [36] was used as an optimizer, the learning rate of which was set to 0.0001. This is a popular alternative for the classical stochastic gradient descent process because it automatically tunes itself and achieves good results fast. The learning rate is an important parameter for neural networks, and controls the adjustment of weights at each iteration, taking into account the loss function. Small values of this parameter increase the accuracy of the algorithm's tuning and potentially reduce the training error.

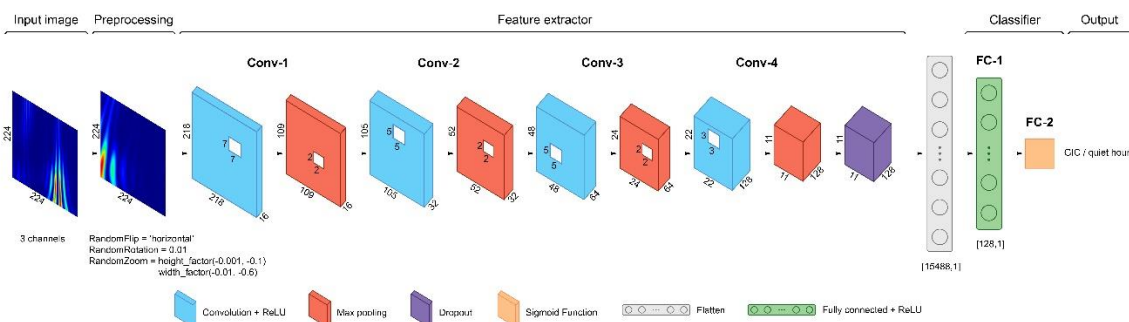


Figure 4. Architecture of proposed CNN for GIC classification in autotransformer neutral.

5. Conclusions

This paper proposes a method for GICs detection in neutral current of autotransformer based on continuous wavelet transform and convolutional neural network. Although similar researches have been performed in the past, the paper is the first to propose an analysis of data of currents in autotransformers neutrals from operating substations in northwestern Russia, based on binary classification of current scalograms. To train the CNN model, a dataset containing 800 scalogram images of two classes: GIC and geomagnetically quiet hours was created. As the result of comparing the performance of four CNNs with different architectures, a model that showed excellent GIC classification performance on the validation set (100.00% accuracy and loss of 0.0115) was chosen. The proposed CNN model, in addition to the main layers, includes preprocessing layers and dropout layer. During further research it is planned to analyze 12 years of data from the GIC monitoring system using the proposed CNN-based classifier to identify lines sections that are most affected by geomagnetic disturbances, as well as to develop measures to protect equipment from the GICs influence.

Author Contributions: Conceptualization, T.A. and V.N.; methodology, T.A. and V.N.; software, T.A.; formal analysis, T.A.; investigation, T.A. and V.N.; writing—original draft preparation, T.A.; writing—review and editing, V.N.; visualization, T.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Russian Science Foundation, grant number 22-29-00413.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw GIC data may be accessed at <http://gic.en51.ru/>. The dataset created and analyzed during the current study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Boteler, D.H. Assessment of Geomagnetic Hazard to Power Systems in Canada. *Nat. Hazards* 2001 232 **2001**, 23, 101–120, doi:10.1023/A:1011194414259.
2. Erinmez, I.A.; Kappenman, J.G.; Radasky, W.A. Management of the Geomagnetically Induced Current Risks on the National Grid Company's Electric Power Transmission System. *J. Atmos. Solar-Terrestrial Phys.* **2002**, 64, 743–756, doi:10.1016/S1364-6826(02)00036-6.
3. Pirjola, R. Effects of Space Weather on High-Latitude Ground Systems. *Adv. Sp. Res.* **2005**, 36, 2231–2240, doi:10.1016/J.ASR.2003.04.074.
4. Liu, C.M.; Liu, L.G.; Pirjola, R. Geomagnetically Induced Currents in the High-Voltage Power Grid in China. *IEEE Trans. Power Deliv.* **2009**, 24, 2368–2374, doi:10.1109/TPWRD.2009.2028490.
5. Marshall, R.A.; Dalzell, M.; Waters, C.L.; Goldthorpe, P.; Smith, E.A. Geomagnetically Induced Currents in the New Zealand Power Network. *Sp. Weather* **2012**, 10, doi:10.1029/2012SW000806.
6. Molinski, T.S. Why Utilities Respect Geomagnetically Induced Currents. *J. Atmos. Solar-Terrestrial Phys.* **2002**, 64, 1765–1778, doi:10.1016/S1364-6826(02)00126-8.
7. Kappenman, J.G. Geomagnetic Disturbances and Impacts upon Power System Operation. In *Electric Power Generation, Transmission, and Distribution: The Electric Power Engineering Handbook*; CRC Press, 2018; pp. 1–22 ISBN 9781315222424.
8. Rajput, V.N.; Boteler, D.H.; Rana, N.; Saiyed, M.; Anjana, S.; Shah, M. Insight into Impact of Geomagnetically Induced Currents on Power Systems: Overview, Challenges and Mitigation. *Electr. Power Syst. Res.* **2021**, 192, 106927, doi:10.1016/J.EPSR.2020.106927.
9. Pulkkinen, A.; Lindahl, S.; Viljanen, A.; Pirjola, R. Geomagnetic Storm of 29-31 October 2003: Geomagnetically Induced Currents and Their Relation to Problems in the Swedish High-Voltage Power Transmission System. *Sp. Weather* **2005**, 3, S08C03, doi:10.1029/2004SW000123.
10. Guillon, S.; Toner, P.; Gibson, L.; Boteler, D. A Colorful Blackout: The Havoc Caused by Auroral Electrojet Generated Magnetic Field Variations in 1989. *IEEE Power Energy Mag.* **2016**, 14, 59–71, doi:10.1109/MPE.2016.2591760.
11. Barannik, M.B.; Danilin, A.N.; Kat'kalov, Y. V.; Kolobov, V. V.; Sakharov, Y.A.; Selivanov, V.N. A System for Recording Geomagnetically Induced Currents in Neutrals of Power Autotransformers. *Instruments Exp. Tech.* **2012**, 55, 110–115, doi:10.1134/S0020441211060121.
12. Viljanen, A. European Project to Improve Models of Geomagnetically Induced Currents. *Sp. Weather* **2011**, 9, S07007, doi:10.1029/2011SW000680.
13. Selivanov, V.N.; Aksenovich, T. V.; Bilin, V.A.; Kolobov, V. V.; Sakharov, Y.A. Database of Geomagnetically Induced Currents in the Main Transmission Line "Northern Transit." *Solar-Terrestrial Phys.* **2023**, 9, 93–101, doi:10.12737/STP-93202311.
14. Wang, S.; Dehghanian, P.; Li, L.; Wang, B. A Machine Learning Approach to Detection of Geomagnetically Induced Currents in Power Grids. *IEEE Trans. Ind. Appl.* **2020**, 56, 1098–1106, doi:10.1109/TIA.2019.2957471.
15. Li, C.; Chen, J.; Yang, C.; Yang, J.; Liu, Z.; Davari, P.; Li, C.; Chen, J.; Yang, C.; Yang, J.; et al. Convolutional Neural Network-Based Transformer Fault Diagnosis Using Vibration Signals. *Sensors* 2023, Vol. 23, Page 4781 **2023**, 23, 4781, doi:10.3390/S23104781.
16. Salles, R.S.; Ribeiro, P.F. The Use of Deep Learning and 2-D Wavelet Scalograms for Power Quality Disturbances Classification. *Electr. Power Syst. Res.* **2023**, 214, 108834, doi:10.1016/J.EPSR.2022.108834.
17. Rawat, W.; Wang, Z. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Comput.* **2017**, 29, 2352–2449, doi:10.1162/NECO_A_00990.
18. Khan, A.; Sohail, A.; Zahoor, U.; Qureshi, A.S. A Survey of the Recent Architectures of Deep Convolutional Neural Networks. *Artif. Intell. Rev.* 2020 538 **2020**, 53, 5455–5516, doi:10.1007/S10462-020-09825-6.
19. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. *J. Big Data* 2021 81 **2021**, 8, 1–74, doi:10.1186/S40537-021-00444-8.
20. Hahnloser, R.H.R.; Sarpeshkar, R.; Mahowald, M.A.; Douglas, R.J.; Seung, H.S. Digital Selection and Analogue Amplification Coexist in a Cortex-Inspired Silicon Circuit. *Nat.* 2000 4056789 **2000**, 405, 947–951, doi:10.1038/35016072.
21. Lee, C.Y.; Gallagher, P.W.; Tu, Z. Generalizing Pooling Functions in Convolutional Neural Networks: Mixed, Gated, and Tree. *Proc. 19th Int. Conf. Artif. Intell. Stat. AISTATS 2016* **2015**, 464–472.
22. Hinton, G.E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R.R. Improving Neural Networks by Preventing Co-Adaptation of Feature Detectors. **2012**.
23. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, 15, 1929–1958, doi:10.5555/2627435.
24. Shorten, C.; Khoshgoftaar, T.M. A Survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, 6, 1–48, doi:10.1186/S40537-019-0197-0/FIGURES/33.
25. Powers, D.M.W. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation. *ArXiv* **2011**, abs/2010.1.

26. Canbek, G.; Taskaya Temizel, T.; Sagioglu, S. PToPI: A Comprehensive Review, Analysis, and Knowledge Representation of Binary Classification Performance Measures/Metrics. *SN Comput. Sci.* **2023**, *4*, 1–30, doi:10.1007/S42979-022-01409-1/TABLES/6.
27. Lee, W.; Lee, D.; Lee, S.; Jun, K.; Kim, M.S. Deep-Learning-Based ADHD Classification Using Children's Skeleton Data Acquired through the ADHD Screening Game. *Sensors* **2023**, *Vol. 23*, Page 246 **2022**, *23*, 246, doi:10.3390/S23010246.
28. Good, I.J. Some Terminology and Notation in Information Theory. *Proc. IEE Part C Monogr.* **1956**, *103*, 200, doi:10.1049/PI-C.1956.0024.
29. Falayi, E.O.; Ogunmodimu, O.; Bolaji, O.S.; Ayanda, J.D.; Ojoniyi, O.S. Investigation of Geomagnetic Induced Current at High Latitude during the Storm-Time Variation. *NRIAG J. Astron. Geophys.* **2017**, *6*, 131–140, doi:10.1016/j.nrjag.2017.04.010.
30. Adhikari, B.; Sapkota, N.; Dahal, S.; Bhattarai, B.; Khanal, K.; Chapagain, N.P. Spectral Characteristic of Geomagnetically Induced Current during Geomagnetic Storms by Wavelet Techniques. *J. Atmos. Solar-Terrestrial Phys.* **2019**, *192*, 104777, doi:10.1016/j.jastp.2018.01.020.
31. Xu, W.-H.; Xing, Z.-Y.; Balan, N.; Liang, L.-K.; Wang, Y.-L.; Zhang, Q.-H.; Sun, Z.-D.; Li, W.-B. Spectral Analysis of Geomagnetically Induced Current and Local Magnetic Field during the 17 March 2013 Geomagnetic Storm. *Adv. Sp. Res.* **2022**, *69*, 3417–3425, doi:10.1016/j.asr.2022.02.025.
32. Aksenovich, T. V.; Bilin, V.A.; Saharov, Y.A.; Selivanov, V.N. Wavelet Analysis of Geomagnetically Induced Currents during the Strong Geomagnetic Storms. *Russ. J. Earth Sci.* **2023**, *22*, 1–12, doi:10.2205/2022ES000825.
33. Torrence, C.; Compo, G.P.; Torrence, C.; Compo, G.P. A Practical Guide to Wavelet Analysis. *BAMS* **1998**, *79*, 61–78, doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.
34. Lee, G.R.; Gommers, R.; Waselewski, F.; Wohlfahrt, K.; O'Leary, A. PyWavelets: A Python Package for Wavelet Analysis. *J. Open Source Softw.* **2019**, *4*, 1237, doi:10.21105/JOSS.01237.
35. Chollet, F. Keras Documentation Available online: <https://keras.io> (accessed on 31 October 2023).
36. Kingma, D.P.; Ba, J.L. Adam: A Method for Stochastic Optimization. *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.* **2014**.

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