1. Zuzanna KRAWCZYK-BORYSIAK, 2. Paweł KLUGE, 3. Andrzej ŁASICA, 4. Przemysław SUL, 5. Maciej CIUBA

Warsaw University of Technology
ORCID: 1. 0000-0002-3897-1078; 2. 0000-0002-7238-3720; 3. 0000-0002-4572-9570; 4. 0000-0002-4327-9334; 5. 0000-0001-6022-5277

doi:10.15199/48.2025.03.49

Methods of classifying voltage surges using Deep Neural Networks

Abstract. *The paper focuses on exploring the potential application of neural networks for the classification of voltage surges compliance with the norm. Three potential neural network architectures were considered for the task - a convolutional neural network (further referred to as CNN), a model combining convolutional and LSTM layers (CNN+LSTM) and a transformer model. The best results were achieved by the simple transformer model (accuracy of 93% on the test dataset), followed by CNN+LSTM model (accuracy: 81%), and CNN (accuracy: 69%).*

Streszczenie*. Artykuł koncentruje się na badaniu potencjalnego zastosowania sieci neuronowych do klasyfikacji zgodności udarów napięciowych z normą. Do tego zadania rozważono trzy potencjalne architektury sieci neuronowych - konwolucyjną sieć neuronową (CNN), model łączący warstwy konwolucyjne i LSTM (CNN+LSTM) oraz model transformatora. Najlepsze wyniki uzyskał prosty model transformatora (dokładność 93% w zestawie danych testowych), następnie model CNN+LSTM (dokładność: 81%) i CNN (dokładność: 69%)* ***(Metody klasyfikacji udarów napięciowych za pomocą głębokich sieci neuronowych)***

Keywords: neural networks, CNN, LSTM, transformer, voltage surges, signal classification
Słowa kluczowe: sieci neuronowe, CNN, LSTM, transformer, udary napięciowe, klasyfikacja sygnałów

Introduction

Insulating systems of high-voltage electrical devices used in the power industry are exposed to various types of exposure during their operation. Overvoltages in the nature of lightning surges, resulting from the direct or indirect impact of atmospheric discharges - lightning, constitute a certain group of these exposures [1]. For this reason, the need to perform and design an acceptance tests of insulating systems of devices that may be subject to such impacts is justified. Insulating materials, insulators, insulating systems of switches, transformers, rotating machines, cables, etc. are tested. The aim is to verify the structural correctness of a given type of device or a specific product by checking the electrical strength of its insulating system with a test voltage of the appropriate shape and value according to standards [2, 3]. Based on the literature [4, 5, 6, 7], the authors decided to use artificial intelligence tools and algorithms to classify correct voltage surges.

The paper focuses on exploring the potential application of neural networks for the classification of voltage surges compliance with the norm. Therefore, the training dataset comprising 269 voltage surges - 134 correct ones and 135 with incorrect parameters was generated. Three potential neural network architectures were considered for the task – a convolutional neural network (further referred to as CNN), a model combining convolutional and Long short-term memory (LSTM) layers (CNN+LSTM) and a Transformer model. The architecture of the models is adjusted in such a way to mitigate the relatively small number of training data. The results show that the neural networks can be successfully applied to the task of voltage surge validation and allow for a quick assessment of whether the generated impulse meets the appropriate standard, with Transformer model achieving most promising results.

The paper is structured as follows: the second and third chapters are focused on the description of normalized voltage surge, and the procedure of training data generation. The fourth chapter describes performed experiments including, data preprocessing steps, the architecture of models used and hyper-parameters tuning as well as models training. Final chapters include discussion of the results, summary of our work and further possible directions of research.

Lighting Voltage Surge Characteristic

The values of times T_1 and T_2 determine the shape of the impulse, which is designated by the symbol T_1/T_2.
An example of a voltage impulse is presented in Figure 1.

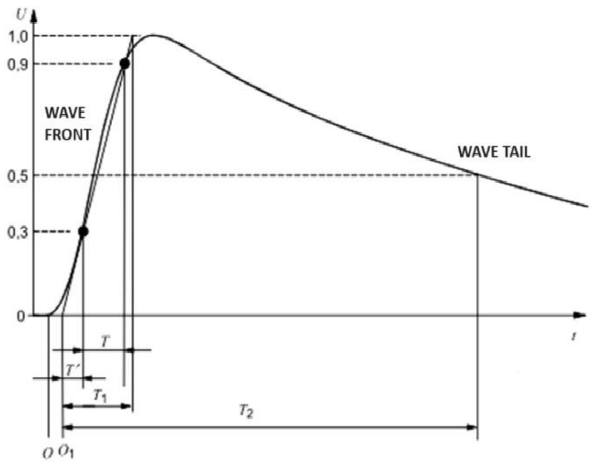

Fig.1. Lightning impulse voltage course as a function of time [1].

In most countries, a standard lightning surge is assumed to have a shape defined by the frontal duration (front time) of T_1 = 1.2 μs and the time to half-value at the tail of the surge T_2 = 50 μs. Therefore, a standard lightning impulse has the shape T_1/T_2 = 1.2/50. The permissible deviations of the front duration should not exceed ± 30\% of the standardized value T_1 = 1.2 μs and the time to half-value ± 20% of the standardized value T_2 = 50 μs. The selection of times T_1 and T_2 was made based on the analysis of the shape of voltage waveforms induced in electrical systems by lightning discharges.

Data acquisition setup

The authors conducted the research at the High Voltage Hall of the Warsaw University of Technology. The voltage surges were generated at a test stand (Figure 3). A detailed electrical diagram of the measuring system is shown in Fig. 2. Voltage surges were recorded using the Dr. Strauss VFT Measuring System with capacitive voltage divider North Star

High Voltage VD-200. During the research, the authors are using known and simplified relationships [10]:

(1) $$T_1 = 3.25\tau_1 = 3.25C_2R_1$$

(2) $$T_2 = 0.7\tau_2 = 0.7C_1R_2$$

where τ_1 and τ_2 means time constants, R_1 is the sum of the R_f resistors, C_2 means load capacitor, R_2 corresponds to the discharge resistance R_d and C_1 is the resultant capacitance of the series-connected generator capacitors. By appropriately modifying the resistance and capacitance of the system, the authors obtained various lightning surges.

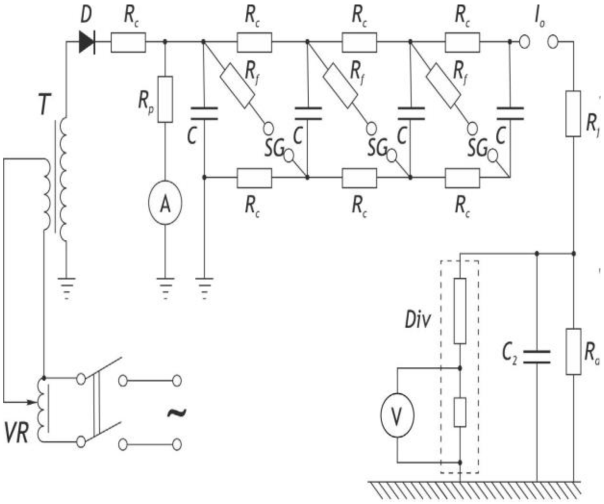

Fig. 2. Electrical diagram of the voltage surge generation system, where: T – transformer, D – rectifier, R_c – charging resistor, SG – spark gap, C – discharge capacitor, C_2 – load capacitor, R_f – front resistor, R_d – discharge resistor, Div - divider, VR - voltage regulator.

Fig. 3. Lightning surge generator station view.

The research involved voltage signals without filtering and denoising. The registered signals consisted of 269 measurements of voltage surges (134 - correct ones and 135 incorrect). During the research, it was decided to add three groups of non-standardized voltage surges. The recorded incorrect signals were characterized by: time T_1 out of range, time T_2 out of range, and time T_1 and T_2 out of the range allowed by standards.

Data pre-processing

The input dataset consisted of 269 measurements of voltage surges where each signal lasted by 200 μs and 50848 measurement points were registered. In order to enable effective models training data obtained in procedure described in previous section where subjected to the preprocessing procedure, consisting of the following steps:

- First, signals were truncated after 40000 points, since further value of the signal was close to zero and negligible.
- In the second step the data were resampled from 40000 points to 1000 in order to mitigate the risk of over-fitting the model by training it with limited number of samples with high resolution.
- Finally, signals where subjected to the z-score normalization, to obtain the mean of a signal equal to 0 and a standard deviation of 1. Time series from both classes (correct one and not) were hard to distinguished visually. Performed transformation improved the differentiation between two classes of signals. Examples of surges after transformation are depicted in Fig. 4.

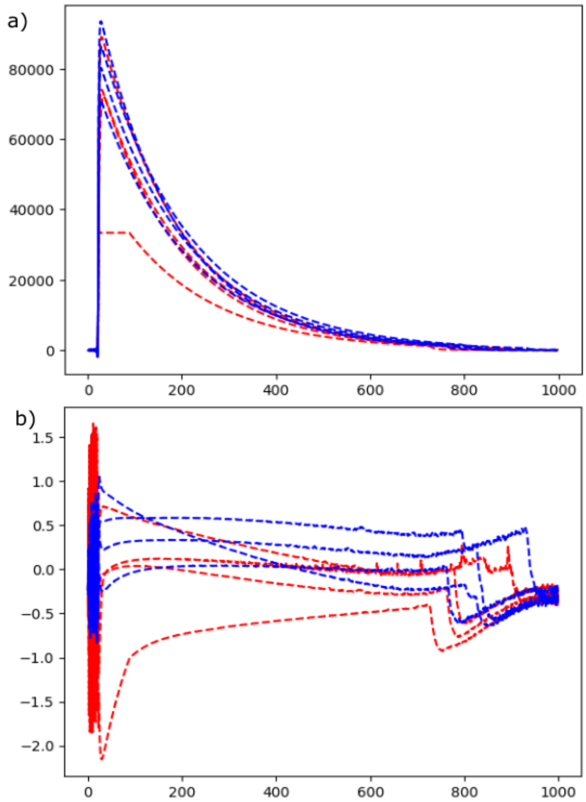

Fig. 4. Example time series a) after resampling to 1000 points b) after z-score normalization. Blue colour - samples in agreement with the standard, red colour - samples of non-standardized surges.

Prior to the training, data were randomly splited into training and testing sets with 80:20 ratio.

Models

We investigated 3 types of models often used in time-series processing tasks [11, 12, 13]: convolutional model, further referred as CNN, model combining convolutional layers with Long Short-term Memory Layers (CNN+LSTM) and transformer based model. Due to small number of training data, in order to avoid over-fitting we constructed models with only few layers.

CNN model consists of 1 up to 3 one dimensional convolutional layers followed by Global Average Pooling layer and one Dense layer. Output layer uses softmax activation function.

CNN+LSTM model includes initial one-dimensional convolutional layer, followed by maxpooling layer and dropout layer added, to increase generalization followed by bidirectional LSTM layer, dense layer with dropout and output dense layer with softmax activation function.

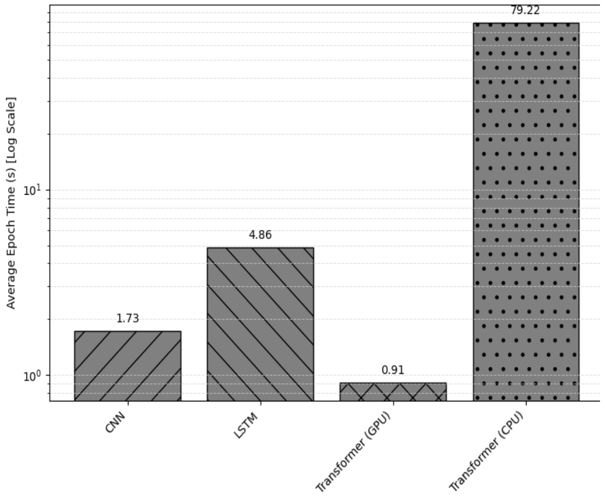

Fig. 5. Average training time of one epoch per model. Times were averaged with exclusion of the initial epoch which usually lasts longer.

Transformer model consists of: Input Layer which passes signal into the network, Token and Position Embedding layer which purpose is to embed signal into higher dimensional space as well as embed the position of the signal in the sequence to make the model aware of the data order. The following layer is Transformer layer which components are Multi-Head Attention block (Multiple heads allow to focus on different parts of sequence in parallel), Feed Forward Network followed by Normalization Layer and Dropout. In further step the output from the transformer block is flattened and passed to the Dense Layer with ReLU activation function. Finally, the output layer uses "sigmoid" function to obtain results.

Implementation

All models where implemented with the use of TensorFlow Keras [8, 9] framework and trained in Google Colab environment. The authors' intention was to create models that could be trained and executed using freely available resources. This goal was successfully achieved: out of 3 models only Transformer one required training on Graphical Processing Unit in order to make it efficient. For that reason we used GPU T4 model with 16GB of memory, other models were trained on standard CPU environment. Average training times are presented in Fig 5.

The Transformer model deployed on the GPU had the fastest training time, averaging below 1 second per epoch. In contrast, training the same model on the CPU was approximately 87 times slower, making it inefficient for multiple experiments. The average training times per epoch for the CNN and LSTM+CNN models were 1.73 and 4.86 seconds, respectively, and were therefore considered acceptable.

Training parameters

In order to obtain most accurate results we performed hyper-parameter tuning for all of three models. In each case we have chosen random parameter search method with 20 trials. The optimization goal was to maximize validation accuracy value (validation set consisted of randomly chosen 20% of training data). Models were trained for maximum 42 epochs with early stopping if validation loss did not decrease for 10 consecutive epochs. The batch size was fixed to the default value of 32 and as optimizer "Adam" was chosen in all cases. The parameters optimized for each models and their best combinations are summarized in Table 1.

Table 1. Models hyper-parameters

Hyperparameters	Range	Best value
CNN		
No. of conv. layers	< 1, 3 >	2
filters no. (conv. layer 1)	< 32, 128 >	64
Kernel size (conv. layer 1)	[3, 5, 7]	7
filters no. (conv. layer 2)	< 32, 128 >	128
Kernel size (conv. layer 2)	[3, 5, 7]	3
Dense units number	< 32, 128 >	64
Learning rate	< $1e^{-4}$, $1e^{-2}$ >	0.00636
LSTM+CNN		
filters no. (conv. layer 1)	< 32, 128 >	80
Kernel size (conv. layer 1)	< 2, 5 >	5
Pool size	< 2, 4 >	3
Dropout rate (CNN)	< 0.1, 0.5 >	0,4
LSTM units number	< 32, 128>	96
Dense units number	< 64, 256 >	160
Dropout rate (Dense)	< 0.1, 0.5>	0.1
Learning rate	<$1e^{-4}$, $1e^{-2}$>	0.00168
Transformer		
Dropout rate (CNN)	< 16, 64 >	48
LSTM units number	< 2, 8 >	8
Dense units number	< 32, 128 >	32
Dropout rate (Dense)	< 0.1, 0.5>	0.3
Learning rate	<$1e^{-4}$, $1e^{-2}$>	0.00021

Table 2. Models Metrics for Test Dataset

Model	**Class**[a]	**Acc.**	**Prec.**	**Recall**	**F1**	**Support**
CNN	0	0.69	0.72	0.74	0.74	31
	1		0.64	0.61	0.61	23
CNN+LSTM	0	0.81	0.78	0.94	0.85	31
	1		0.88	0.65	0.75	23
Transformer	0	0.93	0.94	0.94	0.94	31
	1		0.91	0.91	0.91	23

[a] 0 – incorrect shape, 1 – correct shape, Acc. – accuracy, Prec. - precision

Accuracy obtained on validation set for models with best hyper-parameters was equal to 73% for CNN model, 84% for combined CNN+LSTM model and 97% for Transformer model.

After hyper-parameters tuning step, models where retrained with 5 best parameters combinations for each model type. Best models where evaluated on separate test data set. The evaluation results are presented in Table 2.

Results discussion

The best results for both training and test data were obtained by Transformer model with F1 score equal to 94% for incorrect signal shape and 91% for correct one. The second best result with F1 score equal to 85% for "incorrect" class and 75% for "correct" one was obtained by CNN+LSTM model. However, it is worth noticing that whereas in Transformer and CNN models the values of Precision and Recall are well balanced and does not diverge from one another for CNN+LSTM model the difference between both measures is significant and model tends to classify more samples as incorrect ones. Such behaviour might be caused by imbalanced nature of the test set (higher amount of incorrect samples).

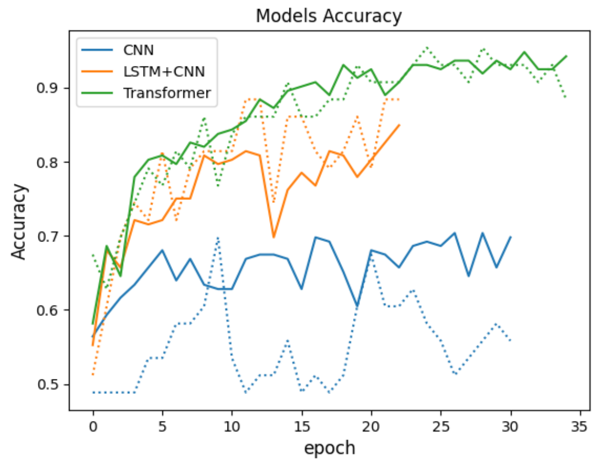

Fig. 6. Accuracy value of best models of each type tracked during training; Solid line ("-") -- training data, dotted line ("..") -- validation data.

The results obtained for CNN model are worse than the ones produced by Transformer and CNN models. The reason of this might be simpler architecture of CNN model without the use of layers which keep information about time dependencies in signal such as LSTM layer or attention structure. Lack of Dropout layer in aforementioned model might decrease generalization abilities.

Figures 6 and 7 show accuracy and loss traced during training of best models from each of three groups (CNN, LSTM+CNN, Transformer). We can observe that none of the models reached the limit of training epochs which was set to 42. (Which means that validation loss was not decreasing for next 10 epochs). Transformer model was trained for the highest number of epochs.

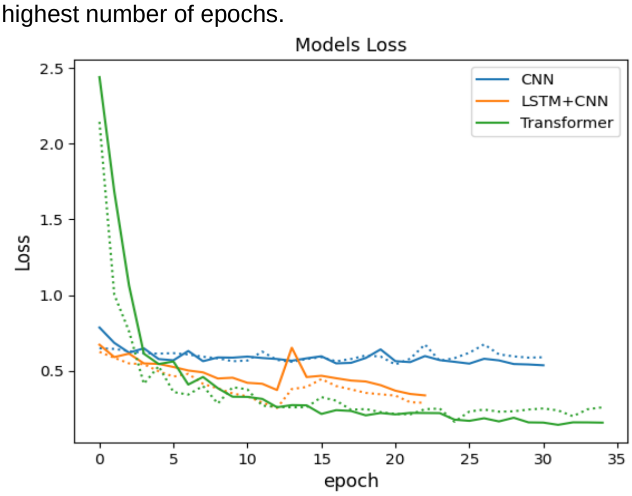

Fig. 7. Loss value of best models of each type tracked during training; Solid line ("-") -- training data, dotted line ("..") -- validation data.

During parameters tuning phase, authors focused on optimization of parameter values in fixed architecture (arbitrary written number and type of layers). Interesting direction of research might be also architecture optimization. Nevertheless, the results obtained by models, in particular Transformer one, with accuracy equals to 93%, show potential in using neural networks for voltage surges classification.

Ease of Use

In the future, the authors plan to create a universal program for classifying various lightning surges and expand the database with additional samples. Part of the data will be created artificially in programs for simulating electrical circuits, such as Matlab Simulink or ETAP.

Summary

The authors attempted to classify real lightning voltage surges. This paper presents three different algorithms used in the research as well as the classification results. The presented results allow to determine the possibility of classifying voltage surges. Despite the difficulty in distinguishing voltage surges meeting the requirements of the standard range, the Transformer model achieved satisfactory classification results.

The number of layers in all models was intentionally set to very small due to the small amount of data and hyper-parameters optimization was performed to avoid over-fitting. In the next step, the authors plan to significantly increase the amount of training data both through measurements collected on different measurement systems and data generated artificially through circuit simulations.

***Authors**: Ph.D. Zuzanna Krawczyk-Borysiak, M. Sc. Paweł Kluge, Ph.D. Andrzej Łasica, Ph.D. Przemysław Sul, M. Sc. Maciej Ciuba, Institute of Theory of Electrical Engineering, Measurement and Information Systems, Faculty of Electrical Engineering, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warszawa, Poland, email: pawel.kluge@pw.edu.pl, zuzanna.krawczyk@pw.edu.pl;*

REFERENCES

[1] Flisowski Z.: Technika wysokich napięć, WNT, Warsaw, 1992.
[2] IEC 60060-1. High-Voltage Test Techniques. Part 1: General Definitions and Test Requirements.
[3] IEC 60060-1. High-Voltage Test Techniques. Part 2: Measurements Systems.
[4] Tong L., Liu Y., Chen Y., Su S. and Liang P.: A CVT Based Lightning Impulse Wave Measuring Method Using Convolutional Neural Network, 2021 IEEE 4th International Electrical and Energy Conference (CIEEC), China, pp. 1-6, 2021.
[5] Figoń P.: Analiza numeryczna przebiegów udarowych — algorytmy obliczeniowe, Przegląd Elektrotechniczny, China, no.10, pp. 201-205, 2015.
[6] Yutthagowith P.: Non-Iterative Technique for Determination of Full Lightning Impulse Voltage Parameters, Energies, 15(12):4199, 2022.
[7] Tuethong P, Kitwattana K, Yutthagowith P, Kunakorn A.: An Algorithm for Circuit Parameter Identification in Lightning Impulse Voltage Generation for Low-Inductance Loads, Energies, 13(15):3913, 2010.
[8] Keras: Deep Learning for Humans [web page] http://keras.io/. [Accessed on 20 May 2024.].
[9] TensorFlow: [web page] https://www.tensorflow.org/. [Accessed on 20 May 2024.].
[10] Hauschild W., Lemke E.: High-Voltage Test and Measuring Techniques, Springer, 2014.
[11] B. Zhao, H. Lu, S. Chen, J. Liu and D. Wu: Convolutional neural networks for time series classification, Journal of Systems Engineering and Electronics, vol. 28, no. 1, pp. 162-169, 2017.
[12] Pham T.D.: Time–frequency time–space LSTM for robust classification of physiological signals, Sci Rep 11, 6936, 2021.
[13] Uchiyama T.: Transformer-Based Time Series Classification for the OpenPack Challenge 2022, 2023 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), Atlanta, GA, USA, pp. 264-266 2023.