JAISCR, 2023, Vol. 13, No. 3, pp. 197 – 210
10.2478/jaiscr-2023-0015

AN INTELLIGENT APPROACH TO SHORT-TERM WIND POWER PREDICTION USING DEEP NEURAL NETWORKS

Tacjana Niksa-Rynkiewicz[1], Piotr Stomma[2], Anna Witkowska[3], Danuta Rutkowska[4], Adam Słowik[5], Krzysztof Cpałka[6], Joanna Jaworek-Korjakowska[7], Piotr Kolendo[8]

[1]*Gdańsk University of Technology, Faculty of Ocean Engineering and Ship Technology, 80-233 Gdańsk, Poland*

[2]*University of Białystok, Institute of Computer Science, 15-328 Białystok, Poland*

[3]*Gdańsk University of Technology, Faculty of Electrical and Control Engineering, 80-233 Gdańsk, Poland*

[4]*University of Social Sciences, Information Technology Institute, 90-213 Łódź, Poland*

[5]*Koszalin University of Technology, Department of Electronics and Computer Science, 75-452 Koszalin, Poland*

[6]*Częstochowa University of Technology, Department of Intelligent Computer Systems, 42-200 Częstochowa, Poland*

[7]*AGH University, Department of Automatic Control and Robotics, Center of Excellence in Artificial Intelligence 30-059 Kraków, Poland*

[8]*Institute of Power Engineering, Department of Power Automation, 80-870 Gdańsk, Poland*

[*]*E-mail: krzysztof.cpalka@pcz.pl*

Submitted: 26th May 2023; Accepted: 27th May 2023

Abstract

In this paper, an intelligent approach to the Short-Term Wind Power Prediction (STWPP) problem is considered, with the use of various types of Deep Neural Networks (DNNs). The impact of the prediction time horizon length on accuracy, and the influence of temperature on prediction effectiveness have been analyzed. Three types of DNNs have been implemented and tested, including: CNN (Convolutional Neural Networks), GRU (Gated Recurrent Unit), and H-MLP (Hierarchical Multilayer Perceptron). The DNN architectures are part of the Deep Learning Prediction (DLP) framework that is applied in the Deep Learning Power Prediction System (DLPPS). The system is trained based on data that comes from a real wind farm. This is significant because the prediction results strongly depend on weather conditions in specific locations. The results obtained from the proposed system, for the real data, are presented and compared. The best result has been achieved for the GRU network. The key advantage of the system is a high effectiveness prediction using a minimal subset of parameters. The prediction of wind power in wind farms is very important as wind power capacity has shown a rapid increase, and has become a promising source of renewable energies.

Keywords: Renewable Energy, Wind Energy, Wind Power, Wind Turbine, Short-Term Wind Power Prediction, Deep Learning, Convolutional Neural Networks, Gated Recurrent Unit, Hierarchical Multilayer Perceptron, Deep Neural Networks

1 Introduction

Renewable energy sources are an alternative to coal-fired power plants. Their role in the production of electrical energy is systematically increasing, and many countries are striving to gradually phase out coal-fired power plants. These actions are in line with the goals of important strategies to combat climate change and improve air quality, which include the "European Green Deal" for the European Union, the "Clean Air Act" for the United States, etc.

Wind power plants play an important role among renewable energy sources. They use the power of wind, which is a common phenomenon and allows for the delivery of large amounts of electrical energy. Wind turbines can be installed on both land and sea.

They generate low operating costs, and can work under different weather conditions. Moreover, they can easily be activated depending on the demand for electrical energy; this process can be automated. In addition, they ensure an increase in energy independence, create new jobs in the local environment, do not emit CO_2, and have a minimal impact on the environment – they stand out due to their low greenhouse gas emissions.

An important issue for wind power plants is Short-Term Wind Power Prediction (STWPP). Such a forecast covers the next few hours (day-ahead). It must take into account the unpredictability of the weather, particularly the variability of the wind.

Wind power prediction helps to reduce the effects of instability of atmospheric conditions, allows planning of power plant operations, provides the possibility of a more stable operation of the energy system, ensures optimization of fossil fuel consumption, labour, and maintenance costs, etc.

The STWPP depends on many factors including wind speed, turbulence (which can significantly affect efficiency), air temperature, atmospheric pressure, season, topographical factors of the terrain, current demand for electrical energy (which can change rapidly), insolation, precipitation level, etc. Precise identification of key factors that determine effective short-term prediction is difficult because many of them are interdependent and their influence can be difficult to quantify.

In this paper, the STWPP problem is considered with application of Deep Neural Networks (DNNs). The impact of prediction time horizon length on the accuracy, and the impact of temperature on prediction effectiveness are analyzed. Simulation studies based on real data from a working wind turbine are described, and results are presented.

The paper is organized as follows: Section 2 provides information about works related to the subject of the STWPP published by various authors. In Section 3, the main characteristics of our approach are described. Section 4 outlines selected aspects of wind turbine operation. A detailed description of the proposed approach is given in Section 5. Simulations, including assumptions and results achieved by the implemented solution, are presented in Section 6. Conclusions are formulated, and plans for future works are outlined, in Section 7.

2 Related works

In this article, DNNs are used to solve the STWPP problem. This kind of neural networks are steadily gaining popularity, so their capabilities are being utilized in various application areas. For example, such networks can be used to identify non-human traffic on a website [11]. Among others, a Recurrent Neural Network (RNN) is employed to monitor a regenerative heat exchanger of a steam turbine power plant [30]. The problem of processing CAPTCHA codes (used on websites) solved by a Convolutional Neural Network (CNN) is considered in [34]. Stacked ensembles of neural networks and autoencoders are tested for intrusion detection in IT services [3]. The authors of paper [18] focus their attention on the problem of improving the efficiency of medical imaging in the case of limited data availability, taking into account interpretation capabilities of the model. In [32], a DNN is used to detect people and human posture points in 2D images. In [28], a CNN is applied to epileptic seizure recognition.

For the STWPP, different methods are employed, not only neural networks, but also statistical models based on regression, various artificial intelligence methods, and hybrid models. In [12], the state-of-the-art in STWPP is presented with a literature overview.

In [49], parametric power curve models, such as the four-parameter logistic model, five-parameter logistic model, and polynomial regression, are analyzed. In [16], a piece-wise linear model is proposed to estimate a power curve based on data of wind speed. In [45], statistical methods are considered with a one-dimensional polynomial dependent on wind speed and a two-variable function using wind speed and direction. In [7, 8], attempts are made to predict wind speed and active power using models such as moving average, weighted moving average, autoregressive moving average, and autoregressive integrated moving average. In [48], it is shown that improving the accuracy of weather forecasting translates into better accuracy of turbine power prediction.

A method for predicting wind speed and direction by use of a one-dimensional CNN is proposed in [14]. The authors of paper [27] focus their attention on using information about wind speed and precipitation, applying methods of time series mapping to image matrices and feature extraction, and employing a DNN for power prediction. In [54], an innovative approach to short-term prediction, based on a set of artificial neural networks, with the PCA (Principal Component Analysis) and FCM (Fuzzy C-Means) clustering algorithm, is presented. In [46], the ELM (Extreme Learning Machines) and a cloud model that takes into account uncertainty aspects are used for power prediction.

In [24], hybrid approaches that combine machine learning and meteorological data-based prediction are described. The authors of article [51] compare the effectiveness of the LSTM (Long Short-Term Memory), CNN, and DBN (Deep Believe Network) in short-term wind speed forecasting. In paper [53], an approach based on the wavelet transform, ELM, and firefly algorithm (population-based optimization), is proposed. In [33], various machine learning techniques, including neural networks, genetic algorithms, and reinforcement learning, are considered in application to the wind power prediction. In [4], the importance of proper selection of input variables, modeling techniques, and methods for evaluating prediction quality, are discussed.

In addition, hybrid methods, particularly those based on population algorithms, are applied to the STWPP. Such algorithms are widely used in various design tasks. In [39], the application of evolutionary algorithms for designing digital minimum-phase filters with non-standard amplitude characteristics and a finite word length is described. In [40] an evolutionary method for designing and optimizing digital combinatorial circuits is employed. In [43], an evolutionary division of VLSI circuits into sub-circuits with a minimal number of connections is presented. In [41, 42], evolutionary methods for designing and optimizing digital IIR filters with non-standard characteristics are proposed.

As for the applications of population-based algorithms to solve the STWPP problem, a hybrid approach that combines the Least Squares Support Vector Regression (LSSVR) and Artificial Bee Colony (ABC) is described in [5]. A hybrid method for wind speed and wind turbine power prediction using fuzzy clustering analysis and the SVM (Support Vector Machines) is presented in [55]. A hybrid approach that combines an MLP (Multi-Layer Perceptron), a genetic algorithm, the SVM, and the linear regression, is depicted in [38]. An innovative approach to the STWPP, based on the kNN (k-nearest neighbors) and PSO (Particle Swarm Optimization), is proposed in [25]. A hybrid approach to wind turbine prediction using a CNN and an improved DE (Differential Evolution) is described in [20]. An algorithm based on a deep CNN, and an evolutionary optimizer that uses the Grey Wolf Optimization (GWO is applied in order to predict the short-term wind power in [15].

Many other publications that propose various modifications of Machine Learning (ML) methods and hybrid approaches to the STWPP are available. The methods of wind power prediction can be analyzed with respect to three factors: physical, statistical, and ML; see [37]. The authors of paper [26] provide an overview of AI-based hybrid approaches for wind power forecasting, and consider Artificial Intelligence (AI) methods with regard to the aforementioned categories; see also [47].

3 Main characteristics of the proposed approach

Taking into account the solutions presented by different authors, and described in Section 2 with regard to the STWPP problem, we propose our approach focusing attention on the DNNs.

Specifically, we are interested in investigating the following aspects: (1) whether it is possible to effectively solve the STWPP problem using a significantly reduced set of input attributes for the DNNs; (2) whether the selection of the type of DNNs is crucial considering the specificity of the STWPP problem; (3) how the prediction time horizon determines the accuracy of the STWPP problem.

Our motivation for considering these issues comes from: (a) the application of various types of DNNs; (b) the use of an extensive set of attributes that describe turbine operating conditions; (c) the incorporation of additional data, such as information from weather services.

In this paper, we apply three types of DNNs (i.e. CNN, Gated Recurrent Unit, and Hierarchical Model based on MLP) to the short-term prediction of wind turbine power. In addition, different prediction horizons are considered. Moreover, focusing our attention on the most reduced set of attributes used for prediction, with a particular emphasis on evaluating the impact of air temperature, can be viewed as an important issue in this approach.

It is worth emphasizing that the simulation studies presented in this article were conducted by use of real data from the Alstom ECO110 turbine with a nominal power of 3 MW operating in a 90 MW farm located in the Pomeranian Voivodeship in Poland. This is very important because prediction results obtained by machine learning models strongly depend on the data. This means that artificial intelligence systems trained on the data gathered in one place (in this case – a geographical location) may not work very well in another location (with different weather conditions).

Therefore, in spite of the fact that there are many publications presenting applications of various methods, including deep learning, to the STWPP problem (see Section 2), most of them employ data collected from wind farms located in countries (and continents) where weather conditions are totally different than in the north part of Poland. Thus, such solutions cannot be directly implemented in wind power plans developed in our country.

The prediction time horizons also should be taken into account because a good system applied in real time to the STWPP problem should guarantee reliable results in the case when the time horizon is extended. Generally, the prediction accuracy decreases as the time horizon increases.

4 Description of selected aspects of wind turbine operation

Power output P of a wind turbine in a wind farm depends on many factors, such as wind speed V, wind direction dir ($\vec{V} = [V, dir]$), blade surface area A, turbine efficiency, air temperature T (an increase in temperature may reduce air density ρ, which increases power output), and atmospheric pressure (a decrease in pressure may also reduce air density, increasing power output). Taking these components into account, the power output of the turbine can be described as:

$$P = f\left(\rho, T, \vec{V}, C_p, A\right), \tag{1}$$

where $f(\cdot)$ is the adopted dynamic model of the turbine, and C_p is a constant power coefficient dependent on the specific turbine (including its efficiency). The correlation plot between power output P and dir, V, and T, generated by the turbine for the data used in the simulations is shown in Figure 1. Based on this plot, it can be concluded that not all factors have a clear impact on P. One such factor is, for example, T.

Assuming that wind speed V is the key factor for power output P of a wind turbine, instantaneous power P (in W) can be estimated as follows:

$$P = \begin{cases} 0 \text{ if } V < V_{CutIn} \\ 0.5 \cdot C_p \cdot \rho \cdot A \cdot V^3 \text{ if } V_{CutIn} \leq V < V_r \\ P_r \text{ if } V_r \leq V < V_{CutOut} \\ 0 \text{ otherwise,} \end{cases} \quad (2)$$

where V_{CutIn} is the minimum wind speed required to start the turbine, V_r is the wind speed at which the turbine achieves maximum efficiency, V_{CutOut} is the maximum wind speed at which the turbine must be shut down, C_p is the aerodynamic power coefficient (the ratio of actual wind power to maximum theoretical wind power), and P_r is the rated power of the turbine (i.e. achieved under nominal operating conditions). The variable values used in equation (2) for the turbine considered in the simulations are shown in Table 1.

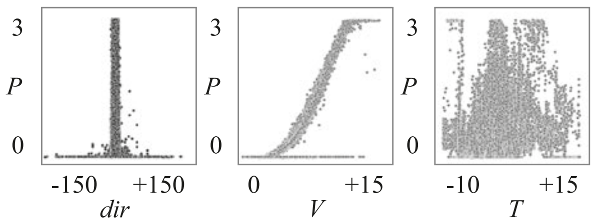

Figure 1. Graphical representation of the correlation between power $P[MW]$ generated by the turbine and $dir\,[^\circ]$, $V\,\left[\frac{m}{s}\right]$, and $T\,[^\circ C]$, occurring in the data used in simulations.

The problem of the STWPP consists of using the attribute values considered in the context of equation (1). Not only the current values of these attributes are taken into account, but also the values in previous moments $(t-1),(t-2),..,(t-past)$. All of them are used to determine the maximum power values in the subsequent moments $(t+1),(t+2),..,(t+horizon)$. Therefore, it is a typical regression problem that can be solved by use of ML methods. In practice, this problem is not easy to solve because:

– Observation time Δt, and the time resolution of the forecast δ, must be properly defined. In the simulations we assumed that $\Delta t \in \{7.5\,\text{min}, 22.5\,\text{min}, 45.0\,\text{min}\}$) and $\delta = 150\,\text{sec}$).

– A different subset of available input attributes can be applied to predict wind power. Some of them may be dependent on each other.

– It is a dynamic problem, and a different number of historical steps can be used for each of the selected input attributes to solve it.

5 Description of the Deep Learning Power Prediction System

The following subsections illustrate architectures of the system created by means of the DNNs, in accordance with the description of wind turbine operation presented in Section 4 and characteristics of our approach outlined in Section 3.

5.1 Power Transmission System

The STWPP should be considered with regard to the Power Transmission System (PTS) shown in Figure 2. The PTS includes the infrastructure that enables the transmission of electrical energy from wind turbines to consumers. It consists of a turbine generating electrical energy (turbines comprising a wind farm), transformers to increase the voltage for reducing transmission power losses, transmission lines, a centralized SCADA (Supervisory Control and Data Acquisition) system for remote monitoring and control of turbine operations, a power prediction system, and an Independent System Operator (ISO) coordinating the operation of the wind farm with other sources of electrical energy in the network. The proposed approach to the STWPP can be a component of the PTS. The SCADA system collects data, such as temperature, wind speed, wind direction, etc. These data can be used in order to train systems based on ML methods. For more information concerning the SCADA system, see [37].

5.2 Deep Learning Prediction System

In the context of this article, the most important block in the PTS is the Deep Learning Power Prediction System (DLPPS), shown in Figure 3. The DLPPS performs initial preprocessing of real data to adapt it for prediction. In particular, data preprocessing includes: equalizing time intervals between data, removing missing data, and determining the

$[V, dir]$ components of the velocity vector $\{V_x, V_y\}$. Simple linear interpolation can be applied to determine missing power values:

$$P(t) = P(t_s) + (P(t_e) - P(t_s)) \cdot \frac{t - t_s}{t_e - t_s}, \quad (3)$$

where $t \in (t_s, t_e)$ and $P(t) \in (P(t_e), P(t_s))$ must be known.

5.3 Deep Learning Prediction Block

Preprocessed data prepared in the DLPPS block can be passed to the DLP (Deep Learning Prediction) block which contains a system that properly defines model $f(\cdot)$ in the form of (1). In this paper, we assume that such a model is expressed by a DNN. During the training procedure, preprocessed measurement data in the following form are used:

$$\{\mathbf{X}(t), \mathbf{Y}(t)\} = \left\{ \begin{array}{c} \left\{ \begin{array}{c} T(t), T(t-1), ..., T(t-past), \\ V_x(t), V_x(t-1), ..., V_x(t-past), \\ V_y(t), V_y(t-1), ..., V_y(t-past), \\ dir(t), dir(t-1), ..., dir(t-past), \\ ... \\ P(t), P(t-1), ..., P(t-past) \end{array} \right\}, \\ \left\{ \begin{array}{c} P(t+1), P(t+2), ..., \\ P(t+horizon) \end{array} \right\} \end{array} \right\}. \quad (4)$$

In our approach, three types of DNNs are employed to construct the DLP block:

- CNN [21] – most popular kind of DNNs. These neural networks apply the convolution operation to process input data, usually images or other spatial data. CNNs consist of convolutional layers that employ filters (kernels) to extract features from input data, and pooling layers that reduce the size of output data from the convolutional layers. Fully connected layers are located at the end of CNNs, and use the resulting features for classification or regression. The architecture of the CNN is illustrated in Figure 4.a.
- GRU (Gated Recurrent Unit) [6] – The GRU is a type of a RNN that was proposed as an extension of the standard LSTM model. This neural network works in a similar way to the LSTM, but it has fewer parameters, and employs reset and update gates to control the information transmitted by the network. Thanks to these gates, the GRU can deal with the problem of the vanishing gradient, which occurs in standard RNNs. Because the GRU works faster than the LSTM, it can be applied in order to process sequences of different lengths. The architecture of the GRU is portrayed in Figure 4.b.
- H-MLP (Hierarchical MLP) [31] – The hierarchical model based on MLP neural networks – a type of DNNs that consists of multiple layers of interconnected MLPs. This model is hierarchical which means that higher layers learn increasingly complex features, which are built on the basis of the features learned in lower layers. This approach allows for a higher level of abstraction in the representation of data. The architecture of the H-MLP is shown in Figure 4.c.

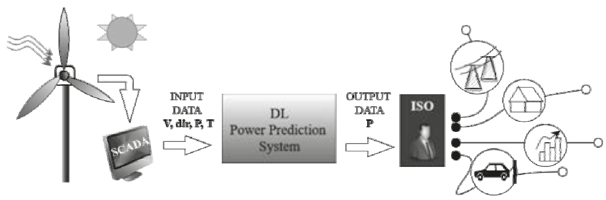

Figure 2. Architecture of the Power Transmission System (PTS).

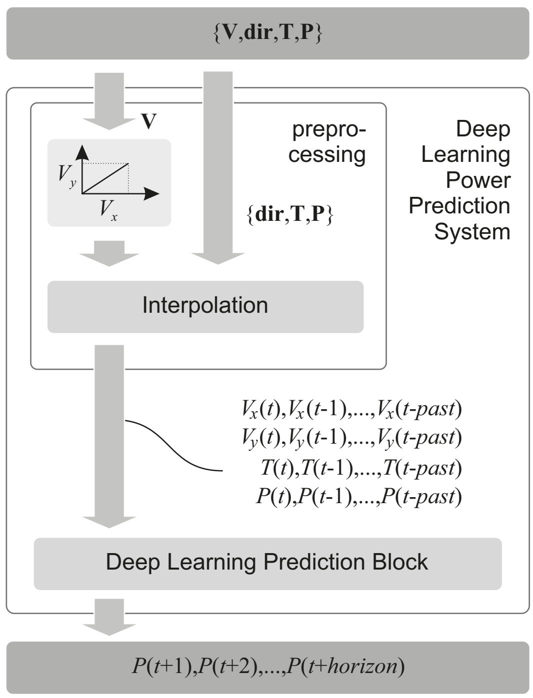

Figure 3. Deep Learning Power Prediction System (DLPPS).

There is a large number of publications covering subjects related to DNNs and their applications, especially with regard to CNNs, including review papers, see eg. [1, 9], in addition [50] concerning the GRU. However, not so many articles refer to the H-MLP; one of the examples presents an approach to automatic language identification based on hierarchical MLP classifiers [23].

The three types of DNNs, depicted above and displayed in Figure 4, have been applied in the simulations reported in Section 6. Nevertheless, it should be emphasized that different models can also be employed – instead of the suggested DNNs variants, as the implementation of the DLP block. For instance, a fuzzy system [44, 52] trained using a population-based algorithm [10, 36] or another option like a neuro-fuzzy architecture with hybrid learning [35], can be considered in this application.

6 Simulations

The simulations conducted by the system described in Section 5, on a real dataset from the SCADA, have been reported in this section which in the following subsections presents assumptions regarding the simulation, and results, respectively.

6.1 Simulation assumptions

As mentioned in Section 4, values of the variables used in equation (2), for the wind turbine considered in the simulations, are shown in Table 1.

Table 1. Technical specification of the wind turbine considered in the simulations.

No.	Parameters of the wind turbine	Values of the parameters
1.	rated power P_r	3.0 MW
2.	type of generator	asynchronous DFIG
3.	rotor diameter	110 m
4.	cut-in wind speed V_{CutIn}	3 $\frac{m}{s}$
5.	rated wind speed V_r	11.5 $\frac{m}{s}$
6.	cut-out wind speed (avg. 10 min) - V_{CutOut}	25 $\frac{m}{s}$
7.	instant cut-out wind speed (3s)	34 $\frac{m}{s}$

Remarks regarding the simulation can be summarized as follows:

- The paper considers data from a wind farm consisting of 30 Alstom ECO110 turbines. One ECO110 turbine has a nominal power of 3.150 MW, produces approximately 195 MWh annually, is designed for medium and high wind speeds, and uses a doubly fed induction generator (DFIG). More information on the DFIG can be found in [29].
- Three variants of DNNs in the DLP block were tested in the simulations: CNN, GRU, H-MLP, in accordance with the information provided in Section 5.3. The hyperparameters of these networks are presented in Table 2.
- For the purpose of the tests performed in this work, 33,983 real-time readings of values from the SCADA system were archived: P, V, dir, and T, during turbine operation at one of the Polish wind farms in the Pomeranian Voivodeship in 2021. The readings were taken every 150 seconds (2.5 minutes). It was assumed that high time resolution could have a positive impact on the performance of the DLP block.
- The training data for the DLP block were generated assuming that the wind farm produces maximum active power under current weather conditions, which is not limited by the farm control system (except for the cases endangering turbine safety, which did not occur during the data acquisition stage).
- The training data for the DLP block include: wind direction dir and speed V, ambient temperature T, and turbine power P. The structure of a segment of these data (limited to 4 decimal places) is shown in Table 3.
- The training dataset, in the form of the structure shown in Table 3, was processed according to the information provided in Section 5. The following number of observations was taken: $horizon \in \{3, 9, 18\}$. The prediction time for the wind turbine power was derived from the number of observations and the time interval between successive readings (i.e. 2.5 minutes): $\Delta t \in \{7.5\,\text{min}, 22.5\,\text{min}, 45.0\,\text{min}\}$. An exemplary segment of the training sequence for $past\_history = 3$ is illustrated in Table 4.

– The simulations were conducted in two variants: with the temperature attribute T and without taking into account the temperature. The purpose of this approach was to attempt to answer the question of whether T, as a derivative of the current weather state, has a significant influence on the value of power generated by the turbine. Doubts regarding the need to consider the temperature attribute were raised in Section 4 with regard to comments concerning Figure 1.

– The Mean Squared Error (MSE) was applied in order to evaluate the effectiveness of prediction while the DNNs were employed. In the case of networks with m neurons in the output layer, the MSE errors were determined as the mean errors of these neurons. The training data were split into a 70:30 ratio into a training and test sequence. Each of the simulations was repeated 50 times, and the corresponding results were averaged.

– The stopping criterion for the learning algorithm, for each of the three variants of DNNs in the DLP block (CNN, GRU, H-MLP), was a constant value of the MSE error over the next 40 learning steps. The maximum number of learning epochs was 100.

– The adaptive moment estimation (ADAM) gradient algorithm, which combines the advantages of the RMSprop (Root Mean Square propagation) and Momentum [19], was used to train each of the DNNs considered in the simulations. Each network was tested for the following variations: $past\_history \in \{3, 6, 9\}$ for $horizon = 3$, $past\_history \in \{11, 18, 27\}$ for $horizon = 9$, and $past\_history \in \{22, 36, 54\}$ for $horizon = 18$. A comparison of the results for the best variants is presented in Table 5.

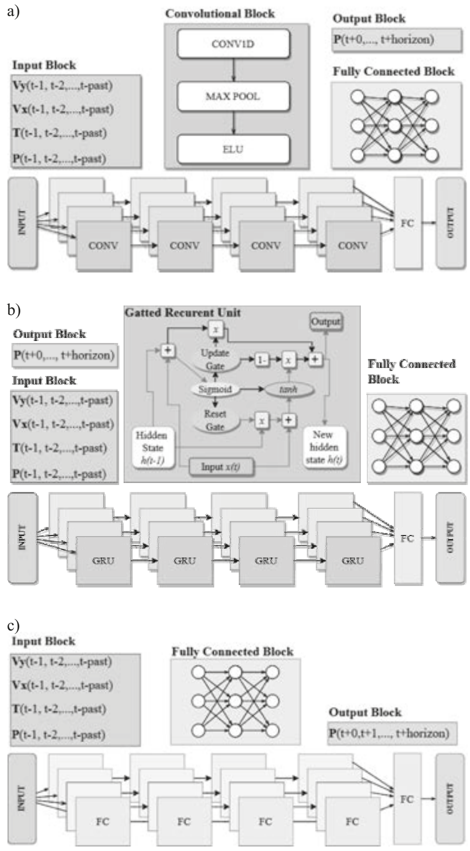

Figure 4. Architectures of the DNNs applied in the DLP block: a) CNN, b) GRU, and c) H-MLP.

Table 2. Hyperpameters of the DNN architecture applied in the DLP block.

No.	Parameter name	Parameter value	Network symbol
1.	batch size	32	CNN
2.	learning rate	0.001	CNN
3.	normalization	z-score	CNN
4.	pooling	max. size 2	CNN
5.	activation function	elu	CNN
6.	kernel size	$[3, 3, 3]$	CNN
7.	filters	$[10, 10, 20]$	CNN
8.	layers	3	CNN
9.	num. of neurons	64-64	GRU
10.	learning rate	0.001	GRU
11.	normalization	z-score	GRU
12.	layers	2	GRU
13.	batch size	32	GRU
14.	activation function	relu	H-MLP
15.	learning rate	0.001	H-MLP
16.	normalization	z-score	H-MLP
17.	layers	5	H-MLP
18.	batch size	32	H-MLP

Table 4. An example of the training sequence for DNNs applied, for $past\_history = 3$.

$V_x(t-1)$	$V_x(t-2)$	$V_x(t-3)$	$P(t-1)$	$P(t-2)$	$P(t-3)$	$V_y(t-1)$	$V_y(t-2)$	$V_y(t-3)$	$P(t)$	$P(t+1)$	$P(t+2)$
...	...	...	...	...	...	...	...	...	...	...	...
-0.4866	-0.5961	1.3565	-0.0320	-0.2251	0.0006	-1.3706	1.1879	0.3487	-0.2674	-0.4575	-0.4199
-0.5961	1.3565	1.1526	-0.2251	0.0006	-0.2673	1.1879	0.3487	0.3004	-0.4575	-0.4199	-0.5103
1.3565	1.1526	0.9900	0.0006	-0.2673	-0.4574	0.3487	0.3004	0.0380	-0.4199	-0.5103	-0.6814
1.1526	0.9900	1.0966	-0.2673	-0.4574	-0.4199	0.3004	0.0380	0.0402	-0.5103	-0.6814	-0.6287
0.9900	1.0966	0.1134	-0.4574	-0.4199	-0.5102	0.0380	0.0402	1.1036	-0.6814	-0.6287	-0.7334
...	...	...	...	...	...	...	...	...	...	...	...

Table 5. Best simulation variants for DNNs applied. The best results are printed in bold.

No.	DNNs	Input attributes	*horizon*/ *past_history*/ Δt [s]	MSE	*horizon*/ *past_history*/ Δt [s]	MSE	*horizon*/ *past_history*/ Δt [s]	MSE
1a.	CNN	$\{V_x, V_y, P\}$	3/3/7.5	0.0745	9/11/22.5	0.1319	18/36/45.0	0.1904
1b.	CNN	$\{V_x, V_y, P, T\}$	3/3/7.5	0.0754	9/11/22.5	0.1366	18/36/45.0	0.1998
2a.	GRU	$\{V_x, V_y, P\}$	3/6/7.5	**0.0727**	9/18/22.5	**0.1295**	18/36/45.0	**0.1817**
2b.	GRU	$\{V_x, V_y, P, T\}$	3/6/7.5	0.0745	9/18/22.5	0.1320	18/36/45.0	0.1875
3a.	H-MLP	$\{V_x, V_y, P\}$	3/6/7.5	0.0751	9/11/22.5	0.1341	18/54/45.0	0.1895
3b.	H-MLP	$\{V_x, V_y, P, T\}$	3/6/7.5	0.0742	9/11/22.5	0.1345	18/54/45.0	0.1986

Table 3. An examplary segment of the real data obtained from the SCADA system.

TimeStamp	*dir*	*V*	*T*	*P*
...	...	...	...	...
2021-01-01 00:00:01	4.7287	5.2717	1.5000	0.2104
2021-01-01 00:02:31	6.6561	5.3789	1.5000	0.2697
2021-01-01 00:05:01	6.6561	5.6060	1.5000	0.2477
2021-01-01 00:07:31	6.6561	5.2746	1.4571	0.2431
2021-01-01 00:10:01	7.3438	4.9269	1.4428	0.1965
...	...	...	...	...

6.2 Simulation results

Results of the simulations, with conclusions, can be summarized as follows:

– The DLP block worked with appropriate accuracy for each tested simulation variant. The best results were obtained for the GRU network; prediction time of 7.5 minutes, 6 historical time steps considered for input attributes, and ambient temperature T omitted (see Table 5).

– Each of the three DNNs performed similarly in both variants: when the ambient temperature attribute was considered and omitted (see rows 1a, 2a, 3a vs. 1b, 2b, 3b, in Table 5). This is also confirmed by the graphical representation of the data shown in Figure 1.

– As expected, the best results were obtained for the shortest tested prediction time (see the MSE column for $horizon = 3$, in Table 5).

– The choice of types of the DNNs in the DLP block does not significantly affect the accuracy of power prediction. This may be due to the specificity of the problem itself, where the dynamics of changes are not high.

6.3 Comparison of the results

A direct comparison of the results obtained from the simulation conducted for our approach to the STWPP problem, with the results of other authors, is not straightforward because: (a) different turbine models can be applied, (b) different prediction horizons are used, (c) different methods for pre-processing data are employed, (d) different datasets collected from wind farms in various locations are utilized.

However, it can be concluded that the results obtained in our work, and results of other authors [2, 13, 17, 22], are similar in terms of accuracy: the deviation of the predicted power from the actual power is not greater than a few percent.

The key advantage of the results presented in this paper demonstrates a high prediction effectiveness achieved using a minimal subset of input attributes.

7 Conclusions

In this paper, we consider an intelligent approach (AI, ML methods) to the Short-Term Wind Power Prediction (STWPP) problem, using Deep Neural Networks (DNNs). Three types of DNNs (the CNN, GRU, and H-MLP) have been employed to construct a wind turbine power prediction block.

Data from a wind turbine, operating on one of the wind farms in Poland (in the Pomeranian voivodeship), have been used to test the effectiveness of this approach. The obtained results are satisfactory.

The key conclusions from the conducted simulations can be summarized as follows. Firstly, the problem of the STWPP is important from a practical point of view, but it is not characterized by high dynamic changes. As a result, each of the applied DNN operated with satisfactory accuracy, so the choice of the network architecture was not crucial.

Secondly, in the short-term prediction, temperature could be removed from the input attributes of the neural networks, and its omission did not significantly affect the prediction accuracy. However, our recommendation - based on the simulations - is to use the GRU.

Thirdly, it was difficult to obtain a reliable turbine power forecast for time periods longer than several minutes. This is due to the potentially large variability of weather conditions characteristic of the coastal climate in winter, which the considered turbine had to cope with.

Our future plans include defining a time-varying model of prediction reliability dependent on weather data from a given location. The behavior of this model could, for example, automatically determine the length of the prediction time horizon.

Acknowledgment

The project is financed under the program of the Polish Minister of Science and Higher Education, the name "Regional Initiative of Excellence", in the years 2019-2023, project number 020/RID/2018/19, the amount of financing PLN 12,000,000.

References

[1] Alzubaidi L., Zhang J., Humaidi A.J. et al.(2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data 8(53), doi.org/10.1186/s40537-021-00444-8

[2] Benítez-Buelga A., Fernández-Blanco P., and Usaola J. (2019). Wind power short-term prediction using LSTM recurrent neural networks. Energies, 12(17), 3338.

[3] Brunner C., Kő A., and Fodor S. (2022). An autoencoder-enhanced stacking neural network model for increasing the performance of intrusion detection. J. of Artificial Intelligence and Soft Computing Research, 12(2) 149-163.

[4] Chaudhary A., Sharma A., Kumar A., Dikshit K., & Kumar N. (2020). Short term wind power forecasting using machine learning techniques. J. of Statistics and Management Systems, 23, 145-156.

[5] Cheng Y., Zhang Z., and Zhou Y. (2018). A population-based wind power short-term prediction approach using hybrid least squares support vector regression with artificial bee colony algorithm. Energy Procedia, 152, 697-703.

[6] Chung J., Gulcehre C., Cho K., and Bengio Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555.

[7] Colak I., Sagiroglu S., Yesilbudak M., Kabalci E., and Bulbul H.I. (2015). Multi-time series and-time scale modeling for wind speed and wind power forecasting part I: Statistical methods, very short-term and short-term applications. In 2015 Int. Conf. on Renewable Energy Research and Applications (ICRERA) (pp. 209-214). IEEE.

[8] Colak I., Sagiroglu S., Yesilbudak M., Kabalci E., and Bulbul H.I. (2015). Multi-time series and-time scale modeling for wind speed and wind power forecasting part II: Medium-term and long-term applications. In 2015 Int. Conf. on Renewable Energy Research and Applications (ICRERA) (pp. 215-220). IEEE.

[9] Emmert-Streib F., Yang Z., Feng H., Tripathi S., and Dehmer M. (2020) An introductory review of deep learning for prediction models with big data. Frontiers in Artificial Intelligence. Sec. Machine Learning and Artificial Intelligence, vol.3, doi.org/10.3389/frai.2020.00004

[10] Gabryel M., Cpałka K., and Rutkowski L. (2005). Evolutionary strategies for learning of neuro-fuzzy systems. Proc. of the I Workshop on Genetic Fuzzy Systems, 119-123.

[11] Gabryel M., Lada D., Filutowicz Z., Patora-Wysocka Z., Kisiel-Dorohinicki M., and Chen G. (2022). Detecting anomalies in advertising web traffic with the use of the variational autoencoder. J. of Artificial Intelligence and Soft Computing Research, 12 (4) 255-256.

[12] Giebel G., Brownsword R., Kariniotakis G., Denhard M., and Draxl C. (2011). The state-of-the-art in short-term prediction of wind power: A literature overview, 2nd edition. ANEMOS.plus.

[13] Guo Z., Song J., and Liu Y. (2018). Wind power short-term prediction based on convolutional neural network. Energies, 11(10), 2634.

[14] Harbola S., and Coors V. (2019). One dimensional convolutional neural network architectures for wind prediction. Energy Conversion and Management, 195, 70-75.

[15] Jalali S.M.J., Ahmmadian S., Khodayar M., Khosravi et.al. (2022). An advanced short-term wind power forecasting framework based on the optimized deep neural network models. Int. J. of Electrical Power and Energy Systems, vol.141, 108143.

[16] Javadi M., Malyscheff A.M., Wu D., Kang C., and Jiang J.N. (2018). An algorithm for practical power curve estimation of wind turbines. CSEE J. of Power and Energy Systems, 4(1), 93-102.

[17] Jia W., Wang H., Sun P., and Li J. (2019). Wind power prediction based on stacked autoencoder and LSTM. Energy Conversion and Management, 194, 78-87.

[18] Karam C., Zini J., Awad M., Saade C., Naffaa L., and Amina M. (2021). A Progressive and cross-domain deep transfer learning framework for wrist fracture detection. J. of Artificial Intelligence and Soft Computing Research, 12(2) 101-120.

[19] Kingma D.P., and Ba J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[20] Li X., Jia H., and Zhang Y. (2020). Short-term wind power prediction based on a novel hybrid algorithm combining the convolutional neural network and the improved differential evolution algorithm. Applied Energy, 275, 115373.

[21] LeCun Y., Bengio Y., and Hinton G. (2015). Deep learning. Nature, 521(7553), 436-444.

[22] Li Y., Cui Y., Li Y., Liu J., and Cao Y. (2020). Wind power forecasting based on EMD-LSTM neural network. Applied Energy, 261, 114441.

[23] Imseng D., Doss M.M., Bourlard H. (2010). Hierarchical multilayer perceptron based language identification. Proc. Interspeech 2010, 2722-2725, doi: 10.21437/Interspeech.2010-721

[24] Lin C.Y., Huang S.M., and Liao Y.C. (2018). A review of wind power point forecasting models: Current status and future perspectives. Renewable and Sustainable Energy Reviews, 82(Pt.1), 1-18.

[25] Lin J., Li Y., Li C., Li J., Li S., and Lin L. (2019). A novel approach for short-term wind power prediction based on improved KNN algorithm and particle swarm optimization. Applied Energy, 236, 350-365.

[26] Lipu M.S.H., et al. (2021). Artificial intelligence based hybrid forecasting approaches for wind power generation: progress challenges and prospects. IEEE Access. vol. 9, pp. 102460-102489.

[27] Liu T., Huang Z., Tian L., Zhu Y., Wang H., and Feng S. (2021). Enhancing wind turbine power forecast via convolutional neural network. Electronics, 10(3), 261.

[28] Ludwig S. (2022). Performance Analysis of data fusion methods applied to epileptic seizure recognition. J. of Artificial Intelligence and Soft Computing Research, 12 (1) 5-17.

[29] McTigue M.F., Ju P., and Krause P.C. (1997). Doubly fed induction generator using back-to-back PWM converters and its application to variable-speed wind-energy generation. IEEE Transactions on Industry Applications, 33(2), 461-468.

[30] Niksa-Rynkiewicz T., Szewczuk-Krypa N., Witkowska A., Cpałka K., Zalasiński M. and Cader, A. (2021). Monitoring regenerative heat exchanger in steam power plant by making use of the recurrent neural network. J. of Artificial Intelligence and Soft Computing Research,11(2) 143-155.

[31] Niksa-Rynkiewicz T., Witkowska A., Głuch J., and Adamowicz M. (2022). Monitoring the gas turbine start-up phase on the platform using a hierarchical model based on Multi-Layer Perceptron networks. Polish Maritime Research, 29, 123-131.

[32] Nguyen H., Nguyen T., Nowak J., Byrski A., Siwocha A. and Le V. (2022). Combined YOLOv5 and HRNet for high accuracy 2D keypoint and human pose estimation. J. of Artificial Intelligence and Soft Computing Research, 12(4) 281-298.

[33] Prasad D.K., Islam M.R., Tabassum-Abbasi, et al. (2018). Wind power prediction using machine learning techniques: A comprehensive review. Renewable and Sustainable Energy Reviews.

[34] Qing K., and Zhang R. (2021). Position-Encoding Convolutional Network to solving connected text Captcha. J. of Artificial Intelligence and Soft Computing Research, 12(2) 121-133.

[35] Rutkowska D. (2002). Neuro-Fuzzy Architectures and Hybrid Learning. Physica-Verlag. A Springer-Verlag Company.

[36] Rutkowski L., and Cpałka K. (2000). Flexible structures of neuro-fuzzy systems. Quo Vadis Computational Intelligence, Studies in Fuzziness and Soft Computing, 54, 479-484.

[37] Singh U., and Rizwan M. (2022). SCADA system dataset exploration and machine learning based forecast for wind turbines. Results in Engineering, vol.16, 100640.

[38] Shan Y., Xiong T., Zhang Z., and Wei X. (2018). A hybrid intelligent approach for short-term wind power prediction. Renewable Energy, 125, 62-72.

[39] Słowik A. (2011). Application of evolutionary algorithm to design minimal phase digital filters with non-standard amplitude characteristics and finite bit word length. Bulletin of the Polish Academy of Sciences-Technical Sciences, 59(2), 125-135.

[40] Słowik A., and Białko M. (2004). Design and optimization of combinational digital circuits using modified evolutionary algorithm. Proc. of 7th Int. Conf. on Artificial Intelligence and Soft Computing, ICAISC 2004, Lecture Notes in Artificial Intelligence. vol. 3070, pp. 468-473.

[41] Słowik A., and Białko M. (2008). Design and Optimization of IIR Digital filters with non-standard characteristics using continuous ant colony optimization algorithm. Proc. of 5th Hellenic Conference on Artificial Intelligence, SETN 2008, Lecture Notes in Artificial Intelligence. vol. 5138, pp. 395-400.

[42] Słowik A., and Białko M. (2007). Design of IIR digital filters with non-standard characteristics using differential evolution algorithm. Bulletin of the Polish Academy of Sciences-Technical Sciences. 55(4), 359-363.

[43] Słowik A., and Białko M. (2006). Partitioning of VLSI circuits on subcircuits with minimal number of connections using evolutionary algorithm. Proc. of 8th Int. Conf. on Artificial Intelligence and Soft Computing, ICAISC 2006, Lecture Notes in Computer Science. vol. 4029, pp. 470-478.

[44] Szczypta J., Przybył A., and Cpałka K. (2013). Some aspects of evolutionary designing optimal controllers. Artificial Intelligence and Soft Computing, Lecture Notes in Computer Science, 7895, 91-100.

[45] Szewczuk-Krypa N., Kolendo P., Głuszek J., Drop M., and Aronowski J. (2022). A new method of wind farm active power curve estimation based on statistical approach. Przegląd Elektrotechniczny, 98(1), 19-26.

[46] Tang J., Ma Z., Chen H., and Yang F. (2019). Probabilistic forecasting of wind power generation using extreme learning machine and cloud model. Applied Energy, 254, 113654.

[47] Tsai W., Hong C., Lin W., Tu C., and Chen C. (2023). A review of modern wind power generation forecasting technologies. Preprints.org 2023, 2023040917.

[48] Wang X., Guo P., and Huang X. (2011). A Review of wind power forecasting models. Energy Procedia, 12, 770-778.

[49] Wang Y., Hu Q., Srinivasan D., and Wang Z. (2019). Wind power curve modeling and wind power forecasting with inconsistent data. IEEE Transactions on Sustainable Energy, 10(1), 16-25.

[50] Yang S., Yu X., and Zhou Y. (2020). LSTM and GRU neural network performance comparison study: Taking Yelp Review Dataset as an example. 2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI), Shanghai, China, 2020, pp. 98-101, doi: 10.1109/IWECAI50956.2020.00027.

[51] Yuan X., Zou W., and Zhang H. (2020). Short-term wind speed forecasting using deep learning: An empirical comparison of long short-term memory, convolutional neural network, and deep belief network. Renewable Energy.

[52] Zalasiński M., Cpałka K., and Hayashi Y. (2015). New fast algorithm for the dynamic signature verification using global features values. Artificial Intelligence and Soft Computing, Lecture Notes in Computer Science, 9120, 175-188.

[53] Zhang H., Wu S., Zhang X., et al. (2021). A novel short-term wind power prediction approach based on wavelet transform and extreme learning machine optimized by firefly algorithm. Energy Conversion and Management.

[54] Zhang X., Li Y., Wang Q., and Wang G. (2019). Short-term wind power prediction using an artificial neural network ensemble based on principal component analysis and fuzzy c-means clustering. Renewable Energy.

[55] Zhang Y., Zhang W., Shi D., and Lu Y. (2018). A hybrid method of wind speed and wind power prediction based on fuzzy clustering analysis and support vector regression. Energy Conversion and Management, 157, 203-212.

Tacjana Niksa-Rynkiewicz received the M.Sc. and Ph.D. degrees from the Częstochowa University of Technology, Częstochowa, Poland, in 2004 and 2011, respectively. She is currently an Assistant Professor with the Faculty of Ocean Engineering and Ship Technology, the Gdansk University of Technology. Dr. Niksa-Rynkiewicz has authored over 10 publications. Her current research interests include computational intelligence, data mining, and expert systems.
https://orcid.org/0000-0001-9202-8454

Piotr Stomma received his M.Sc.Eng degree from the Interdisciplinary Centre of Mathematical Modelling at the University of Warsaw, Poland, in 2022. Since October 2022, he has been working as a research assistant in the Institute of Computer Science, Division of Bioinformatics, at the University of Białystok in Poland. His primary areas of interest include graph clustering, dimensionality reduction, statistical feature selection, interpretable machine learning, and predictive analytics. Currently, he is involved in a project funded by the National Science Centre of Poland, which focuses on discovering synergistic relationships in the microbiome and human exposome.
https://orcid.org/0000-0002-2603-8205

Anna Witkowska holds an M.Sc. in mathematics and computer science from the University of Gdańsk, Poland in 2001. She received her Ph.D. degree from the Technical University of Warsaw in 2011, and the D.Sc. (habilitation) degree in automation, electronic and electrical engineering from the Technical University of Gdańsk, in 2020. Currently, she is a professor of the Gdańsk University of Technology, Faculty of Electrical and Control Engineering. Her research interests include automation, especially adaptive and robust control of nonlinear systems as ocean vehicles and control allocation methods in ship dynamic positioning systems.
https://orcid.org/0000-0001-9594-6832

Danuta Rutkowska is a professor of computer science. She graduated from Wrocław University of Science and Technology, Wrocław, Poland, from where she also received her Ph.D. in automation and then D.Sc. in computer science. In 2002 she was given the title of professor conferred by the President of Poland. She is an author or co-author of numerous publications, mostly in computational intelligence, including several books, book chapters, and many scientific papers. Her research interests are in the area of artificial intelligence, especially computational intelligence, focusing on artificial neural networks, fuzzy systems, genetic/evolutionary algorithms, as well as hybrid intelligent systems, i.e. neuro-fuzzy or genetic-neuro-fuzzy systems, and their applications. After working for many years at the Częstochowa University of Technology, Częstochowa, Poland, she is currently a professor in the Institute of Information Technology at the University of Social Sciences in Łódź, Poland. Since 2022 she has been serving as vice-rector for science at this university.
https://orcid.org/0000-0003-0217-2589

Adam Słowik received the B.Sc. and M.Sc. degrees in computer engineering and electronics in 2001 and the Ph.D. degree with distinction in 2007 from the Department of Electronics and Computer Science, Koszalin University of Technology, Koszalin, Poland. He received the Dr. habil. degree in computer science (intelligent systems) in 2013 from the Department of Mechanical Engineering and Computer Science, Czestochowa University of Technology, Czestochowa, Poland. Since October 2013, he has been an Associate Professor in the Department of Electronics and Computer Science, Koszalin University of Technology. His research interests include soft computing, computational intelligence, and, particularly, bioinspired optimization algorithms and their engineering applications. He is a reviewer for many international scientific journals. He is an author or coauthor of over 100 refereed articles in international journals, two books, and conference proceedings, including one invited talk. Dr. Slowik is an Associate Editor of the IEEE Transactions on Industrial Informatics. He is a member of the program committees of several important international conferences in the area of artificial intelligence and evolutionary computation. He was a recipient of one Best Paper Award (IEEE Conference on Human System Interaction- HSI 2008).
https://orcid.org/0000-0003-2542-9842

Krzysztof Cpałka received the M.Sc. and Ph.D. degrees from the Częstochowa University of Technology, Częstochowa, Poland, in 1997 and 2002, respectively. In 2010 he obtained a D.Sc. degree in computer science at the Systems Research Institute of the Polish Academy of Sciences in Warsaw, Poland. Since 2010, he has been a professor at the Department of Intelligent Computer Systems, Częstochowa University of Technology. Krzysztof Cpałka is the author of two books and over 100 refereed papers. His research interests include population-based algorithms, neural networks, fuzzy systems, and their applications.
https://orcid.org/0000-0001-9761-118X

Joanna Jaworek-Korjakowska, Univ. Professor, Director of the Centre of Excellence in Artificial Intelligence, and Deputy Head of the Department of Automatic Control and Robotics at the AGH University in Krakow, Poland. She is an expert at the Confederation of Laboratories for Artificial Intelligence Research in Europe (CLAIRE), a member of IEEE, Polish Artificial Intelligence Society, and International Dermoscopy Society as well as an alumnus of the TOP 500 Innovator programme at Stanford University, USA. Her main research interests focus on computer vision, data mining, artificial intelligence especially deep learning methods, anomaly detection as well as clustering. J. Jaworek has been awarded Honorable Mention Award during the CVPR'19 conference (ISIC workshop) and Bekker Fellowship'22 to conduct research at Stanford University, USA.
https://orcid.org/0000-0003-0146-8652

Piotr Kolendo graduated with a Master's degree in 2010 and obtained a Doctor of Engineering degree in 2016 from the Faculty of Electrical and Control Engineering at Gdańsk University of Technology. He is currently employed at the Gdańsk Branch of the Institute of Power Engineering. His research interests focus on area voltage regulation, particularly ARNE/ARST group control systems, and multi-criteria optimization. He has authored over 20 publications in national and international journals. Piotr Kolendo has developed the regulation concept and implemented group voltage control systems in major power plants across Poland and Eastern Europe, including Bełchatów, Vilnius, Kozienice, Turów, Opole, Dolna Odra, Żarnowiec, Żar-Porąbka, Żydowo, Jaworzno, Włocławek, combined heat and power plants in Żerań, Gorzów, Karolin, Gdańsk, Ostrołęka, Stalowa Wola, Zielona Góra, and the largest Polish refinery, PKN Orlen in Płock.
https://orcid.org/0000-0001-9292-5980