2023 International Conference for Technological Engineering and its Applications in Sustainable Development (ICTEASD) | 979-8-3503-3647-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICTEASD57136.2023.10585156

Wind energy forecasting based on integration of CNN and Bidirectional RNN

1st Dr.R. Udayakumar
Department of CS & IT
Kalinga University, India
rsukumar2007@gmail.com,
deancsit@kalingauniversity.ac.in

2nd S. Jayasree
S&H
Prince Shri Venkateshwara Padmavathy Engineering College, Chennai - 127
jayasree.s_chem@psvpec.in

3rd Shruti Bhargava Choubey
Sreenidhi Institute of Science and Technology
Yamnampet, Ghatkesar, Hyderabad, 501301
shrutibhargava@sreenidhi.edu.in

4th M. Sasikumar
Computer Science and Engineering
Vel Tech Rangarajan, Dr. Sagunthala R&D Institute of Science and Technology, Chennai-600062 Tamil Nadu, India
sasikumarm@veltech.edu.in

5th V. Shanmuganeethi
Computer Science and Engineering
NITTTR Chennai
shanneethi@nitttrc.edu.in

6th Kabulov Ilyos
Tashkent Institute of Irrigation and Agricultural Mechanization Engineers National Research University, Tashkent, Uzbekistan, Western Caspian University
Baku, Azerbaijan
bobur.toshbekov.wm@gmail.com

7th Hassan M. Al-Jawahry
Department of computers Techniques engineering
College of technical engineering
The Islamic university, Najaf, Iraq
hassanaljawahry@gmail.com

8th Gummagatta Yajaman Vybhavi
GRIET, Hyderabad,
Telangana, India
Gummagatta.Yajaman.Vybhavi@gmail.com

***Abstract*—The variability and intermittency of integrating large-scale wind generation pose a significant risk to the reliability and integrity of the electricity grid. Precise wind power forecasting is crucial for ensuring the reliability of wind power grid integration. This research presents a Wind Energy Forecasting (WEF) method that combines a Convolutional Neural Network with a Bidirectional Recurrent Neural Network (CNN-BiRNN). The Convolutional Neural Network (CNN) analyzes the spatial characteristics present in weather observations, extracting significant features. Meanwhile, the Bidirectional Recurrent Neural Network (BiRNN) manages the temporal relationships within the serial time-series information. By using this holistic strategy, the model can successfully represent the intrinsic temporal and spatial trends in wind energy production, leading to more precise and reliable forecasts. A refined ResNet150-based CNN architecture has been utilized to obtain profound latent characteristics that are highly representative and selective, resulting in enhanced accuracy. An assessment of the CNN-BiRNN model and the individual CNN and BiRNN models was conducted using weather data and a wind energy dataset from a wind farm. The study aimed to assess the performance of these models in multi-step WEF. The findings demonstrate that the suggested CNN-BiRNN has superior capability in extracting spatial and temporal features compared to the conventional structural model. The comparison indicates that the CNN-BiRNN model surpasses the performance of separate CNN and BiRNN models with reduced error at different time intervals, highlighting its potential to improve the accuracy of WPF.**

Keywords— Wind Energy Forecasting, Convolutional Neural Network, Bidirectional Recurrent Neural Network, ResNet150, Accuracy

I. Introduction

Concerns regarding the sustainability of renewable energy sources, like solar and wind power, are growing, and there have been notable shifts in the mix of energy sources used to generate electricity [1]–[2]. Globally, these resources are growing at the fastest rate per year, which indicates a rapid energy transition. As an illustration, the capacity of wind power globally increased from 178 GW in 2009 to 745 GW in 2020. Concurrently, the solar power capacity increased from 38 GW to 592 GW [3]. Because of the increased economic benefits of solar and wind power, further growth in their penetration is anticipated.

The weather-dependent variability of these energy resources, however, poses a risk to power system operations' economic efficiency and dependability, which could result in large-scale social and financial losses [4]. Forecasting variable renewable energy (VRE) is the most basic and useful front-end application among supply-side variability handling techniques. Its accuracy makes the VRE's grid integration safe and affordable [5]. It has been established that wind power's extremely unpredictable qualities make it less predictable than solar power. Furthermore, unlike solar generators, wind generators are typically installed as wind power plants as opposed to distributed generators. This geographic smoothing and aggregation contributes to the reduction of operating reserve requirements, which is especially advantageous for bulk power system operations. Thus, a number of studies have been carried out to look into ways to improve the WEF and the impact of aggregated wind power on the power system [6].

Three categories exist for WEF methods: CNN, RNN-based method, physical method, and conventional statistical method. In order to complement each other, a hybrid model combining more than two of the aforementioned methods has been studied [7]. The numerical weather prediction system (NWP), which represents the mathematically expressive model based on diverse geographical and meteorological data, is the foundation of the physical approach, which expands upon it. This approach has limitations for short-term forecasting because it is challenging to compile all the necessary geographical or meteorological data, even though it works well for medium-term forecasting periods longer than three hours [8].

II. Literature Survey

Decomposition-based hybrid WEF model by Jaseena and Kovoor (2021) uses deep bidirectional LSTM networks. The method used Empirical Mode Decomposition (EMD) to analyze wind speed time data and deep bidirectional LSTM networks to predict. The model accurately reproduces the complex patterns of wind speed fluctuations using spectral and temporal properties [10]. Deep bidirectional LSTM networks were trained using wind speed data. The model accurately predicts wind speeds, demonstrating its ability to identify temporal relationships and improve forecasts.

Nguyen and Phan (2022) developed an hourly wind speed prediction method. They use EMD, CNN-Bi-LSTM, and GA optimization in a hybrid model [11]. The authors start by breaking down the wind speed time series into its intrinsic mode functions using Ensemble EMD. Prediction is then done with a CNN with Bi-LSTM architecture and GA optimization. The hybrid model is trained using wind speed data. The output numbers accurately predict wind speeds for the next hour. The benefits include better prediction, resilience to complex and unpredictable factors, and wind pattern adaptation. However, parameter adjustment and optimization computing complexity can be drawbacks.

Garg and Krishnamurthi (2023) developed a CNN encoder-decoder LSTM model to predict sustainable wind power. CNNs extract and encode features, and then LSTM layers model sequences [12]. The model is designed to predict wind power output using historical data. You teach the CNN encoder-decoder LSTM model to learn from wind power measurements. The model's output values accurately predict wind power production, showing its ability to capture complex spatial and temporal relationships. Its ability to process sequential input and capture hierarchical characteristics is beneficial. However, large datasets and processing resources may be drawbacks.

Anu Shalini and Sri Revathi (2022) studied renewable energy power production forecasting. CNN-based Bidirectional LSTM was used for deep learning. The suggested method uses a CNN to extract features and a BiLSTM network for sequential modeling [13]. The model is trained using renewable power production data. The output values accurately predict power production, demonstrating the model's ability to capture complex renewable energy generation patterns. Precision in predicting renewable energy generation makes it ideal for power system integration. However, a lot of annotated data and computing power may be needed.

Paramasivan studied deep learning-based RNNs in 2021 to improve wind energy forecasts. Many papers using RNNs for wind energy forecasting were reviewed [14]. Execution focused on condensing findings and methodology from examined investigations. The output values show how well RNNs capture temporal relationships, improving wind energy prediction. Integrating current data and pattern recognition in wind energy forecasts using RNNs is beneficial. The study's review-oriented design may limit implementation details, which can have drawbacks.

Zhang and Wang (2023) developed a new wind speed prediction method. Multi-head attention-based probabilistic CNN-BiLSTM uses many methods to make accurate forecasts [15]. The method captures spatial and temporal wind speed relationships using multi-head attention mechanisms and a hybrid CNN-BiLSTM architecture. The model is trained using past wind speed data, considering probabilistic predictions. Probabilistic wind velocity forecasts reveal uncertainty. Benefits include better uncertainty management and prediction accuracy. However, computational complexity and attention mechanism calibration may be drawbacks.

Using temporal and spatial feature extraction, Zhen et al. (2020) developed a hybrid DL model for wind power predictions [16]. Integrating a DL model that extracts temporal-spatial features improves wind power estimates. Wind power data is used to instruct the model. The model captures complex spatiotemporal patterns to predict wind power production accurately. Integrating temporal-spatial characteristics improves precision. Drawbacks may include model complexity and the need for large datasets.

Ko et al. (2020) developed a deep concatenated residual network with biLSTM to predict wind power one hour ahead [17]. The technology captures short-term and long-term wind power data relationships using a deep concatenated residual network and BiLSTM layers. Wind power data is used to train the model during implementation. The output numbers accurately predict wind power for the next hour. Short-term wind power prediction is more precise, making the model suitable for real-time applications. However, computational complexity and parameter adjustment may be drawbacks.

The knowledge acquired from these studies provides a foundation for further progress in creating precise and dependable models for WEF, which is essential for maximizing the integration of renewable energy into power systems.

III. Wind Energy Forecasting Based on the Integration of CNN and Bidirectional RNN

The integration of CNN and BiRNN in the WEF technique (shown in Fig. 1) is a sophisticated approach that enhances the precision of predictions in the renewable energy sector.

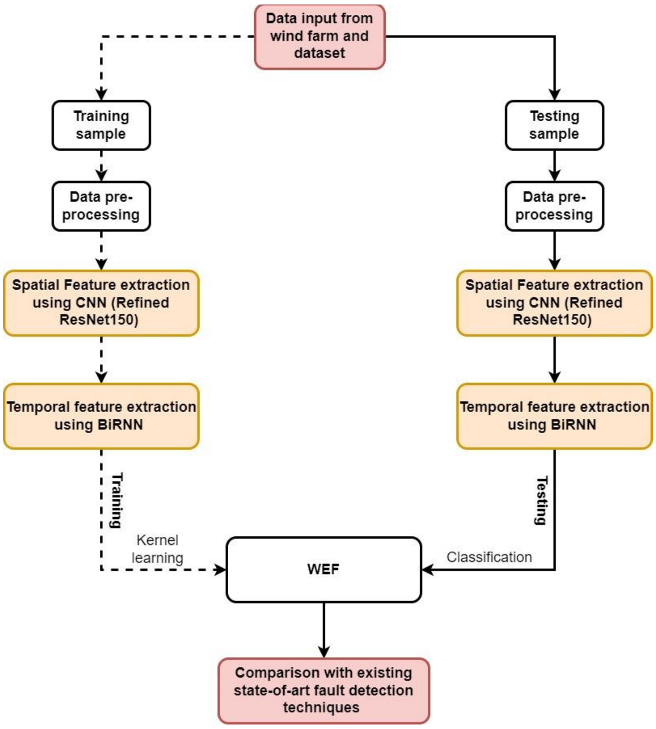

Fig. 1. Framework for wind energy forecasting based on the integration of CNN and Bidirectional RNN

Initially, data was collected, and pre-processing of the dataset was done. This novel approach utilizes CNN to extract intricate spatial features from meteorological data, encompassing wind speed, temperature, and pressure maps. The BiRNN models the temporal dependencies of wind patterns by incorporating information from both preceding and subsequent time intervals, leading to a holistic comprehension. The model can effectively comprehend the complex spatial and temporal patterns associated with wind energy generation by integrating these two neural network architectures. This technique showcases outstanding expertise in gathering intricate patterns within the data and also provides more accurate and robust predictions. It emphasizes the potential of combining spatial and temporal feature learning to improve WEF capabilities.

A. Refined ResNet150-based CNN architecture

A refined ResNet150-based CNN architecture has been utilized to obtain profound latent characteristics that are highly representative and selective, resulting in enhanced accuracy.

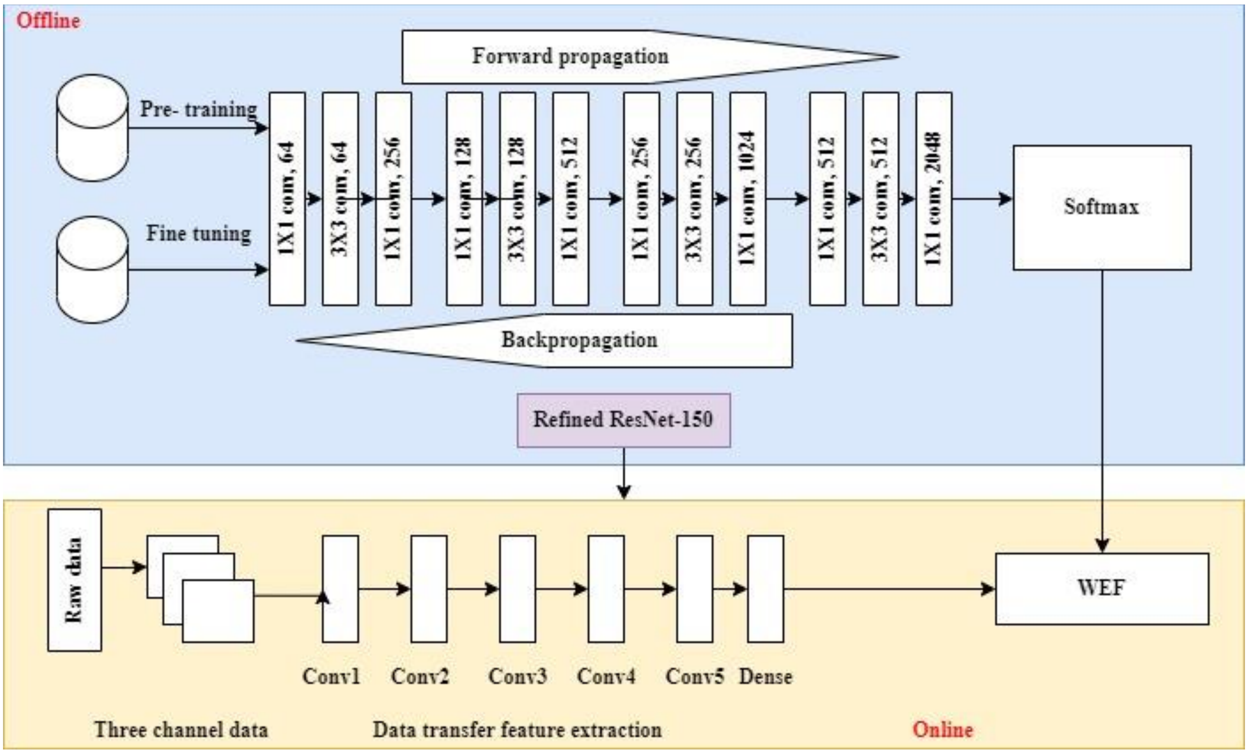

Fig. 2. Refined ResNet150-based CNN architecture using Transfer Learning (TL)

Fig. 2 depicts Refined ResNet150-based CNN architecture using Transfer Learning (TL). ResNet, short for Residual Network, is a CNN architecture that facilitates training neural networks with many layers. This is achieved by incorporating residual connections, which enable the network to learn residual functions instead of directly learning the desired underlying mapping. ResNet150 is an enhanced iteration of the initial ResNet model, featuring 150 layers. This enables the creation of more complex network structures and enhances overall performance. Inadequate data can impede the effectiveness of ML/DL models in WEF. Lack of adequate data can lead to the problems of overfitting or underfitting, which in turn diminishes the accuracy of the model. This study tackles the issue by employing the Refined ResNet150 framework to extract spatial features using TL. The extracted features are anticipated to possess greater representativeness and discriminability, resulting in enhanced accuracy in the WEF.

The process consists of two distinct stages: offline training and online detection. During the offline stage, the raw training data produces tri-channel information. The three-channel information is subsequently employed to refine the pre-trained ResNet-150 model, which generates profound transfer features. The features are subsequently employed to categorize whether the monitoring goal has attained the state of failure. The framework of three-channel data remains consistent in both the offline and online phases. The profound transfer characteristics are subsequently generated by inputting this three-channel data into the deep transfer algorithm. The detections are obtained through the logistic regression technique, specifically the softmax classification. This is a comprehensive elucidation of every constituent of the procedure.

B. The architecture of DL

DL is an effective method for identifying potential correlations among disparate datasets. It systematically reveals latent attributes within the initial data and employs them for regression or clustering. The correlation between the depth of a neural network and its efficiency is crucial. The neural networks acquire increasingly complex and refined features through layering. Therefore, the higher the number of tiers in a network, the more efficient it becomes in categorizing and predicting. However, the authors found that the accuracy of identification may reach a maximum or decrease when the length of a system increases. This phenomenon is referred to as DL model degradation. It suggests that the process of training deep neural networks is challenging. The fundamental principle of the structure is that when a complex system is built by adding new layers to a simple design, the worst-case circumstance is that the extra layers do not gain any further functionality and duplicate the behavior of the outside system. This suggests that the newly introduced dimension replicates the old layer without any changes. Given the circumstances, the deep network is expected to perform at a minimum level equal to or better than the shallower system.

The fully connected layer is described explicitly by Equation (1).

$$K_{fc} = f_{W,b}^{C\rightarrow}(K_c)$$
$$\hat{O} = soft\_\max(K_{fc}) \tag{1}$$

where K_{fc} represents the outcome of the new fully interconnected layers and K_c represents the result of the pre-trained convolutional level.

During the offline training stage, the parameters of the newly added layers are adjusted, while the parameters of the pre-trained convolutional layers are kept fixed. To achieve the objective of modeling updates, fine-tuning involves performing guided WEF. The outcome of the profound transfer system is represented by$\breve{O}$. The SoftMax algorithm utilizes logistic regression to determine the class of the incoming information. The SoftMax equation calculates the probability, denoted as p_x, that the specimen belongs to a particular group. In Equation (2), the term e $\exp(Z_x)$denotes the output of a particular neuron within the fully interconnected neural networks. The expression $\sum_{y=0}^{N} \exp(Z_y)$ represents the sum of all neuronal responses in fully interconnected neuronal networks.

$$p_x = \frac{\exp(Z_x)}{\sum_{y=0}^{N} \exp(Z_y)} \tag{2}$$

During the training phase, the cross-entropy loss functions (LF) calculate the difference between the forecasting result and the actual detection of the wind energy. Equation (3) displays the LF.

$$LF = \frac{1}{M}\sum_{x=0}^{M} -\{O_x * \log(p_x) + (1 - O_x) * \log(1 - O_x)\} \tag{3}$$

In Equation (3), O_x is the labeling of the x^{th} specimen, 1 is a positive class, and 0 is the negative class. $p_x$signifies the probability that the x^{th} sampling will be positive.

In the offline stage, information from additional elements is used to build a three-channel model. Lastly, the study uses the three-channel data to refine the ResNet-150 model that was previously learned to get spatial transfer properties.

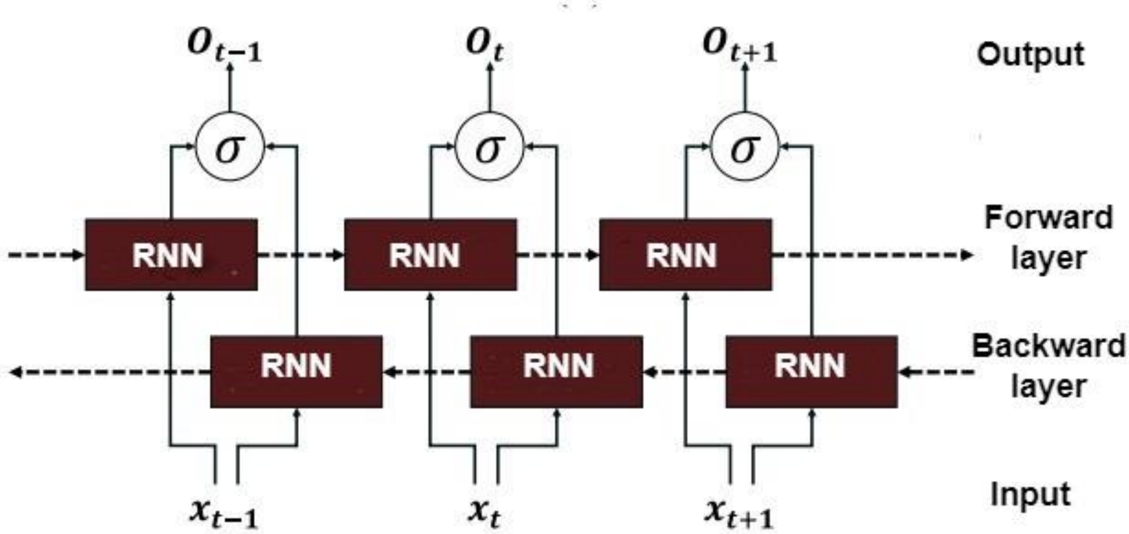

Fig. 3. The architecture of BiRNN

The architecture of BiRNN is shown in Fig. 3. The accuracy of a traditional RNN may be increased by bidirectional learning. The idea that the output is a component of the continuous correlation rather than the exclusive result of earlier inputs has led to the adoption of bidirectional learning. BiRNN trains its parameters in both forward and backward directions to comprehend the context. RNNs are trained only in the forward route; this training approach can capture characteristics or patterns in both directions. Compared to traditional RNN, BiRNN performed better and showed more sequential learning accuracy.

The BiRNN architecture manages the temporal ties in the serial time-series data. This comprehensive approach enables the model to capture the inherent temporal and geographical patterns in wind energy output, producing more accurate and consistent predictions. Ultimately, the components train the detection algorithm based on SoftMax. The temporal deep features are extracted using the BiRNN method. Finally, WEF results are acquired by applying these comprehensive properties to the SoftMax classifier classification model.

IV. Results and discussion

This study introduces the wind energy database from ERCOT with the temperature dataset from Dresden [18]. Every database consists of hourly mean information, including just one feature and no additional factors. The values of each dataset undergo pre-processing to be scaled within the range of 0 to 1. Subsequently, every dataset is partitioned into training and testing subsets. The training set is used to adjust the specifications of the prediction models, while the test set assesses the efficacy of the chosen model.

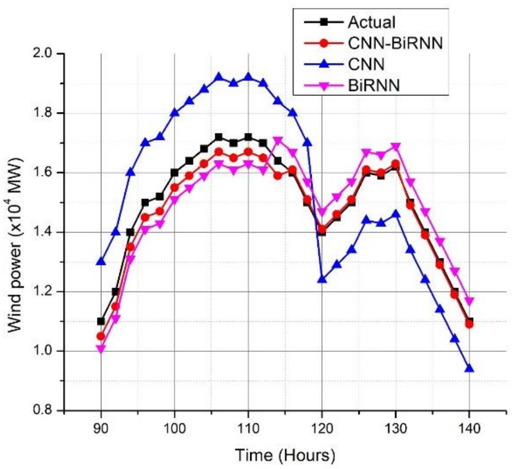

Fig. 4. Wind Power Forecasting output (x104 MW) results using CNN, BiRNN, and the proposed CNN-BiRNN along with original values

Fig. 4 illustrates the WPF output (x104 MW) results using CNN, BiRNN, and the proposed CNN-BiRNN along with original values. The actual wind power outputs are used as the definitive measure for assessment. At time point 90, the observed output is 1.1 x104 MW, while the CNN-BiRNN model forecasts 1.05 x104 MW, the CNN model forecasts 1.3 x104 MW, and the BiRNN model forecasts 1.01 x104 MW. The tendency persists for successive time intervals. The suggested CNN-BiRNN routinely exhibits precise forecasts, closely matching the real values. This indicates the efficacy of integrating CNN with BiRNN for wind power prediction. The comparison demonstrates that the CNN-BiRNN model surpasses the performance of separate CNN and BiRNN models at different time intervals, highlighting its potential to improve the accuracy of WPF.

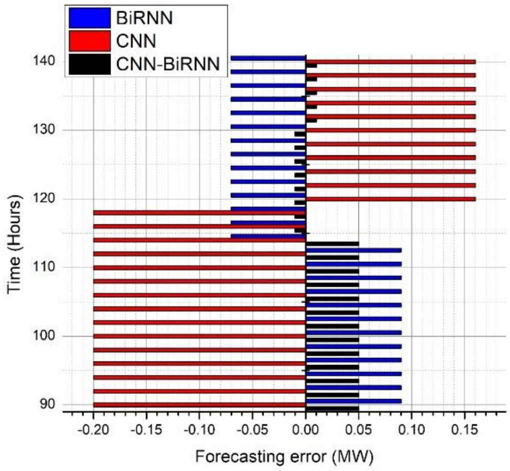

Fig. 5. Forecasting error (MW) of BiRNN, CNN, and the proposed CNN-BiRNN for WPF

Fig. 5 depicts the forecasting error (MW) of BiRNN, CNN, and the proposed CNN-BiRNN for WPF. The forecasting error is the difference between the projected and actual wind power outputs. At time point 90, the CNN-BiRNN model has an error of 0.05 MW, whereas the CNN model has an error of -0.2 MW, and the BiRNN model has an error of 0.09 MW. The tendency persists for successive time intervals. The suggested CNN-BiRNN consistently maintains a low error, demonstrating its capacity for precise wind energy projections. The proposed CNN-BiRNN reduces errors and improves the accuracy of WEF compared to using separate CNN and BiRNN models. The presence of negative errors at some periods suggests a little underestimating. However, in general, the CNN-BiRNN model surpasses or equals the performance of the other models, emphasizing its effectiveness in WEF.

V. Conclusion

This study introduces a WEF technique integrating a Convolutional Neural Network with a Bidirectional Recurrent Neural Network (CNN-BiRNN). The CNN examines the spatial attributes of weather data, identifying notable elements. The BiRNN handles the temporal dependencies in the sequential time-series data. By implementing this comprehensive approach, the model effectively captures the inherent patterns in generating wind energy, including temporal and geographical aspects. As a result, the predictions generated are characterized by enhanced accuracy and dependability. A sophisticated ResNet150-based CNN architecture has been used to extract deep latent features that are both highly representative and selective, leading to improved accuracy. An evaluation was performed on the CNN-BiRNN model and the separate CNN and BiRNN models, utilizing meteorological data and a wind energy dataset obtained from a wind farm. The research sought to evaluate the efficacy of these models in WEF systems. The results indicate that the proposed CNN-BiRNN outperforms the traditional structural model in extracting spatial and temporal data. The comparison reveals that the CNN-BiRNN model outperforms the standalone CNN and BiRNN models by achieving lower error rates at various intervals. This emphasizes the model's potential to enhance the accuracy of WPF.

REFERENCES

[1] Du, E., Zhang, N., Hodge, B. M., Wang, Q., Lu, Z., Kang, C., ... & Xia, Q. (2018). Operation of a high renewable penetrated power system with CSP plants: A look-ahead stochastic unit commitment model. IEEE Transactions on Power Systems, 34(1), 140-151.

[2] Khodayar, M., Mohammadi, S., Khodayar, M. E., Wang, J., & Liu, G. (2019). Convolutional graph autoencoder: A generative deep neural network for probabilistic spatio-temporal solar irradiance forecasting. IEEE Transactions on Sustainable Energy, 11(2), 571-583.

[3] Dhabi, A. (2020). Irena. Renewable energy statistics.

[4] Rehmani, M. H., Reisslein, M., Rachedi, A., Erol-Kantarci, M., & Radenkovic, M. (2018). Integrating renewable energy resources into the smart grid: Recent developments in information and communication technologies. IEEE Transactions on Industrial Informatics, 14(7), 2814-2825.

[5] Sinsel, S. R., Riemke, R. L., & Hoffmann, V. H. (2020). Challenges and solution technologies for the integration of variable renewable energy sources—a review. renewable energy, 145, 2271-2285.

[6] Jia, M., Shen, C., & Wang, Z. (2018). A distributed probabilistic modeling algorithm for the aggregated power forecast error of multiple newly built wind farms. IEEE Transactions on Sustainable Energy, 10(4), 1857-1866.

[7] Liu, H., Mi, X. W., & Li, Y. F. (2018). Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy conversion and management, 156, 498-514.

[8] Zhu, R., Liao, W., & Wang, Y. (2020). Short-term prediction for wind power based on temporal convolutional network. Energy Reports, 6, 424-429.

[9] Ezzat, A. A. (2020). Turbine-specific short-term wind speed forecasting considering within-farm wind field dependencies and fluctuations. Applied Energy, 269, 115034.

[10] Jaseena, K. U., & Kovoor, B. C. (2021). Decomposition-based hybrid wind speed forecasting model using deep bidirectional LSTM networks. Energy Conversion and Management, 234, 113944.

[11] Nguyen, T. H. T., & Phan, Q. B. (2022). Hourly day ahead wind speed forecasting based on a hybrid model of EEMD, CNN-Bi-LSTM embedded with GA optimization. Energy Reports, 8, 53-60.

[12] Garg, S., & Krishnamurthi, R. (2023). A CNN encoder decoder LSTM model for sustainable wind power predictive analytics. Sustainable Computing: Informatics and Systems, 38, 100869.

[13] Anu Shalini, T., & Sri Revathi, B. (2022). Power generation forecasting using deep learning CNN-based BILSTM technique for renewable energy systems. Journal of Intelligent & Fuzzy Systems, 43(6), 8247-8262.

[14] Paramasivan, S. K. (2021). Deep learning based recurrent neural networks to enhance the performance of wind energy forecasting: A review. Revue d'Intelligence Artificielle, 35(1).

[15] Zhang, Y. M., & Wang, H. (2023). Multi-head attention-based probabilistic CNN-BiLSTM for day-ahead wind speed forecasting. Energy, 278, 127865.

[16] Zhen, H., Niu, D., Yu, M., Wang, K., Liang, Y., & Xu, X. (2020). A hybrid deep learning model and comparison for wind power forecasting considering temporal-spatial feature extraction. Sustainability, 12(22), 9490.

[17] Ko, M. S., Lee, K., Kim, J. K., Hong, C. W., Dong, Z. Y., & Hur, K. (2020). Deep concatenated residual network with bidirectional LSTM for one-hour-ahead wind power forecasting. IEEE Transactions on Sustainable Energy, 12(2), 1321-1335.

[18] https://www.nrel.gov/docs/fy23osti/79498.pdf

[19] Aftab, Muhmmad Waleed, Sajjad Hussain, Aftab Husain, Umar Ali Khan, and Hamza Kundi. "A Deep Reinforcement-Learning-Based Relay Selection for Underwater Sensors Network." Journal of Smart Internet of Things (2023).

[20] Lalit Kumar Tyagi, Rama Kant, Anish Gupta "A Comparative Analysis of Various Local Feature Descriptors in Content-Based Image Retrieval System" in International Journal of Physics, Dec 2020.