IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 6, 2021, accepted February 17, 2021, date of publication February 19, 2021, date of current version March 2, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3060654

A Short-Term Load Forecasting Method Using Integrated CNN and LSTM Network

SHAFIUL HASAN RAFI[1], NAHID-AL-MASOOD[2], (Senior Member, IEEE), SHOHANA RAHMAN DEEBA[3], (Member, IEEE), AND EKLAS HOSSAIN[4], (Senior Member, IEEE)

[1]Department of Electrical, Electronic and Communication Engineering, Military Institute of Science and Technology, Dhaka 1216, Bangladesh
[2]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka 1205, Bangladesh
[3]Department of Electrical and Computer Engineering, North South University, Dhaka 1229, Bangladesh
[4]Oregon Renewable Energy Center (OREC), Department of Electrical Engineering and Renewable Energy, Klamath Falls, OR 97601, USA

Corresponding author: Nahid-Al-Masood (nahid@eee.buet.ac.bd)

ABSTRACT In this study, a new technique is proposed to forecast short-term electrical load. Load forecasting is an integral part of power system planning and operation. Precise forecasting of load is essential for unit commitment, capacity planning, network augmentation and demand side management. Load forecasting can be generally categorized into three classes such as short-term, midterm and long-term. Short-term forecasting is usually done to predict load for next few hours to few weeks. In the literature, various methodologies such as regression analysis, machine learning approaches, deep learning methods and artificial intelligence systems have been used for short-term load forecasting. However, existing techniques may not always provide higher accuracy in short-term load forecasting. To overcome this challenge, a new approach is proposed in this paper for short-term load forecasting. The developed method is based on the integration of convolutional neural network (CNN) and long short-term memory (LSTM) network. The method is applied to Bangladesh power system to provide short-term forecasting of electrical load. Also, the effectiveness of the proposed technique is validated by comparing the forecasting errors with that of some existing approaches such as long short-term memory network, radial basis function network and extreme gradient boosting algorithm. It is found that the proposed strategy results in higher precision and accuracy in short-term load forecasting.

INDEX TERMS Short-term load forecasting, convolutional neural network, long-short-term memory network, Bangladesh power system, evaluation metrics.

I. INTRODUCTION

Accurate forecasting of electrical load is crucial for formulating the planning and operational strategies of power generation, transmission and distribution systems. Unit commitment and scheduling of the power plants significantly depend on the precise forecasting of load [1]. Also, operational cost cannot be estimated without accurate load prediction. Load forecasting is generally classified into three categories. These are short-term load forecasting, which predicts load for next few hours to few weeks; midterm load forecasting, which usually covers a week to a year; and long-term load forecasting, which predicts load for more than a year [2]. Day-ahead load prediction, which is imperative for power system operation and control, is done through short-term load forecasting.

Short-term load forecasting (STLF) is performed via various techniques in the literature. These procedures can be broadly divided into two groups such as traditional and artificial intelligence-based techniques. In traditional approaches, statistical methods are mostly used [3]. These include multiple linear regression [4], [5], exponential smoothing [3] and autoregressive integrated moving normal (ARIMA) [6] algorithms. Note that due to characteristics of non-linear features of time series univariate load data, the above-mentioned techniques do not always satisfactorily perform in STLF [6]–[8].

To mitigate this issue, machine learning based strategies are evolved and extensively used in STLF [9] These methods encompass clustering method [10], support vector machine

The associate editor coordinating the review of this manuscript and approving it for publication was Vahid Vahidinasab.

(SVR) [11], fuzzy logic framework [12], artificial neural network (ANN) [13]–[16], radial basis functional network (RBFN) [17] and hybrid methods [18], [19].

It has been reported in [20] that a kernel-based SVR model can be used for forecasting electric load. The developed model provides a novel approach for selecting kernel function of SVR model. The performance of the network is assessed by the real-world case of the Australia and California Power Grid. In [21], a method of hourly load forecasting using fuzzy logic has been presented. The load forecasting has been done using the one-year data from the large-scale power system. The proposed methodology uses fuzzy rules to incorporate historical load data with time and day, where day means weekend or weekdays. The aim of this work is to determine the probable load curve of a particular day by observing one-year data from large-scale industry.

In [22], an effective method based on ANN strategies fortified by a wavelet de-noising algorithm is developed to predict the short-term load demand. The obtained results from the proposed approach exhibit that it considerably increases the accurateness of prediction. In recent times, deep learning approaches have drawn a special attention because of having a greater number of hidden layers, which enables these models to deal with the complicated non-linear patterns [23], [24]. In [25], a wind power forecasting (WPF) of six-month-ahead is presented using a machine learning based algorithm. To train the WPF model, tree-based algorithms is applied, which includes decision tree, bagging, random forest, boosting (AdaBoost), gradient boosting, and extreme gradient boosting (XGBoost). The proposed framework provides prominent performance and accuracy in wind power forecasting of Ghadamgah wind farm.

In addition, recurrent neural network (RNN) is employed for STLF because of its effective learning ability to capture the non-stationary load data pattern. In [26], a RNN algorithm is implemented for domestic load forecasting that gives better performance in terms of root mean square error (RMSE). A modern load forecasting methodology using RNN is introduced in [27], which deploys a concept of one-step-ahead. The proposed method gives acceptable performance in low power demand and high power demand region. It also exhibits the smaller fluctuations in different regions compared to the other models. Note that vanishing gradient and exploding gradient problem arise in RNN, which reduce the prediction accuracy. Moreover, gated recurrent unit network (GRUN) has been extensively used in the recent years due to the absence of vanishing gradient problem [28]. In [29], a GRUN based algorithm is proposed for STLF with multi-source data. The mean average percentage error (MAPE) obtained from this network is minimum, which outperforms other current methods. In addition, long short-term memory (LSTM) network has been taken into account to solve gradient vanishing problems [30]. In [31], an effective methodology using LSTM network is developed to make a precise forecasting that is capable of handling more complex time series load data with long-term dependencies. The proposed technique outstrips other existing models. In conjunction with the above techniques, convolutional neural network (CNN) has been frequently used in the field of load prediction because of its excellent ability to capture the trend of load data [32]. In [33], a load forecasting methodology based on CNN model is introduced and compared with various artificial neural networks. From the outcomes, it is reported that CNN based method results in good accuracy in STLF. Moreover, for improving the forecasting accuracy, time-dependent convolutional neural network (TD-CNN) and cycle-based LSTM (C-LSTM) network have been implemented in the domain of STLF [34]. In [35], time-cognition CNN (TCMS-CNN) based multi-step STLF procedure is proposed, which significantly improves the prediction ability after extracting substantial and complex features from the electric load sequences. It results in precise outcomes in probabilistic forecasting that enables robust simplification in electricity market bidding and spot price calculation. An effective hybrid technique is implemented in [36] for STLF in smart cities via modified grasshopper optimization algorithm (MGOA) and locally weighted support vector regression (LWSVR). Here, LWSVR is used for getting stable performance and higher accuracy. In addition, MGOA is implemented after some modification in traditional GOA. Then, the performance of the developed LWSVR-MGOA is evaluated using different real word dataset.

In light of the above discussion, both LSTM and CNN have the capability to provide satisfactory results in short-term forecasting of load [37]. Therefore, it is logical to anticipate that integration of CNN and LSTM will further reduce the forecasting error. Based on the recent works, it is found that existing STLF approaches may not be always suitable due to abrupt change in load demand. Also, in Bangladesh power system, no fruitful method is yet implemented for forecasting short-term load. Therefore, to mitigate the current challenges and limitations of the existing techniques, further research is still required.

Note that authors in [35] only consider individual household load forecasting of a Smart Grid Smart City (SGSC) project. In contrast, the proposed algorithm of this paper can forecast the load of an entire power system of a country in different time horizon. To this end, this paper intends to make the following contributions.

- A hybrid methodology to forecast short-term electrical load is proposed. The methodology is based on the integration of CNN and LSTM network to take into account the advantages of both techniques.
- The developed methodology is applied to Bangladesh power system (BPS) for day-ahead and weak-ahead forecasting of load. Note that currently, short-term load forecasting is usually done via anecdotal evidences in BPS. Therefore, actual and forecasted demands may considerably mismatch that imposes additional challenges for network operators.

The outcome of this paper will be beneficial for utilities to perform short-term load forecasting more accurately. The rest of the paper is organized as follows. Section II describes the forecasting architecture. Then, Section III outlines the proposed methodology. Next, Section IV shows the steps to implement the proposed methodology Afterwards, Section V contains simulation results and analyses, and finally section VI concludes the paper.

II. FORECASTING ARCHITECTURE

A. LSTM ARCHITECTURE

A particular form of RNN network known as LSTM network, can store past data in its memory unit. It is very much effective for time series data prediction. A LSTM network contains four basic components viz. cell, input gate, output gate and forget gate. Information is transferred by the cell over random time intervals. The gates trace the flow of the input and output data from the cell. Gradient vanishing and exploding problems obtained from RNN can be mitigated using LSTM network. The basic configuration of a LSTM network is shown in Fig. 1. The nodal outputs of a LSTM network are computed as follows:

$$i_t = \sigma(W_i.\left[h_{t-1,}x_t\right] + b_i) \tag{1}$$

$$f_t = \sigma(W_f.\left[h_{t-1,}x_t\right] + b_f) \tag{2}$$

$$\widehat{C}_t = \tanh(W_c.\left[h_{t-1,}x_t\right] + b_c) \tag{3}$$

$$C_t = \mathrm{f_t} * C_{t-1} + i_t * \widehat{C}_t \tag{4}$$

$$o_t = \sigma(W_o.[h_{t-1}, x_t] + b_0) \tag{5}$$

$$h_t = o_t * \tanh(c_t) \tag{6}$$

where input variable at time t is denoted by x_t. Further, W_i, W_f, W_c, W_o are known as weight matrices, i_t stands for input gates; f_t, O_t and C_t represent forget gate, output gate, and cell output, respectively. The sigmoid activation function is specified by σ and h_t is assigned for calculating hidden state outputs at time t in vector structure. Besides, b_i, b_c, b_f and b_0 refer to the biased values of different gates.

B. CNN ARCHITECTURE

A special type of neural network called CNN is recently used to address STLF. This network is capable of handling long time series data and gives promising performance. It basically integrates two functions to form a third function using convolutional operation [32]. This operation can be computed as follows.

$$s = (x * w) \tag{7}$$

where input function is denoted by x, w stands for weighting function. Output of the convolution operation called feature map in two-dimensional axis can be represented as below.

$$s = (I * K)\,(i, j) = \sum\nolimits_m \sum\nolimits_n I\,(m, n)\,K(i - m, j - n) \tag{8}$$

A CNN comprises of three sections, namely convolutional layer, pooling layer and dense layer. Moreover, convolutional layer consists of three stages. In the first stage, convolution operation takes place with the help of linear activation. The second stage is called detector stage, which detects the activation function used in the first stage and rectified it into a linear function.Then, pooling operation is taken place in the third stage. To reduce the dimensionality of the time series data, a pooling function is used which shortens the training time of the network. It is also needed for down sampling the obtained feature map without any change of the depth. Maximum pooling operation of the pooling layer generally select the maximum values obtained from the convolution layer and launch it on the maximum pooling window. It differs from convolutional layers' parameter. However, various pooling operation such as average pooling and minimum pooling are mostly used when needed.

It can be summarized that values obtained from the convolutional layer is sent to the pooling layer for the purpose of the feature extraction. At the end of pooling stage, the output data are required to be flattened and forwarded to the dense layer. Eventually, it makes a one dimensional output sequence. This model used back propagation algorithm for extending the learning scheme of the network. Internal architecture of CNN is shown in Fig. 2.

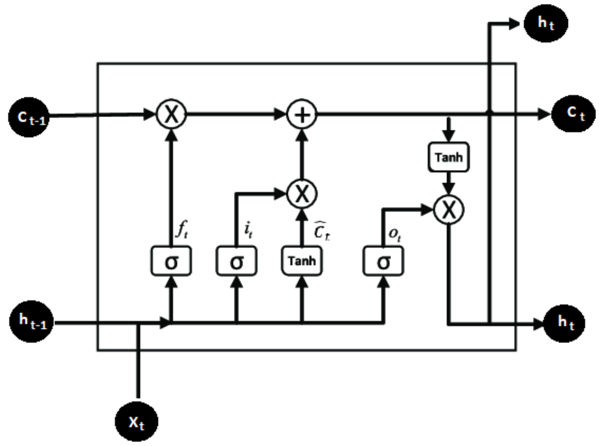

FIGURE 1. LSTM architecture, showing the three gates (input, output and forget).

C. HYBRID CNN-LSTM NETWORK

LSTM and CNN are both outlined to deliver a high precision forecasting. To this extent, it intends to implement a hybrid neural network framework that is able to extract and facilitate various hidden features of the load sequences to provide accurate load forecasting. CNN-LSTM hybrid network basically comprises of a convolution neural network module, a long short-term memory module and a feature-fusion module. A CNN-LSTM architecture is shown in Fig. 3. The original historical electric load dataset is reassigned into two different datasets. In addition, the CNN module is used to capture the local trend of the load data pattern. It also flattened down the samples into a single one-dimensional vector, which is used as a single input time step to the LSTM layer. Generally,

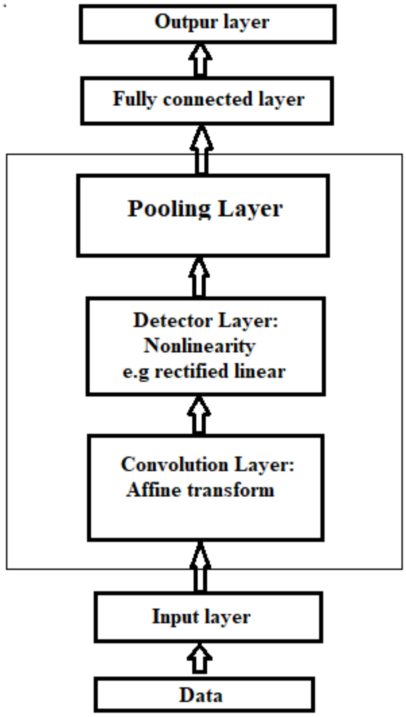

FIGURE 2. Internal architecture of CNN, showing the step by step procedure.

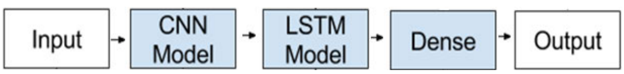

FIGURE 3. A model of CNN-LSTM, representing general block diagram of the network.

the LSTM module is designed to learn the long-term dependencies of the load dataset. The final forecasting is produced after a complete connection through dense layers.

III. PROPOSED METHODOLOGY

The steps of the proposed methodology are briefly described as follows.

A. STEP 1: DATA FRAMING

In this step, historical load data are collected from a particular region and null values are checked. The missing values need to be adjusted with the trend of the load pattern obtained from previous record. Thus, very good quality dataset can be obtained and it has very less impact on performance of forecasting. Then, a load dataset has to be divided into training and test sets for evaluating the proposed model. Collected dataset are split into different standard weeks. Reformation of this data frame is very much effective for defining the model, which can predict the power consumption for the week-ahead and month- ahead.

B. STEP 2: CONSTRUCTING MULTISTEP TIME SERIES

In the proposed model, electric load dataset has to be transformed in the shape of [sample, time steps, features]. At first, per sample, seven time steps are taken having one feature for total daily power consumption of seven days. The information of this pattern is not sufficient to train the network. Therefore, it is needed to create more training information by changing the problem to predict the next seven days given the prior seven days, irrespective of the standard week. Secondly, dataset is needed to be flattened at first and make eight-time series sequences. Then, it is obligatory to iterate over the time steps and divide the dataset into overlapping window, where it moves along one-time step and predicts the subsequent seven days. However, the test information from the data set remains the same in every case.

C. STEP 3: BUILDING FORECASTING MODEL

It is needed to implement an encoder-decoder CNN-LSTM model, which basically deals with the one-dimensional data in the three-dimensional pattern. The CNN block of the proposed model is defined via two convolutional layers, where convolution done with the help of the kernel filter. The first convolution layer reads the input series and projects its sequences on to the features windows. The second convolution layer is operated for amplifying the features obtained from the first layer. In the proposed model, the number of feature maps is 64 per each convolutional layer with three-time step kernel filter. A maximum pooling layer is generally used for getting the values after two times convolutions in the convolution layers. It is actually used for simplifying the input features. In the proposed model, maximum pooling operation has to be done by taking 1/4 of the values with the original sequence. The results obtained from this operation are then flattened into a long vector that is used as input to the decoding process of LSTM unit followed by a dense layer. This layer is used to provide the output. The developed model for load forecasting is shown in the Fig. 4.

D. STEP 4: TRAINING THE PROPOSED MODEL

Proposed CNN-LSTM architecture is built using Keras, an open-source neural network library, which is written in Python. Afterwards, the model needs to be tested with different unseen data set to check the general applicability of the model and enhance the performance. In the proposed framework, the network is trained with the following hyper parameters.

- Type of convolution: One dimensional
- No of filter: 64 with kernel size 3
- Activation: Rectified linear unit (RELU) for CNN, LSTM and Dense layer
- Optimizer: Adam
- No. of hidden layer: 200 for LSTM
- No. of training iterations (epochs): 20
- Batch size: 16

With the above-mentioned hyper parameters, the proposed model is verified from over-fitting problem by selecting low bias and low variance. Flow chart of the proposed technique is shown in Fig. 5.

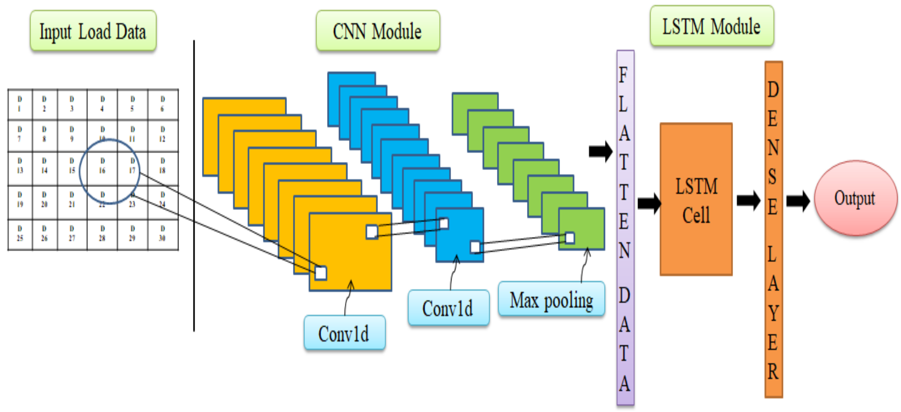

FIGURE 4. Proposed CNN-LSTM model; representing data selection mode, feature extraction mode from CNN module, decoding process from LSTM module.

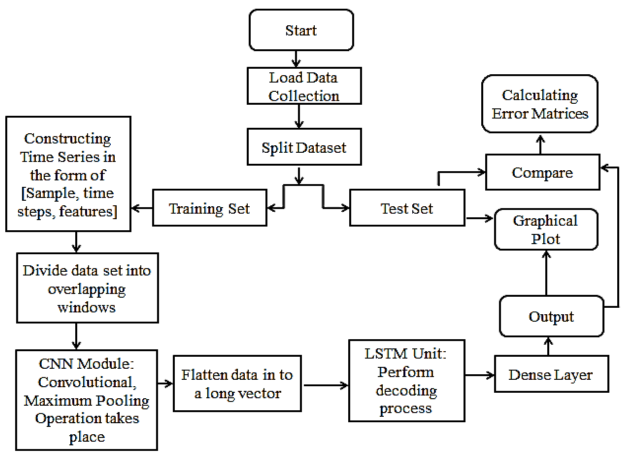

FIGURE 5. Flow chart of the proposed technique; data selection, partitioned process, training set, test set, forecasting outcome and validation.

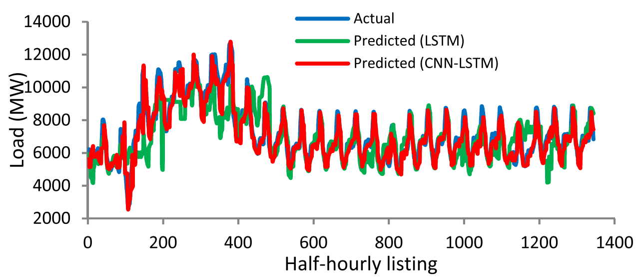

FIGURE 6. Comparison of load forecasting results of BPS using different networks for February 2019.

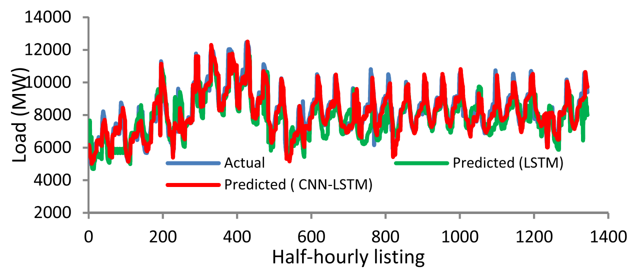

FIGURE 7. Comparison of load forecasting results of BPS using different networks for March 2019.

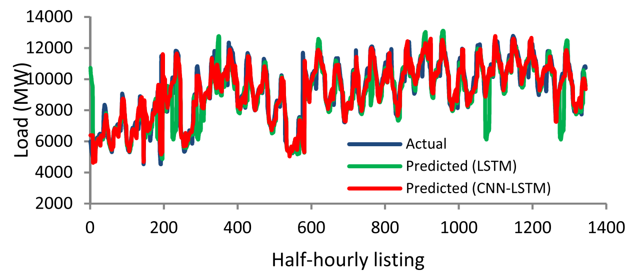

FIGURE 8. Comparison of load forecasting results of BPS using different networks for July 2019.

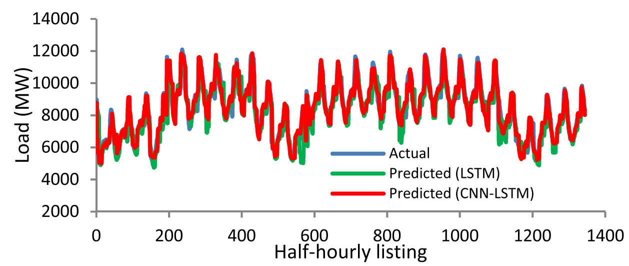

FIGURE 9. Comparison of load forecasting results of BPS using different networks for October 2019.

IV. IMPLEMENTATION OF PROPOSED METHODOLOGY

The following steps are followed while implementing the proposed methodology to forecast electrical load.

A. STEP 1

Historical time series half-hourly electric load data of BPS over six years (January 2014 – December 2019) are collected. For the training process of the model, data of the first five years (January 2014 – December 2018) are used. Then, sixth year data (January 2019 – December 2019) are utilized for evaluating the performance of the proposed model. For training the developed CNN-LSTM network, the collected data are divided into different segments with 7 days' interval from

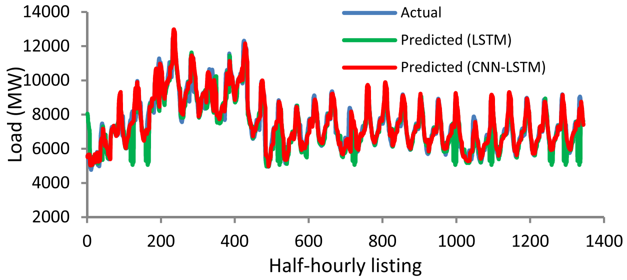

FIGURE 10. Comparison of load forecasting results of BPS using different networks for November 2019.

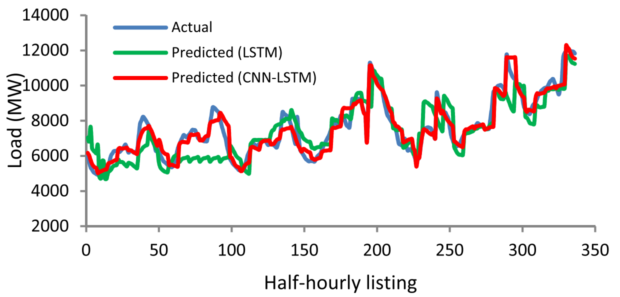

FIGURE 11. Comparison of load forecasting results of BPS using different networks for 01-07 March 2019.

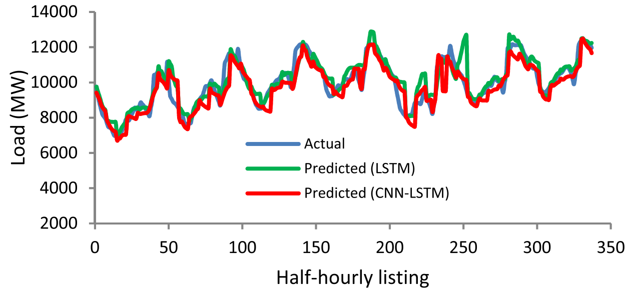

FIGURE 12. Comparison of load forecasting results of BPS using different networks for 15-21 July 2019.

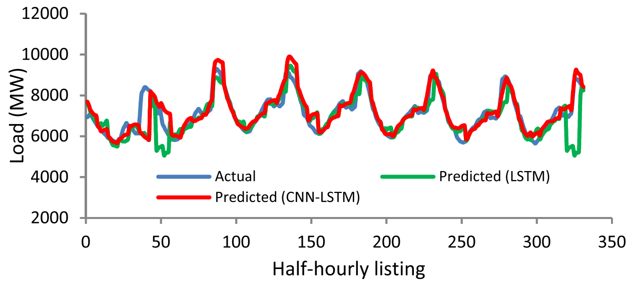

FIGURE 13. Comparison of load forecasting results of BPS using different networks for 15-21 November 2019.

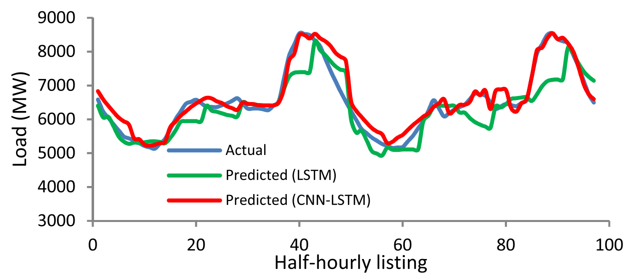

FIGURE 14. Comparison of load forecasting results of BPS using different networks for 15-16 January 2019.

very beginning to with a view to create more training features and avoid over-fitting problem. Data used for implementation of the proposed model is sufficient to learn the load dependencies. Thus, the model mitigates the under-fitting and over-fitting issues.

B. STEP 2

The training data set of seven days' interval is organized as: [sample, time-steps, feature], where no. of sample = 260, time steps = 7 and features = 8.

C. STEP 3

Processed data from step 2 is applied in the proposed CNNLSTM model.

D. STEP 4

Proposed model is then trained with the processed electric load data of BPS with the layer specification described in step 4 in the previous section.

V. RESULTS AND ANALYSES

A. FORECASTING OUTCOMES

The trained proposed CNN-LSTM model and LSTM model can predict load over different time horizons such as 24 hours, one week and one month at half hourly intervals. The time series load data from January 2014 to December 2019 contains a total of 105168 data points (where each point holds half hourly information). It is divided into 260 samples and 8 features with 7 time steps for producing more information to train the network.

First five years load dataset (i.e. from January 2014 to December 2018) is selected to train the proposed scheme. On the other hand, test procedure is performed with the dataset of 2019 (i.e. January 2019 - December 2019).

Forecasting results of BPS for various time horizons using the developed algorithm are shown in Fig. 6 to Fig. 19. Also, for comparison purpose, loads are predicted using LSTM model. It can be observed form Fig. 6 to Fig. 19 that both CNN-LSTM and LSTM model can predict the actual load demand trend. However, the forecasting outcomes obtained from CNN-LSTM are closer to the actual load patterns compared to that of LSTM technique.

B. EVALUATION METRICS

To validate the proposed methodology, its performances are compared to that of LSTM, RBFN and XGBoost approaches in terms of error metrics (also known as evaluation metrics). To this end, three types of error metrics such as Mean Average Error (MAE), Root Mean Squared Error (RMSE) and Mean

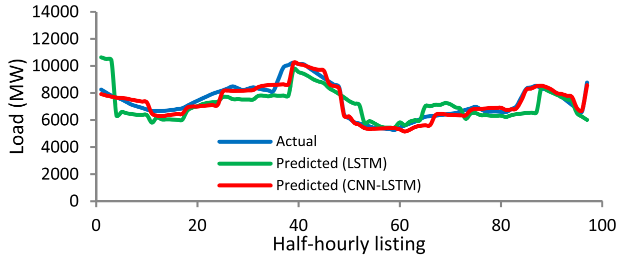

FIGURE 15. **Comparison of load forecasting results of BPS using different networks for 11-12 March 2019.**

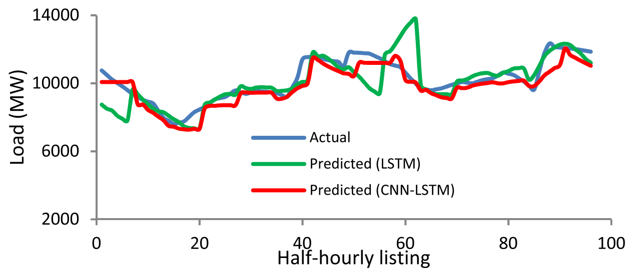

FIGURE 16. **Comparison of load forecasting results of BPS using different networks for 07-08 July 2019.**

Absolute Percentage Error (MAPE) are taken into account. A better forecasting is indicated by the relatively lower values of these metrics.

The mathematical expressions of the above error metrics are given as follows.

$$MAE = \frac{1}{N}\sum\nolimits_{L=1}^{N} |(F_L - Y_L)| \tag{9}$$

$$MAPE = \frac{\sum_{L=1}^{N}\left|\frac{(F_L - Y_L)}{Y_L}\right|}{N} \times 100 \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N}\sum\nolimits_{L=1}^{N} (F_L - Y_L)^2} \tag{11}$$

where N is the number of data points in forecasted load, F_L denotes the magnitude of forecasted load and Y_L stands for the magnitude of actual load at any instant.

In addition, another indicator named coefficient of determination (R^2) is considered for performance comparison. It can be calculated as follows.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(F_L - Y_L)^2}{\sum_{i-1}^{N}(F_L - A_L)^2} \tag{12}$$

where A_L indicates the mean value of the observations. Higher forecasting accuracy is obtained when the value of R^2 is closer to 1.

C. PERFORMANCE CALCULATION USING EVALUATION METRICS

By evaluating the MAE, RMSE, MAPE and R^2 values, the forecasting performances of various methods are compared. It is found that the proposed CNN-LSTM network outperforms LSTM, RBFN and XGBoost models in all time horizons.

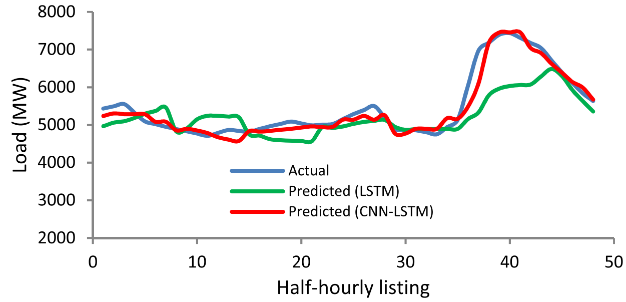

FIGURE 17. **Comparison of load forecasting results of BPS using different networks for 01 April 2019.**

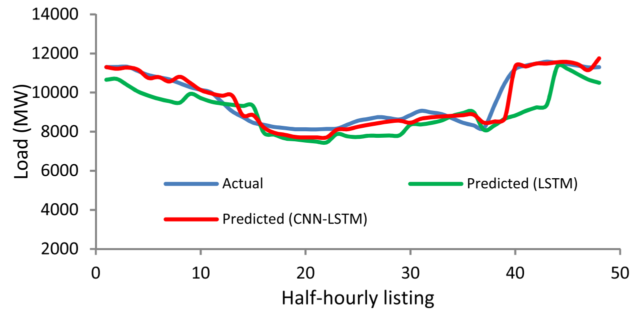

FIGURE 18. **Comparison of load forecasting results of BPS using different networks for 15 July 2019.**

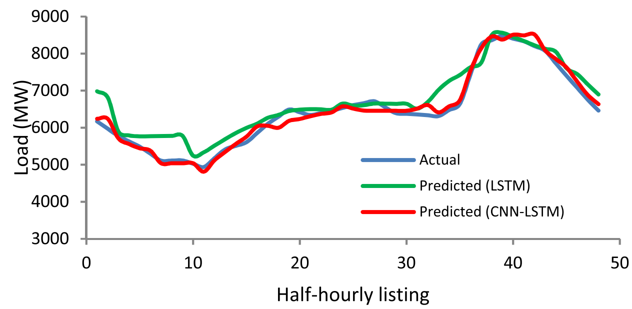

FIGURE 19. **Comparison of load forecasting results of BPS using different networks for 07 September 2019.**

Note that a computer with the following specifications is used to carry out simulations – Processor: Intel® core (TM) i7-8550(1.9 Hz), RAM: 8 GB. The run time for CNN-LSTM for 24 hour long forecast is approximately 25 minutes. In contrast, LSTM, RBFN, XGboost take 20, 17 and 13 minutes, respectively. Therefore, computational burden for CNN-LSTM is reasonable compared to other methods. Moreover, high configuration computer will take less time to compute the results. The detailed results are given below.

1) MONTHLY COMPARISON OF EVALUATION METRICS

The performance of the proposed CNN-LSTM network is compared with LSTM network, RBFN, XGBoost algorithm in terms of monthly MAE, RMSE, MAPE and R^2 in Table 1.

TABLE 1. Monthly MAE, RMSE, MAPE and R^2 for LSTM, CNN-LSTM, RBFN and XGBOOST.

Observation Period	MAE				RMSE				MAPE (%)				R^2			
	LSTM	CNN-LSTM	RBFN	XGboost	LSTM	CNN-LSTM	RBFN	XGboost	LSTM	CNN-LSTM	RBFN	XGboost	LSTM	CNN-LSTM	RBFN	XGboost
Jan 19	486.49	304.97	558.49	664.56	577.94	429.79	615.15	775.45	5.71	4.23	6.23	7.98	0.911	0.926	0.897	0.879
Feb 19	889.05	370.18	905.84	954.87	1260.03	538.77	818.23	1080.57	12.81	5.34	9.23	11.87	0.853	0.923	0.789	0.775
Mar 19	734.98	381.50	791.53	881.32	962.59	573.09	954.45	1276.53	9.76	4.73	8.25	10.56	0.854	0.927	0.795	0.775
Apr 19	664.96	402.77	705.43	779.90	1036.38	610.37	895.51	1236.56	8.70	4.88	7.79	10.54	0.885	0.927	0.803	0.790
May 19	692.62	471.63	801.71	850.87	1110.09	733.13	1020.19	1375.65	8.43	5.47	7.03	10.06	0.889	0.905	0.845	0.815
Jun 19	649.99	426.56	720.03	764.98	1122.29	736.35	1115.19	1453.65	8.05	4.94	9.04	11.05	0.890	0.908	0.864	0.817
July 19	559.02	489.24	590.41	687.85	859.10	645.65	947.79	1234.59	6.58	5.65	5.97	8.07	0.903	0.905	0.887	0.873
Aug 19	621.80	437.53	689.59	645.63	939.49	668.87	850.19	1123.74	7.81	5.31	7.89	9.39	0.894	0.904	0.871	0.852
Sep 19	614.79	438.94	650.69	692.58	929.80	669.94	1213.45	1523.47	7.71	5.35	7.91	9.50	0.896	0.904	0.870	0.853
Oct 19	476.99	291.16	555.12	509.36	713.112	406.45	790.19	995.69	6.09	3.49	6.85	8.01	0.909	0.960	0.810	0.870
Nov 19	474.16	298.62	535.39	538.79	821.09	464.95	805.27	1076.54	7.13	3.93	7.95	9.97	0.898	0.961	0.805	0.854
Dec 19	430.14	335.55	515.64	604.67	606.02	498.14	920.87	1123.34	6.13	4.72	5.81	8.87	0.909	0.927	0.798	0.870
Average	561.15	387.39	668.32	714.62	911.49	581.29	912.21	1457.76	7.91	4.84	7.49	9.66	0.891	0.92	0.836	0.835

TABLE 2. Weekly MAE, RMSE, MAPE and R^2 for LSTM, CNN-LSTM, RBFN and XGBOOST.

Observation Period	MAE				RMSE				MAPE (%)				R^2			
	LSTM	CNN-LSTM	RBFN	XGboost	LSTM	CNN-LSTM	RBFN	XGboost	LSTM	CNN-LSTM	RBFN	XGboost	LSTM	CNN-LSTM	RBFN	XGboost
01-07 Jan 19	567.99	119.03	690.45	762.52	803.53	331.91	970.34	1072.45	8.99	6.65	9.20	12.40	0.911	0.926	0.896	0.841
01-07 Feb 19	1133.49	494.83	1278.79	1352.67	1552.89	718.33	1780.76	1935.35	16.64	6.86	18.58	21.43	0.874	0.910	0.781	0.758
01-07 Mar19	727.63	318.64	968.65	1120.32	944.99	445.34	1132.54	1321.53	10.48	4.26	13.20	15.79	0.831	0.901	0.776	0.733
01-07 Apr 19	549.86	356.17	735.23	925.29	858.77	552.55	998.64	1147.57	8.12	5.26	10.48	13.25	0.889	0.903	0.777	0.735
01-07 May 19	657.59	358.08	878.54	990.37	1220.78	591.73	1390.94	1576.34	8.82	4.88	11.25	13.78	0.869	0.883	0.819	0.789
08-14 Jun 19	609.90	446.49	820.76	975.54	1049.54	882.45	1190.59	1360.46	7.56	5.55	9.75	12.72	0.861	0.887	0.841	0.797
15-22 July 19	523.35	433.68	725.25	910.38	716.04	577.57	860.37	987.37	5.52	4.14	7.42	9.78	0.883	0.879	0.853	0.843
23-29 Aug 19	652.43	422.81	868.85	982.65	968.67	607.37	1180.56	1372.85	7.13	4.28	9.35	12.10	0.864	0.884	0.852	0.839
01-07 Sep 19	728.19	365.88	969.98	1121.69	1155.09	548.76	1360.38	1580.38	9.92	4.92	11.24	13.75	0.872	0.873	0.842	0.827
08-14 Oct 19	482.92	286.49	658.91	827.57	774.84	394.61	934.86	1030.53	6.36	3.42	8.74	11.02	0.885	0.940	0.811	0.803
15-21 Nov 19	386.92	297.45	545.37	721.87	690.64	470.34	780.37	956.38	6.08	4.18	8.39	10.89	0.872	0.933	0.842	0.827
22-28 Dec 19	416.52	279.59	609.63	813.59	594.52	400.14	734.64	939.26	6.64	4.38	8.86	11.35	0.896	0.911	0.782	0.858
Average	619.73	348.26	812.53	958.705	944.19	543.42	1109.58	1273.37	8.52	4.89	10.54	13.19	0.875	0.902	0.823	0.805

For instance, with the forecasted results in January 2019, the LSTM network provides a MAE of 486.49, while the proposed CNN-LSTM offers a MAE of 304.97. For the month of February 2019, LSTM offers a MAE of 889.05,

TABLE 3. MAE, RMSE, MAPE and R^2 of 48 hours for LSTM, CNN-LSTM, RBFN and XGBOOST.

Observation Period	MAE				RMSE				MAPE (%)				R^2			
	LSTM	CNN-LSTM	RBFN	XGBoost	LSTM	CNN-LSTM	RBFN	XGBoost	LSTM	CNN-LSTM	RBFN	XGBoost	LSTM	CNN-LSTM	RBFN	XGboost
15-16 Jan 19	355.97	201.21	568.65	680.34	503.85	290.17	633.75	730.52	5.54	3.03	7.94	9.85	0.905	0.916	0.887	0.865
11-12 Mar 19	692.20	262.17	920.39	1050.38	910.44	360.40	1115.98	1387.43	9.58	3.69	11.32	13.79	0.845	0.913	0.769	0.725
06-07 May 19	848.83	347.19	1030.32	1215.54	1493.41	619.53	1730.76	1920.43	11.63	4.77	14.20	16.52	0.834	0.907	0.755	0.715
07-08 July 19	618.83	439.01	858.43	1080.76	976.78	575.28	1180.53	1378.48	6.19	4.53	8.17	10.59	0.872	0.918	0.788	0.772
04-05 Sep19	732.71	249.17	890.43	1134.64	1112.42	341.84	1356.86	1520.73	9.15	2.86	11.23	13.63	0.853	0.886	0.805	0.775
20-21 Nov 19	386.50	190.95	570.32	778.53	606.73	236.63	805.74	1030.54	5.72	2.71	7.87	9.72	0.881	0.918	0.864	0.827
Average	605.84	281.62	806.42	990.03	933.94	403.98	1137.27	1328.02	7.97	3.59	10.12	12.35	0.865	0.909	0.811	0.779

TABLE 4. MAE, RMSE, MAPE and R^2 of 24 hours for LSTM, CNN-LSTM, RBFN and XGBOOST.

Observation Period	MAE				RMSE				MAPE (%)				R^2			
	LSTM	CNN-LSTM	RBFN	XGBoost	LSTM	CNN-LSTM	RBFN	XGBoost	LSTM	CNN-LSTM	RBFN	XGBoost	LSTM	CNN-LSTM	RBFN	XGBoost
01 Jan 19	493.71	211.92	578.45	730.54	652.79	249.90	862.45	995.43	8.82	3.65	11.32	13.49	0.901	0.916	0.886	0.821
06 Feb 19	642.76	239.56	858.54	938.68	865.10	347.95	953.13	1162.54	6.67	2.31	8.78	9.92	0.864	0.930	0.792	0.768
02 Mar 19	793.11	230.04	920.86	1110.76	918.56	286.09	1130.38	1352.85	12.72	3.38	14.75	16.10	0.843	0.907	0.786	0.743
01 Apr 19	418.05	136.18	626.54	896.43	585.16	200.50	631.54	887.43	7.75	2.56	9.45	10.75	0.879	0.913	0.797	0.745
02 May 19	407.96	150.39	608.23	721.95	502.86	206.25	728.54	927.64	5.82	5.21	6.87	8.92	0.875	0.897	0.825	0.802
11 Jun 19	403.92	175.74	849.34	975.32	579.69	210.82	778.92	965.35	4.99	2.19	6.78	7.89	0.871	0.891	0.851	0.806
15 July 19	775.62	192.31	928.26	1078.54	954.05	258.19	1181.76	1372.87	8.72	2.31	10.78	11.78	0.893	0.885	0.867	0.851
23 Aug 19	775.62	265.20	975.47	1136.49	954.05	394.45	1198.52	1420.74	8.72	2.99	11.80	12.86	0.874	0.894	0.862	0.849
07 Sep 19	878.13	353.07	1048.53	1248.54	1287.83	353.07	1368.53	1580.63	12.29	3.99	13.49	15.70	0.883	0.887	0.850	0.833
08 Oct 19	480.82	335.58	620.65	735.76	636.21	435.76	821.53	1133.84	5.12	3.62	6.97	8.68	0.895	0.930	0.831	0.810
18 Nov 19	312.11	252.61	678.93	819.64	406.64	370.54	591.84	732.68	4.16	3.37	6.12	7.87	0.882	0.943	0.853	0.832
23 Dec 19	300.40	242.38	528.54	649.32	377.38	192.81	623.96	872.63	4.76	3.03	6.52	8.23	0.903	0.921	0.791	0.865
Average	556.85	232.08	768.5	920.16	726.69	292.19	905.93	1117.05	7.55	3.22	9.46	11.02	0.880	0.909	0.833	0.810

where MAE obtained from CNN-LSTM is 370.18. Similarly, the proposed method results in less MAE with the data forecasted data in October 2019. The similar trend is also evident for other months. On average, the CNN-LSTM provides 173.76, 330.2 and 3.07% less MAE, RMSE and MAPE respectively compared to LSTM network. The improvements are graphically illustrated in Fig. 20 to Fig. 22. Also, R^2 values are higher for CNN-LSTM compared to other methods. Therefore, CNN-LSTM is found to be more effective in short-term electrical load forecasting.

2) WEEKLY COMPARISON OF EVALUATION METRICS

The weekly comparison of MAE, RMSE, MAPE and R^2 among the proposed CNN-LSTM, LSTM, RBFN and XGBoost algorithm is presented in Table 2. For example, with the forecasted data in 01-07 Januray 2019, the LSTM network provides a RMSE of 803.53, while the proposed

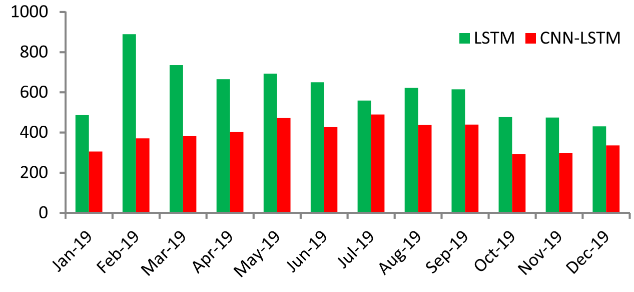

FIGURE 20. Comparison of MAE obtained from LSTM and proposed CNN-LSTM for 30 days' prediction.

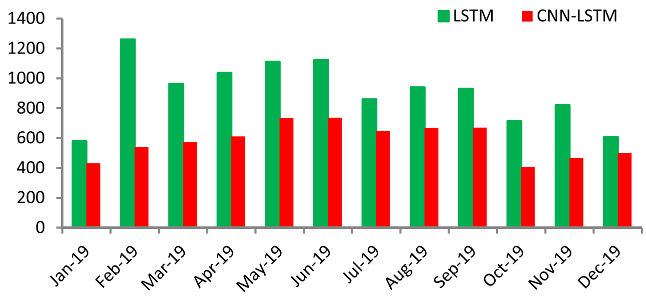

FIGURE 21. Comparison of RMSE obtained from LSTM and proposed CNN-LSTM for 30 days' prediction.

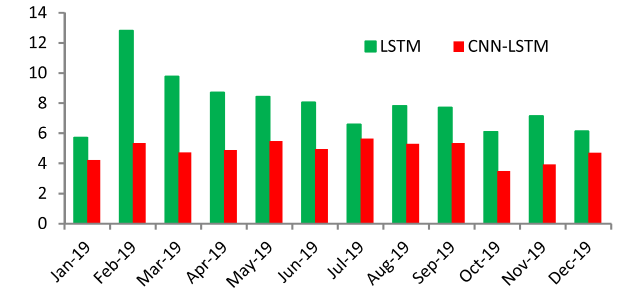

FIGURE 22. Comparison of MAPE obtained from LSTM and proposed CNN-LSTM for 30 days' prediction.

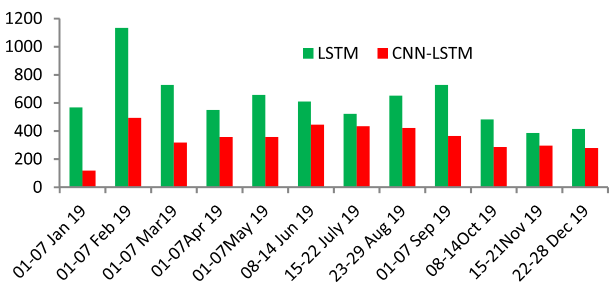

FIGURE 23. Comparison of MAE obtained from LSTM and proposed CNN-LSTM for 07 days' prediction.

CNN- LSTM offers a RMSE of 331.91. Similarly, the proposed method offers less RMSE in 08-14 October 2019. The similar trend is also noticed in other weeks. On average, the proposed method provides 271, 400.77, and 3.63% less MAE, RMSE and MAPE respectively than the LSTM network. The improvements of all metrics are depicted in Fig. 23 to Fig. 25.

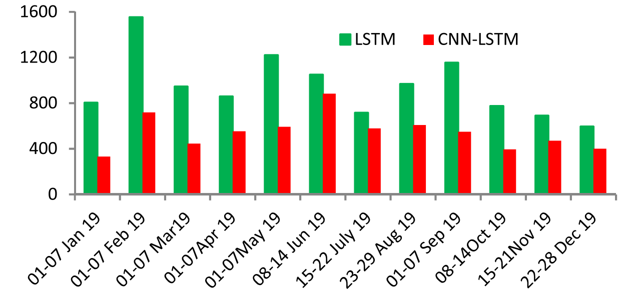

FIGURE 24. Comparison of RMSE obtained from LSTM and proposed CNN-LSTM for 07 days' prediction.

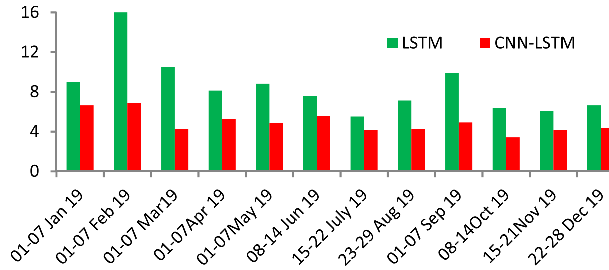

FIGURE 25. Comparison of MAPE obtained from LSTM and proposed CNN-LSTM for 07 days' prediction.

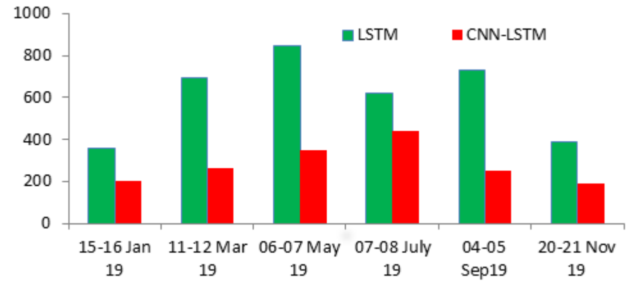

FIGURE 26. Comparison of MAE obtained from LSTM and proposed CNN-LSTM for 02 days.

3) 48 HOURS COMPARISON OF EVALUATION METRICS

Table 3 demonstrates the selected 48 hours' comparison of MAE, RMSE, MAPE and R^2 for the proposed CNN-LSTM, LSTM, RBFN, XGBoost algorithms. For instance, with the forecasted data in 15-16 January 2019, the LSTM network provides RMSE of 503.85, while the proposed CNN-LSTM offers a RMSE of 290.17. Similarly, the proposed method offers less MAPE in 04-05 September 2019. The similar trend is also revealed in other forecasting results. On average, the proposed method provides 324.22, 529.96 and 4.37% less MAE, RMSE and MAPE respectively than LSTM network. The improvements of all metrics are shown in Fig. 26 to Fig. 28.

4) COMPARISON OF EVALUATION MATRICES FOR 24 HOURS

Daily comparison of MAE, RMSE, MAPE and R^2 for the proposed CNN-LSTM, LSTM, RBFN and XGBoost algorithm network are shown in Table 4. To give an example,

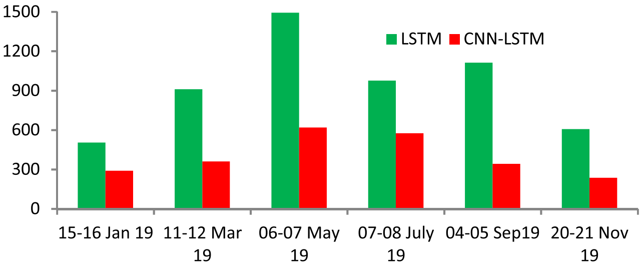

FIGURE 27. Comparison of RMSE obtained from LSTM and proposed CNN-LSTM for 02 days.

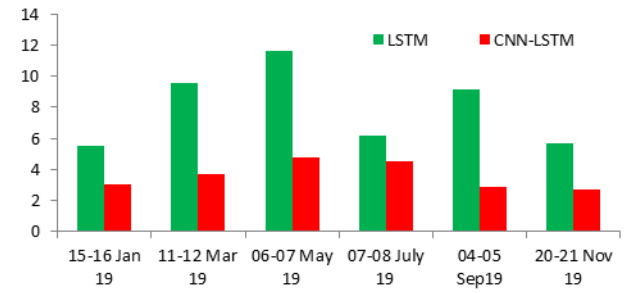

FIGURE 28. Comparison of MAPE obtained from LSTM and proposed CNN-LSTM for 02 days.

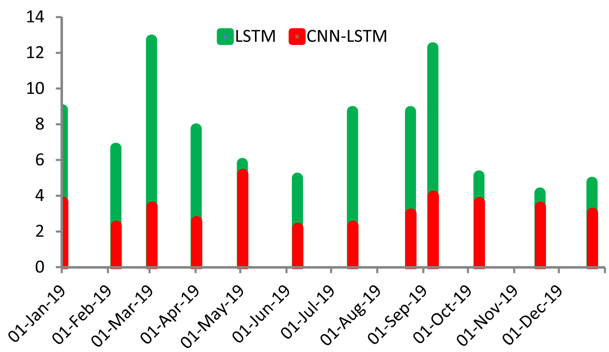

FIGURE 29. Comparison of MAE obtained from LSTM and proposed CNN-LSTM for 24 hours.

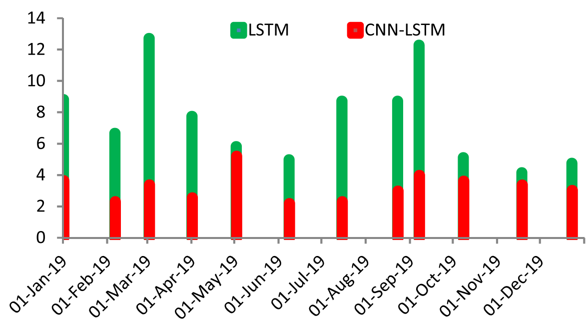

FIGURE 30. Comparison of RMSE obtained from LSTM and proposed CNN-LSTM for 24 hours.

with the forecasted data in 01 January 2019, the LSTM network provides an RMSE of 652.79, whereas the proposed CNN-LSTM offers a RMSE of 249.90. Similarly, the proposed method offers less RMSE in 07 September 2019.

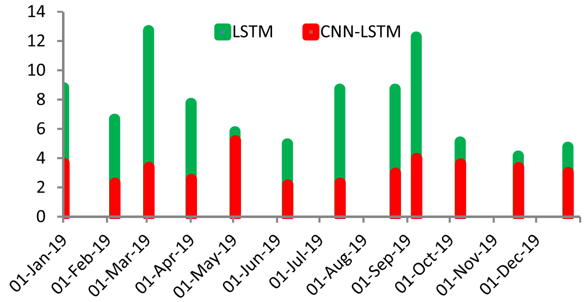

FIGURE 31. Comparison of MAPE obtained from LSTM and proposed CNN-LSTM for 24 hours.

The similar trend is also observed in other cases. On average, the proposed method provides 324.7693, 434.4985 and 4.3275% less MAE, RMSE and MAPE respectively than the LSTM network. The results are graphically shown in Fig. 29 to Fig. 31.

VI. CONCLUSION

This research work proposes an approach using CNN-LSTM model for short-term electrical load forecasting. The proposed methodology is implemented with the hidden features of CNN and LSTM networks to acquire the advantages of both modules. The performance of the developed model is evaluated by investigating the electrical load forecasting of Bangladesh power system. To ensure the stability and effectiveness, various segments of the dataset are executed accordingly. Furthermore, the robustness and general applicability of the developed network are tested.

It can be revealed from simulations that the proposed CNN-LSTM model gives the lowest values of MAE, RMSE and MAPE compared to LSTM, RBFN and XGboost models. Also, the R^2 values of CNN-LSTM are the highest among four models. In all validation cases, the forecasting outcomes using CNN-LSTM outperform LSTM, RBFN and XGBoost algorithms. Towards the end, it can be illustrated that the proposed CNN-LSTM model can deal with the long sequence time series electric load data and predict the future load demand over a considerable period. In the future work, highly precise load forecasting framework can be developed by using GRU integrated with CNN network.

REFERENCES

[1] Z. Yu, Z. Niu, W. Tang, and Q. Wu, "Deep learning for daily peak load forecasting–A novel gated recurrent neural network combining dynamic time warping," *IEEE Access*, vol. 7, pp. 17184–17194, 2019.

[2] M. Jacob, C. Neves, and D. V. Greetham, "Short-term load forecasting," in *Forecasting and Assessing Risk of Individual Electricity Peaks* (Mathematics of Planet Earth). Cham, Switzerland: Springer, 2020, pp. 5–37.

[3] W. Christiaanse, "Short-term load forecasting using general exponential smoothing," *IEEE Trans. Power App. Syst.*, vol. PAS-90, no. 2, pp. 900–911, Mar. 1971.

[4] W. Charytoniuk, M. S. Chen, and P. Van Olinda, "Nonparametric regression based short-term load forecasting," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 725–730, May 1998.

[5] K.-B. Song, Y.-S. Baek, D. H. Hong, and G. Jang, "Short-term load forecasting for the holidays using fuzzy linear regression method," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 96–101, Feb. 2005.
[6] C.-M. Lee and C.-N. Ko, "Short-term load forecasting using lifting scheme and ARIMA models," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 5902–5911, May 2011.
[7] J. Kim, S. Cho, K. Ko, and R. R. Rao, "Short-term electric load prediction using multiple linear regression method," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids (SmartGridComm)*, Aalborg, Denmark, Oct. 2018, pp. 1–6.
[8] J. W. Taylor, "Short-term electricity demand forecasting using double seasonal exponential smoothing," *J. Oper. Res. Soc.*, vol. 54, no. 8, pp. 799–805, Aug. 2003.
[9] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, and H. Zareipour, "Energy forecasting: A review and outlook," *IEEE Open Access J. Power Energy*, vol. 7, pp. 376–388, 2020.
[10] A. Jain and B. Satish, "Clustering based short term load forecasting using support vector machines," in *Proc. IEEE Bucharest PowerTech*, Bucharest, Romania, Jun./Jul. 2009, pp. 1–8.
[11] Y. Chen, P. Xu, Y. Chu, W. Li, Y. Wu, L. Ni, Y. Bao, and K. Wang, "Short-term electrical load forecasting using the support vector regression (SVR) model to calculate the demand response baseline for office buildings," *Appl. Energy*, vol. 195, pp. 659–670, Jun. 2017.
[12] H. H. Çevik and M. Çunkaş, "Short-term load forecasting using fuzzy logic and ANFIS," *Neural Comput. Appl.*, vol. 26, no. 6, pp. 1355–1367, Aug. 2015.
[13] K. B. Sahay and M. M. Triapthi, "Day ahead hourly load and price forecast in ISO New England market using ANN," in *Proc. IEEE INDICON*, Mumbai, India, Dec. 2013, pp. 1–6.
[14] B. S. Bisht and R. M. Holmukhe, "Electricity load forecasting by artificial neural network model using weather data," *IJEET Trans. Power Syst.*, vol. 4, no. 1, pp. 91–99, 2013.
[15] E. Banda and K. A. Folly, "Short term load forecasting using artificial neural network," in *Proc. IEEE Lausanne Power Tech*, Shimla, India, Jul. 2007, pp. 1–5.
[16] P.-H. Kuo and C.-J. Huang, "A high precision artificial neural networks model for short-term energy load forecasting," *Energies*, vol. 11, no. 1, p. 213, Jan. 2018.
[17] W.-Y. Chang, "Short-term load forecasting using radial basis function neural network," *J. Comput. Commun.*, vol. 03, no. 11, pp. 40–45, 2015.
[18] L. Ekonomou, C. A. Christodoulou, and V. A. Mladenov, "A short-term load forecasting method using artificial neural networks and wavelet analysis," *Int. J. Power Syst.*, vol. 1, pp. 64–68, 2016.
[19] S. Motepe, A. N. Hasan, and R. Stopforth, "Improving load forecasting process for a power distribution network using hybrid AI and deep learning algorithms," *IEEE Access*, vol. 7, pp. 82584–82598, 2019.
[20] J. Che and J. Wang, "Short-term load forecasting using a kernel-based support vector regression combination model," *Appl. Energy*, vol. 132, pp. 602–609, Nov. 2014.
[21] P. Ganguly, A. Kalam, and A. Zayegh, "Short-term load forecasting using fuzzy logic," in *Proc. Int. Conf. Res. Educ. Sci.*, Ankara, Turkey, 2017, pp. 1–5.
[22] L. Ekonomou, C. A. Christodoulou, and V. A. Mladenov, "Short-term load forecasting method using artificialneural Netw. wavelet analysis," *Int. J. Power Syst.*, vol. 1, pp. 64–68, Jul. 2016.
[23] G. Merkel, R. Povinelli, and R. Brown, "Short-term load forecasting of natural gas with deep neural network regression," *Energies*, vol. 11, no. 8, p. 2008, Aug. 2018.
[24] S. Bouktif, A. Fiaz, A. Ouni, and M. Serhani, "Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches," *Energies*, vol. 11, no. 7, p. 1636, Jun. 2018.
[25] A. Ahmadi, M. Nabipour, B. Mohammadi-Ivatloo, A. M. Amani, S. Rho, and M. J. Piran, "Long-term wind power forecasting using tree-based learning algorithms," *IEEE Access*, vol. 8, pp. 151511–151522, 2020.
[26] H. Shi, M. Xu, and R. Li, "Deep learning for household load forecasting—A novel pooling deep RNN," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5271–5280, Sep. 2018.
[27] L.-Y. Wei, "A study of the hybrid recurrent neural network model for electricity loads forecasting," *Int. J. Acad. Res. Accounting, Finance Manage. Sci.*, vol. 7, no. 2, pp. 21–29, Apr. 2017.
[28] L. Kuan, Z. Yan, W. Xin, C. Yan, P. Xiangkun, S. Wenxue, J. Zhe, Z. Yong, X. Nan, and Z. Xin, "Short-term electricity load forecasting method based on multilayered self-normalizing GRU network," in *Proc. IEEE Conf. Energy Internet Energy Syst. Integr. (EI2)*, Beijing, China, Nov. 2017, pp. 1–5.
[29] Y. Wang, M. Liu, Z. Bao, and S. Zhang, "Short-term load forecasting with multi-source data using gated recurrent unit neural networks," *Energies*, vol. 11, no. 5, p. 1138, May 2018.
[30] S. Muzaffar and A. Afshari, "Short-term load forecasts using LSTM networks," in *Proc. 10th Int. Conf. Appl. Energy*, Beijing, China, 2018, pp. 2922–2927.
[31] M. S. Hossain and H. Mahmood, "Short-term load forecasting using an LSTM neural network," in *Proc. IEEE Power Energy Conf. Illinois (PECI)*, Champaign, IL, USA Feb. 2020, pp. 1–6.
[32] S. H. Rafi and Nahid-Al-Masood, "Short term electric load forecasting using high precision convolutional neural network," in *Proc. Int. Conf. Comput., Electr. Commun. Eng. (ICCECE)*, Kolkata, India, Jan. 2020, pp. 1–7.
[33] K. Amarasinghe, D. L. Marino, and M. Manic, "Deep neural networks for energy load forecasting," in *Proc. IEEE 26th Int. Symp. Ind. Electron. (ISIE)*, Edinburgh, U.K., Jun. 2017, pp. 1483–1488.
[34] L. Han, Y. Peng, Y. Li, B. Yong, Q. Zhou, and L. Shu, "Enhanced deep networks for short-term and medium-term load forecasting," *IEEE Access*, vol. 7, pp. 4045–4055, 2018.
[35] Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu, and Z. Zhu, "Multi-scale convolutional neural network with time-cognition for multi-step short-term load forecasting," *IEEE Access*, vol. 7, pp. 88058–88071, 2019.
[36] E. E. Elattar, N. A. Sabiha, M. Alsharef, M. K. Metwaly, A. M. Abd-Elhady, and I. B. Taha, "Short term electric load forecasting using hybrid algorithm for smart cities," *Appl. Intell.*, vol. 50, pp. 3379–3399, Oct. 2020.
[37] M. Alhussein, K. Aurangzeb, and S. I. Haider, "Hybrid CNN-LSTM model for short-term individual household load forecasting," *IEEE Access*, vol. 8, pp. 180544–180557, 2020.

SHAFIUL HASAN RAFI received the B.Sc. degree in electrical and electronic engineering (EEE) from the Rajshahi University of Engineering and Technology (RUET), Rajshahi, Bangladesh, in 2014, and the M.Sc. degree in electrical electronic and communication engineering from the Military Institute of Science and Technology (MIST), Dhaka, Bangladesh, in 2020. He is currently an Assistant Professor with the School of Electrical, Electronic and Communication Engineering, MIST. His research interests include power system analysis and renewable energy.

NAHID-AL-MASOOD (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical and electronic engineering (EEE) from the Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 2008 and 2010, respectively, and Ph.D. degree in power and energy systems from The University of Queensland, Brisbane, QLD, Australia, in 2017. He is currently working as an Associate Professor with the Department of Electrical and Electronic Engineering (EEE), BUET. His research interests include power system modeling and analysis, and grid integration of renewable energy.

SHOHANA RAHMAN DEEBA (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical and electronic engineering (EEE) from the Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 2009 and 2013, respectively, and the Ph.D. degree in power and energy systems from The University of Queensland, Brisbane, QLD, Australia, in 2017. She is currently working as an Assistant Professor with the Department of Electrical and Computer Engineering, North South University, Dhaka. Her research interests include power system modeling, distribution system analysis, and renewable energy integration.

EKLAS HOSSAIN (Senior Member, IEEE) received the B.S. degree in electrical and electronic engineering from the Khulna University of Engineering and Technology, Bangladesh, in 2006, the M.S. degree in mechatronics and robotics engineering from the International Islamic University of Malaysia, Malaysia, in 2010, and the Ph.D. degree from the College of Engineering and Applied Science, University of Wisconsin Milwaukee (UWM). He is working as an Associate Researcher with the Oregon Renewable Energy Center (OREC). He has been working in the area of distributed power systems and renewable energy integration for last ten years. He is currently involved with several research projects on renewable energy and grid tied microgrid system with Oregon Tech, as an Assistant Professor with the Department of Electrical Engineering and Renewable Energy, since 2015. He has published a number of research articles and posters in this field. His research interests include modeling, analysis, design, and control of power electronic devices, energy storage systems, renewable energy sources, integration of distributed generation systems, microgrid and smart grid applications, robotics, and advanced control systems. He is a Senior Member of the Association of Energy Engineers (AEE). He is a registered Professional Engineer (PE) in the state of Oregon, USA. He is also a Certified Energy Manager (CEM) and Renewable Energy Professional (REP). He is a Winner of the Rising Faculty Scholar Awardfrom the Oregon Institute of Technology, in 2019, for his outstanding contribution in teaching. His dedicated research team, is looking forward to explore methods to make the electric power systems more sustainable, cost-effective, and secure, through extensive research and analysis on energy storage, microgrid systems, and renewable energy sources. He currently serves as an Associate Editor for IEEE Access.