IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 5, 2019, accepted December 18, 2019, date of publication December 30, 2019, date of current version July 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2963045

Short-Term Prediction of Residential Power Energy Consumption via CNN and Multi-Layer Bi-Directional LSTM Networks

FATH U MIN ULLAH, (Student Member, IEEE), AMIN ULLAH, (Student Member, IEEE), IJAZ UL HAQ, (Student Member, IEEE), SEUNGMIN RHO, (Member, IEEE), AND SUNG WOOK BAIK, (Member, IEEE)

Intelligent Media Laboratory, Digital Contents Research Institute, Sejong University, Seoul 143-747, South Korea

Corresponding author: Sung Wook Baik (sbaik@sejong.ac.kr)

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019M3F2A1073179).

ABSTRACT Excessive Power Consumption (PC) and demand for power is increasing on a daily basis, due to advancements in technology, the rise in electricity-dependent machinery, and the growth of the human population. It has become necessary to predict PC in order to improve power management and co-operation between the energy used in a building and the power grid. State-of-the-art Energy Consumption Prediction (ECP) methods are limited in terms of predicting the energy effectively, due to various challenges such as weather conditions and the dynamic behaviour of occupants. Thus, to overcome the drawbacks of these methods, we present an intelligent hybrid technique that combines a Convolutional Neural Network (CNN) with a Multi-layer Bi-directional Long-short Term Memory (M-BDLSTM) method using three steps. When applied to short-term power ECP, this approach helps to provide efficient power management i.e. it can assist the supplier to produce the optimum amount of power. The first step in our proposed method integrates the pre-processing and data organisation mechanisms to refine the data and remove abnormalities. The second step employs a deep learning network, where the sequence of refined data is fed into the CNN via the M-BDLSTM network to learn the sequence pattern effectively. The third step generates the ECP/PC by comparing actual and predicted data series and evaluates the prediction using error metrics. The proposed method achieves better prediction results than existing techniques, thus demonstrating its effectiveness. Furthermore, it achieved the smallest value of the Mean Square Error (MSE) and Root Mean Square Error (RMSE) for individual household dataset using 10-fold Cross Validation (CV) and a hold-out (CV) method.

INDEX TERMS Artificial intelligence, deep learning, power consumption, CNN, bi-directional LSTM, and short-term energy consumption.

I. INTRODUCTION

The demand for power is currently expanding on a worldwide scale, due to the widespread use of transportation such as engine-powered vehicles, large-scale machinery, and electronic market trading. This demand for power is a vital aspect of modern day life that can be fulfilled through the networked control of microgrids [1]. Similarly, the continuing expansion of Smart Meter Infrastructure (SMI) over worldwide has laid the groundwork for the introduction of active power energy systems into smart grids [2]. This deployment has created opportunities for short-term energy forecasting, and it is very important to apply for green atmosphere, and for efficiency of use, particularly for household consumers [3]. The demand for power energy is affected by consumers' behaviour and their use of electrical appliances. Therefore, the power grid authorities are considering the need to develop and invent new ways of efficiently handling the consumption of power in industrial and residential buildings that will regulate energy demand. In addition, although smart residential buildings can offer occupants facilities such as operating different electronic devices remotely via mobile apps, these devices

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

TABLE 1. Energy balance relating to overall electricity consumption per capita in South Korea [5].

Electricity	Overall	Per capita	Compared to Europe per capita
Consumption	507.6bn kwh	9862.7 kwh	5514.7 kwh
Production	526.0bn kwh	10220.3 kwh	5929.2 kwh

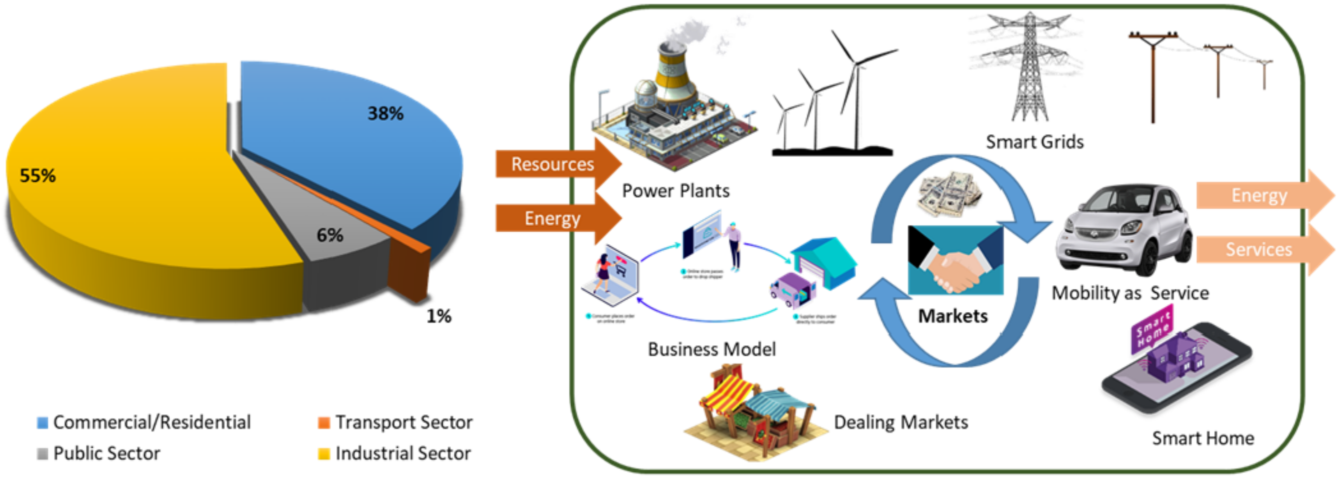

FIGURE 1. Energy consumption in South Korea on the annual basis illustrates that the large amount of energy is mostly consumed in industrial sector because in industrial sector different heavy machineries are installed for production in the form of electric, light, and water energy. The power energy consumption from resources to industries, business markets, transport and smart homes should be synchronized by strong cooperation. This flow is better shown in the figure.

consume relatively high amounts of energy due to the sensors that are required. Home appliances are often used improperly due to this sort of mismanagement, and a large amount of energy is wasted annually. Reducing this wastage of energy is very important in order to conserve it for future use by efficient prediction of consumption. Some energy prediction algorithms are used in the power management sector to predict demand over the next years, months, weeks, days, and hours, in order to produce the required amount of energy. However, there are certain factors related to the building structure that can influence energy use, such as climate level conditioning, the materials used in construction, and sub-level systems for heating, lighting, and ventilation [4]. With this in mind, management on the demand side can guide energy consumers in a procedural way by regulating the load due to devices or the behaviour of the occupants, allowing electricity consumption to be subsidized [6]. Commercial and residential buildings account for 30% to 40% of the entire energy consumed in a smart city [7]. Current trends illustrate that this percentage may increase in the near future, and energy consumption and penetration is increasing across the globe [8]. Forecasting this energy is essential and has become a challenging problem that depends on the complexity of the building's infrastructure behaviour and various uncertainties.

The management and preservation of energy in the power sectors are key phenomena in the literature associated with South Korea [9]. Residential locations in South Korea have been getting smarter, owing to the use of remote devices for communication, and the demand for energy in cities is growing rapidly day by day. According to statistics [10], energy consumption in the commercial and residential, transport, public and industrial sectors in South Korea is 38%, 1%, 6%, and 55%, respectively. These figures are illustrated in Fig. 1, where the energy consumption cycle between consumers and sources is also shown. Furthermore, South Korea can fully provide itself with self-produced energy and the overall production of electrical energy is currently 526bn kWh, representing 104% of the country's requirements [5]. The total energy balance in South Korea is given in Table 1, which shows the total energy consumption and production and a comparative analysis with Europe on a per capita basis.

The prediction of power usage is essential in order to allow for efficient use and deployment. In view of this, numerous prediction methods have been developed by researchers to improve the quality of the power grid and optimise energy usage. They have used historical data [11] in conjunction with machine learning and deep learning algorithms such as an Artificial Neural Network (ANN) [12], Support Vector Machine (SVM) [13], Adaptive Neuro-fuzzy Inference System (ANFIS) [14], and Extreme Learning Machine (ELM) [15] for prediction. The following sub-sections describe statistical machine learning and deep-learning-based methods in more detail.

A. STATISTICAL MACHINE LEARNING METHODS

In the past few years, statistical methods have remained popular for modelling PC and ECP. For instance, Wang *et al.* [16] predicted energy usage over a very short duration for two educational buildings in Florida, using a Random Forest (RF) approach to predict the hourly electricity consumption. Similarly, Candanedo *et al.* [17] performed several experiments on the prediction of energy, using Multiple Linear Regression (MLR), SVM, RF, and a Gradient Boosting Machine (GBM), and confirmed that GBM worked better than all other methods. Some researchers have also suggested that energy usage can also be affected by individual behaviour. For instance, Wang and Ding [18] developed a model based on the occupants' behaviour, and analysed the relationship between the behaviour and devices that consumed energy. Zhong *et al.* [19] used Support Vector Regression (SVR) based on a vector field to accurately predict the energy consumption in an office located in China. An approach used in [12] developed simple and multiple regression, including quadratic regression, to predict the hourly and daily energy consumption in residential buildings. Moreover, Cai *et al.* [20] accurately classified the energy consumption ratings for 16,000 residential houses on the basis of data collected from the entire region. They summarised the electricity usage patterns using data mining techniques and performed clustering using the k-means algorithm, where the electricity usage was divided using the centres of the obtained clusters, and applied SVM for classification.

B. DEEP LEARNING METHODS

Most research communities are using deep-learning-based methods due to their excellent results and applications in the field of computer vision and energy consumption. Recent attempts have given favourable results; for instance, Fayaz *et al.* [21] predicted ECP over the short term in residential buildings using a Feedforward and Backpropagation Neural Network (FFBPNN) on the normalised data for ECP. Similarly, Liu *et al.* [22] proposed a Lyapunov-based Decentralised Control (LBC) method for power and voltage regulation. Using this procedure, they obtained better stability on the power system. Tian *et al.* [23] used Generative Adversarial Networks (GANs) for a parallel prediction scheme for energy consumption. This scheme used a limited number of time series data to produce parallel data using a GAN, and then generated mixed data consisting of original and artificial data. The generated mixed data were used for the ECP model. Li *et al.* [1] proposed a dual-layered network with optimum power for an islanded microgrid. Almalaq and Zhang [24] proposed a hybrid approach based on deep learning by combining the genetic algorithm with LSTM for energy prediction, while Fayaz and Kim [25] improved the FFBPNN method to predict energy consumption using a Deep Extreme Learning Machine (DELM) and ANFIS. In addition to these approaches, Cui *et al.* [26] modelled the load among the distributed network via voltage disturbance. Another method presented in [27] used a deep recurrent neural network with LSTM for the forecasting of aggregated power load and photovoltaic power, and then applied a particle swarm algorithm to optimise load dispatched by the connected grid. Kim and Cho [28] used a hybrid approach combining a CNN with LSTM for short-term ECP. Likewise, Li *et al.* [29] developed an evolutionary algorithm called Teaching-learning-based Optimisation (TLBO) to predict short-term residential energy consumption. They further modified the algorithm by combining it with an ANN to give further improvement. To forecast energy use in a public building, Ruiz *et al.* [30] used an Elman neural network and optimised the model weights using the genetic algorithm.

The approaches mentioned above provide rough directions and features for ECP or PC, such as weather information and power data information that are used to calculate power energy demand through efficient prediction. These methods are also less proficient in the prediction of PC, due to their disorganised data usage. We can observe from the literature discussed above that the majority of prediction methods are based on statistical features and utilized machine learning techniques for ECP and PC. To overcome these problems and challenges, we introduce an effective deep-learning-based method, due to its high accuracy and performance in the areas of image processing [31], energy and power systems [32], [33] and computer vision such as smoke detection [34], [35], action recognition [36], [37], violence detection [38], and video summarisation [39], [40]. The key contributions of this work are as follows:

1. Sometimes, due to weather conditions and the occupants' behaviour, the collected energy consumption data contains abnormalities and redundancy, leading the system to give a wrong prediction for the energy used. Thus, the power management between the consumer and supplier fails. To address this issue, data are passed to a refinement step to remove noise and handle the missing values. Next, the mean of the data is calculated, and the data are organised using a rolling window sequence algorithm, and are then fed into the network.
2. ECP is a challenging task within smart grid planning and electricity marketing, and statistical machine learning techniques such as linear regression or clustering are limited in their effectiveness. In view of this, we add several CNN layers for feature extraction.
3. We propose a M-BDLSTM sequential learning method that obtains CNN features and efficiently predicts the power consumption. M-BDLSTM learns the input sequence in both the forward and backward directions, and concatenates both interpretations at the end to generate the output power ECP.
4. Furthermore, we experimentally demonstrate the superiority of the proposed method over state-of-the-art techniques using 10-fold and hold-out CV method in evaluation. Our method achieves the smallest values of the evaluation metrics, for instance the MSE and RMSE, thus confirms the accurate prediction of energy consumption.

The remainder of the manuscript is organised as follows: Section II explains the proposed method; Section III discusses the experimental evaluation; and the conclusion is presented in Section IV.

II. PROPOSED METHOD

Predicting power consumption in residential buildings is very important in assisting a smart grid to manage and preserve energy in order to ensure efficient use and avoid wastage. Correct prediction of power use is challenging due to the influence of various unpredictable circumstances or the noisy disordering of data, and the methods used sometimes yield wrong predictions. In addition, various techniques have been developed based on traditional networks and are unable to efficiently predict energy demand [41]. Traditional systems have issues related to short-term memory and learning from scratch. These issues can be easily solved using LSTM, which is a specific type of Recurrent Neural Network (RNN) that has received a great deal of attention in the field of machine learning. This network typically contains contextual cells that act as long- or short-term memory cells. These cells modulate the output of the LSTM, and are an important property, since the prediction obtained from the network depends on the historical context of the input. LSTM preserves previous information that has been learnt already, but it is essential to preserve information from the previous and the next states for learning. For this purpose, we need M-BDLSTM, since this network learns the input sequence in both the forward and backward directions. We present a three-step method combining M-BDLSTM with a CNN, as briefly illustrated in Fig. 2. The proposed method is discussed in the following sub-sections.

A. DATA PROCUREMENT AND PRE-PROCESSING

This section provides a detailed discussion of the process of data collection and refinement. The parameter values used throughout the paper are given in Table 2. To collect energy data, smart meters are installed at the edge of the main board where the wires from all appliances are brought to a single point. Data are typically collected monthly or annually, which introduces abnormalities and noise into the data due to climate, metering problems, and individual mistakes or measurement errors. These data need to be refined before being passed to the training process. For this purpose, various smoothing filters can be applied to clean the data, such as LOESS, LOWESS, RLOESS, RLOWESS, Savitsky-Golay, or a moving average filter [25]. We removed the noise by considering the previous time resolution values and applying a moving average filter that is widely used by researchers [13] for smoothing. The data are also usually scattered and need to be brought into a given range, so the data mean is calculated in order to give same range of values for each sample. Fig. 3 illustrates the detailed flow of the data, clearly highlighting each variable and its duration.

TABLE 2. Parameters and their abbreviations.

Abbreviation	Description
ECP	Energy Consumption Prediction
PC	Power Consumption
LSTM	Long-Short Term Memory
M-BDLSTM	Multi-layer Bi-directional LSTM
CNN	Convolutional Neural Network
SMI	Smart Meter Infrastructure
MSE	Mean Square Error
RNN	Recurrent Neural Network
RMSE	Root Mean Square Error
MAE	Mean Absolute Error

B. LSTM AND M-BDLSTM

RNNs [42] typically experience short-term memory problems, particularly when processing and storing a long sequence from the earlier to the later steps. If this process is continued, the RNN may leave out important information from the startup sequence. In addition, RNNs suffer from the problem of vanishing gradient during backpropagation [43], [44]. The gradient updates the neural network weights, thus shrinking during propagation. Its value is extremely small, as it does not contribute to the learning process in an RNN, and the layers that receive a small gradient update stop learning. In this case, the RNN forgets the longer sequences that it has seen previously. An LSTM [45] can provide a solution for short-term memory. An LSTM has internal structures called gates [46] that regulate the processing of information. In LSTM, development and a structural review has performed in [47] regarding the internal memory cell state SST-1. These cell states interact with the intermediate output HST-1 and the successive input XST to regulate the internal state vector that needs to be updated. These are also maintained or expunged on the basis of the previous time step. Internally, the LSTM structure defines input nodes $\boldsymbol{N_{ST}}$ such as input gates $\boldsymbol{I_{ST}}$, forget gates $\boldsymbol{F_{ST}}$, and the output gate $\boldsymbol{O_{ST}}$. These nodes can be formulated as follows:

$$F_{ST} = \sigma(W_{FX}\, X_{ST} + W_{FH}\, H_{ST-1} + B_F) \tag{1}$$
$$I_{ST} = \sigma(W_{IX}\, X_{ST} + W_{IH}\, H_{ST-1} + B_I) \tag{2}$$
$$N_{ST} = \phi(W_{NX}\, X_{ST} + W_{NH}\, H_{ST-1} + B_N) \tag{3}$$
$$O_{ST} = \sigma(W_{OX}\, X_{ST} + W_{OH}\, H_{ST-1} + B_O \tag{4}$$
$$S_{ST} = G_T \ominus I_{ST} + S_{ST-1} \ominus F_{ST}) \tag{5}$$
$$H_{ST} = \phi(S_{ST-1})\ominus O_{ST} \tag{6}$$

In Eqs. 1 to 6, the weight matrices are represented by $\boldsymbol{W_{FX}}, \boldsymbol{W_{FH}}, \boldsymbol{W_{IX}}, \boldsymbol{W_{IH}}, \boldsymbol{W_{NX}}, \boldsymbol{W_{NH}}, \boldsymbol{W_{OX}}$, and $\boldsymbol{W_{HO}}$, which are the corresponding inputs to the network activation function. Here, Θ represents elementwise multiplication, while $\boldsymbol{\sigma}$ stands for the sigmoid activation function and $\boldsymbol{\phi}$ represents the tanh function. The LSTM for a single time step is shown in Fig. 4(a). We also evaluated a Bi-directional LSTM (BDLSTM) in terms of its performance for power energy consumption. The concept of a (BDLSTM) arose from a bidirectional RNN [48], which processes the input data sequence in both the forward and backward directions

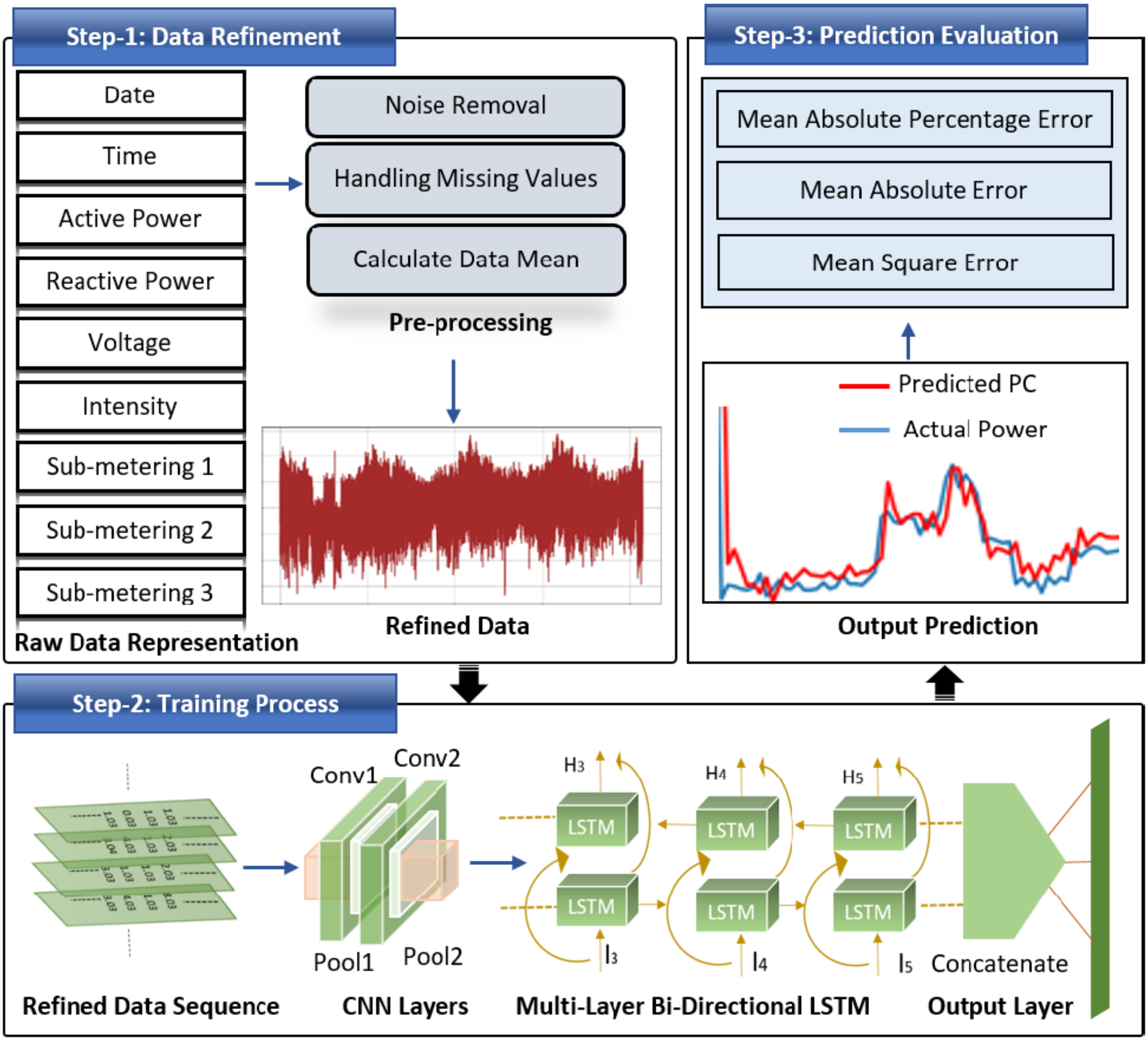

FIGURE 2. **Framework of the three steps of the proposed method. Step 1 is based on data acquisition, where the data are collected and preprocessed to refine and remove abnormalities. The refined data sequence is passed to step 2, where a CNN with an M-BDLSTM network is employed. Finally, step 3 provides the final ECP and an evaluation based on error metrics.**

within the hidden layers. BDLSTM has shown promising results in many fields of computer vision and ECP, such as classification [37], summarisation [39], and load forecasting [49]. We therefore structured the BDLSTM layers that consist of forward and backward layers, as illustrated in Fig. 4(b). In the forward layer, the output sequence $\boldsymbol{I}_T^{>}$ is iteratively calculated via inputs as a positive sequence from T−n to T−1, while in the backward layer, the output sequence $H_T^{<}$ is calculated via reversed inputs such as from T−n to T−1. Finally, LSTM generates the output vector, O_T, where the element is found using Eq. 7.

$$O_T = \sigma(I_T^{>}, H_T^{<}) \tag{7}$$

In Eq. 7, $\boldsymbol{\sigma}$ combines the output sequence, and is also called the concatenating, average, or summation function.

C. CNN WITH M-BDLSTM NETWORK

Due to the encouraging results from CNNs in the field of computer vision, researchers are using this approach for ECP [50] and load forecasting [51] and have produced promising results. Inspired by this, we combine a CNN with M-BDLSTM in the proposed method as a series connection network to predict energy consumption. This network extracts and learns complex features from the PC variables and stores them in a cell state memory before forwarding them to the next layers. The upper layers of the network

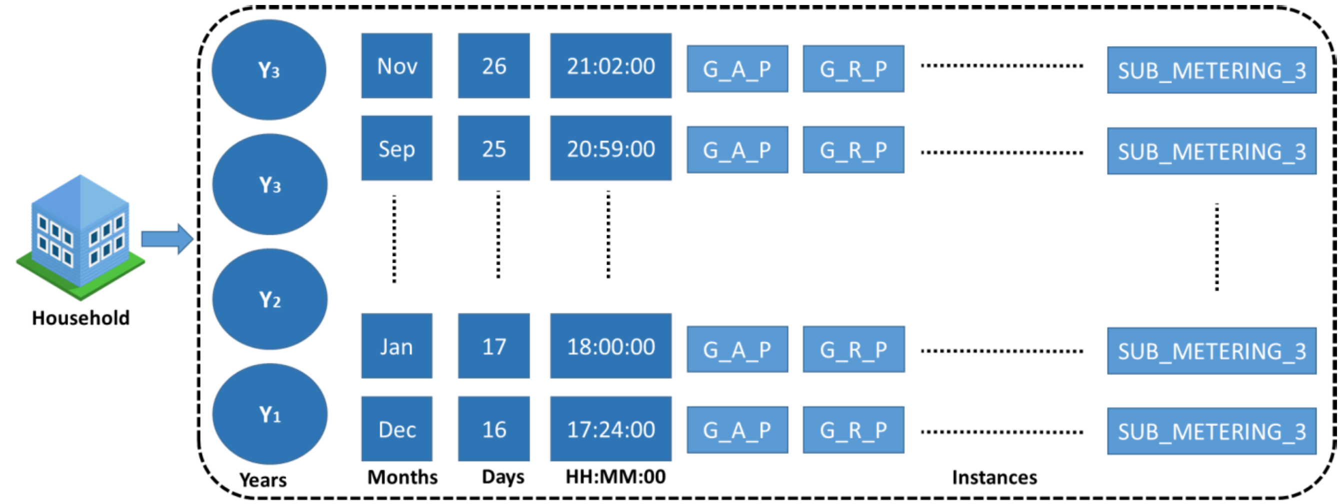

FIGURE 3. Energy data collection from a household, with years divided into months and days. The data are recorded over four years and have a resolution of minutes.

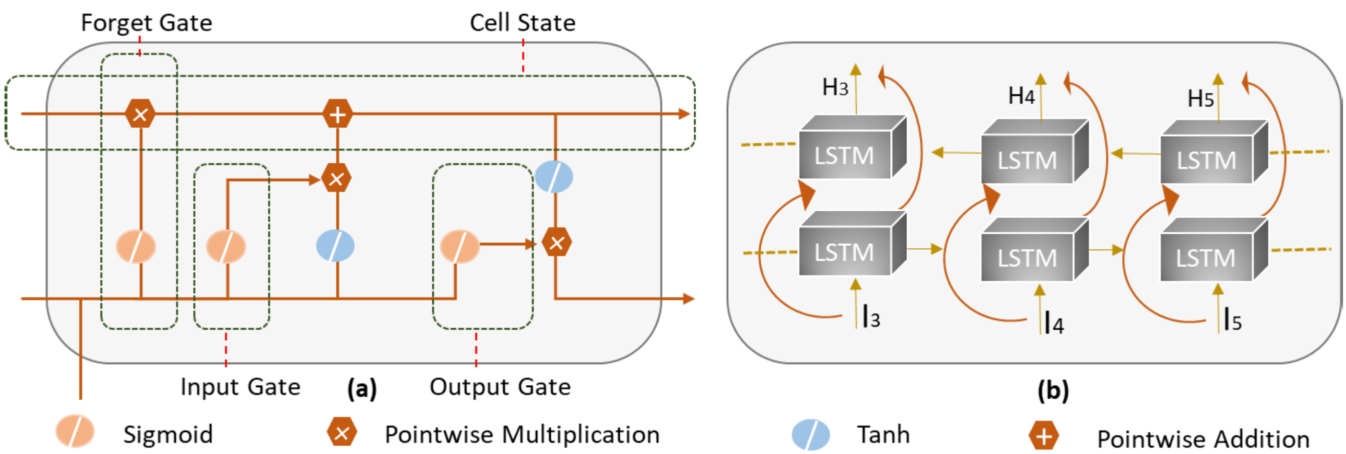

FIGURE 4. (a) The internal mechanism of the LSTM, consisting of different gates and LSTM cells that perform different types of operation, allowing the gates and cells to preserve and forget information; (b) the multi-layer BDLSTM, which obtains the input data sequence and processes it in both the forward and backward directions.

consist of CNNs, and these layers obtain several variable values such as power, voltage, and sub-metering, while other characteristics such as the time, date, occupants' behaviour and occupancy information are modelled as meta information. The variable values collected through sensors are obtained via an input CNN layer, where the output layer extracts local higher-level features into a multi-layer BDLSTM that consists of several hidden layers and an output layer. The hidden layers comprise of a convolutional layer, a ReLu layer, and a pooling layer. Subsequently, this convolution operation extracts prominent information from a multivariate time series sequence and this is passed to the next layer. For instance, if $\boldsymbol{b_i^0} = \{\boldsymbol{b_1},\boldsymbol{b_2},\ldots\ldots,\boldsymbol{b_m}\}$ is the input variable vector and m represents the normalised sequence unit, then Eq. 8 represents the vector $\boldsymbol{C_{ij}^1}$ of the previous layer, $\boldsymbol{d_j^1}$ is the bias for the jth feature map, and w, f, and σ are the weight, filter index value and activation function, respectively. The convolution layer is followed by a max-pooling layer that reduces the network computation costs and parameters.

$$C_{ij}^1 = \sigma(d_j^1 + \sum\nolimits_{f=1}^{M} w_{m},_{j}^{1}\, b_{i+m-1},_{j}^{0}) \tag{8}$$

$$P_{ij}^i = may_{r\in R^{i\times S+k,j}}^{i-1} \tag{9}$$

Eq. 9 represents the max-pooling operation, where $\boldsymbol{S}$ represents the stride, and k represents the pooling size, which should be less than the size of the input y. Furthermore, M-BDLSTM acts as series connection after CNN, and retrieves long short-term information of input sequences related to PC and process them in both the forward and backward directions. This preserves long-term information memory units that update the hidden state of the previous layer. The output obtained from the CNN layers is then fed

TABLE 3. Description and features of household power consumption dataset.

#	Attributes	Units	Remarks
01	Date	dd/mm/yyyy	This variable is given in days, months and years, where the values for the day, month and year are integers ranging from 1 to 30, 1 to 21 and 2006 to 2010, respectively.
02	Time	hh/mm/ss	The time is given in hours, minutes, and seconds (hh/mm/ss), where the hour values range from 0 to 23, and minute values range from 1 to 60.
03	Global Active Power (GAP)	Kilowatts	Total average active power of the household for each minute.
04	Global Reactive Power (GRP)	Kilowatts	Total average reactive power of the household for each minute.
05	Voltage (V)	Volts	Total average voltage for each minute.
06	Global Intensity (GI)	Amperes	Total average current intensity for each minute.
07	Sub-metering 1 (S1)	Watt-hours	Active energy related to kitchen, including dishwasher, oven and microwave.
08	Sub-metering 2 (S2)	Watt-hours	Active energy related to laundry room, including washing machine, tumble-drier and refrigerator.
09	Sub-metering 3 (S3)	Watt-hours	Active energy related to electric water-heater and air-conditioner.

TABLE 4. Detailed quantitative analysis of household power consumption dataset.

Attributes	Date	GAP	GRP	Voltage	GI	S1	S2	S3
Minimum	16/12/2006	0.076	0.000	223.200	0.200	0.000	0.000	0.000
Maximum	26/11/2010	11.122	1.390	254.150	48.400	88.000	80.000	31.000
Std. Dev.	---	1.055	0.113	3.239	4.435	6.139	5.794	8.436
Average	---	1.089	0.124	240.844	4.618	1.117	1.289	6.453

into the gate units. It should be noted that unlike traditional RNN approaches, this network works well for the prediction of PC, and addresses the vanishing gradient problem. This network contains memory cells that update its status by the activation function to regulate continuous values. In addition, the internal mechanism and the operations carried out by the gates, such as the forget, input and output gates, are modelled in Eqs. 1 to 6. The tanh function works to reduce the values to the range $[-1, 1]$ in order to regulate the output of the network. A fully connected layer is then used to produce a prediction of energy consumption over a given period of time. Finally, we add a layer that flattens the feature vector called flatten layer. The functions used in the network are discussed in the previous section.

D. NETWORK ARCHITECTURE

The architecture structure and use settings are changed, using different variables according to the layer adjustment and network performance. Modifications to the size of the kernel, the number of filters and number of strides are performed during the experimental evaluation, and broadly confirm its effect. This setting affects the rate of learning and the network performance, which depend on the nature of the learning data [52]. The network consists of convolutional layers, pooling layers, M-BDLSTM layers and a dense layer [53]. The optimum settings at which the network performed best are applied, the number of filters in the convolutional layer is 64, with the kernel size of three, followed by an activation layer. Each convolution layer is followed by a max-pooling layer with kernel size two. A time-distributed layer is then applied, followed by an M-BDLSTM layer, with 70 neurons in the first layer and 100 in the second. Finally, we add a dense layer for the final output.

III. EXPERIMENTAL EVALUATION

In this section, we provide a detailed discussion of the dataset and the experimental setup, an evaluation of the proposed method and a comparison with state-of-the art techniques.

A. INDIVIDUAL HOUSEHOLD POWER CONSUMPTION DATASET

We evaluated the proposed method using an individual household dataset that is publicly available from the UCI machine learning repository [54], which contains data on electric power consumption between 2006 and 2010. It consists of 2,075,259 instances, including 25,979 missing values (1.25% of the total data), and these missing values are handled in the pre-processing step. This dataset also contains measurements of electric power consumption at a one-minute sampling rate over a period of almost four years. In this dataset, the global active power represented by sub-metering 1, 2 and 3 represent the total active power energy consumed at every minute, given in watt-hours. The electrical quantities are also given with the sub-metering values, and are collected via the

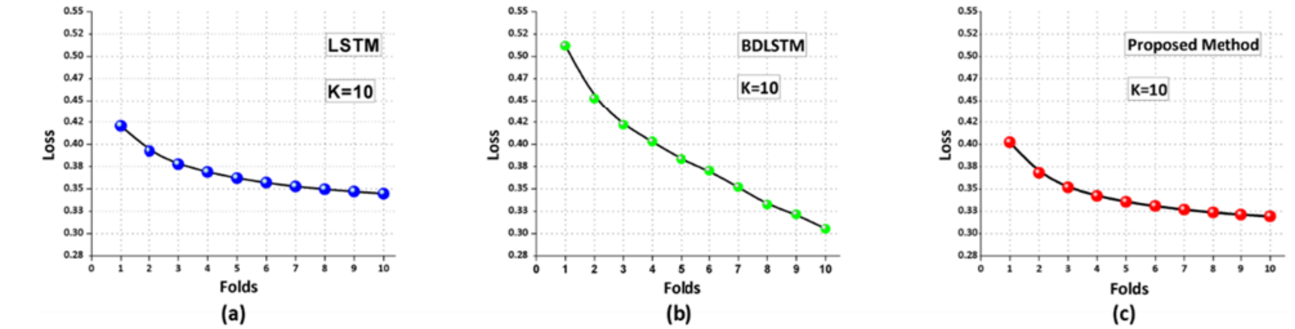

FIGURE 5. Variation in loss with different folds for the power consumption dataset: (a) variation in loss for LSTM for different numbers of folds. The MSE obtained for the first fold is 0.4207, which gradually changes as it iterates with the loss decreasing until by the tenth fold, the obtained MSE becomes 0.3446; (b) results for BDLSTM, which gives 0.5116 MSE for the starting fold, decreasing to 0.3295 by the tenth fold; (c) change in loss for the proposed CNN with M-BDLSTM model, which achieves the smallest value for MSE for the tenth fold of 0.3193, thus demonstrating the effectiveness of the prediction model.

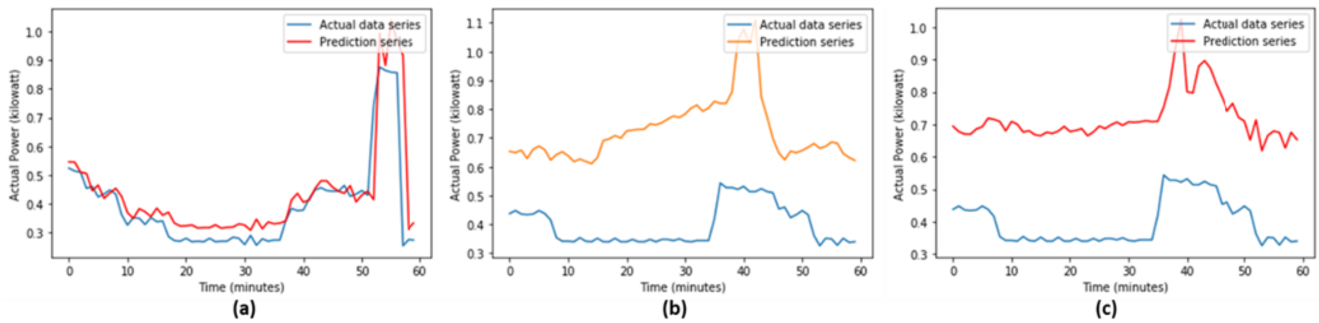

FIGURE 6. Graph obtained for power consumption prediction. The power is given in kilowatt, where (a) shows the prediction results for a CNN with M-BDLSTM. We note that the proposed method gives a pattern that closely corresponds to the actual power consumed. The graphs in (b) and (c) illustrate the results for BDLSTM and LSTM, respectively.

installed sensors. To test the proposed method, we used a 60-minute sequence of power consumption to predict next 60 minutes, where the input time data series was converted into a two-dimensional window sequence and was fed as input to the proposed network for training. In Table 3, we describe the variables used in the individual household dataset with remarks. A total of nine variables with their respective units are shown, and these form the power consumption data. We also give quantitative details of the power consumption dataset in Table 4 given in the literature [55].

From Table 4, we can see that the maximum active power consumed is 11.122 Kilowatts, while the minimum active power is 0.076 Kilowatts. This means that sub-metering 1 consumes more energy compared to other attributes; this is because sub-metering 1 is related to kitchen devices, which are highly energy consuming.

B. EXPERIMENTAL SETUP

We analysed and verified the effectiveness of the proposed CNN with M-BDLSTM model using different kinds of experiments to evaluate its performance. Most existing approaches use either the hold-out CV method or a k-fold CV method to evaluate performance. In view of this, the proposed method was evaluated using 10-fold CV, and achieved the lowest value of MSE for predicting power consumption compared to state-of-the-art techniques. We implemented our system in Python (version 3.5) in the Keras deep learning framework, with Tensorflow at the backend and utilised RMSprop as an optimizer. As we were dealing with a regression problem, we evaluated the quality of the prediction model using the three basic metrics of MSE, RMSE, and Mean Absolute Error (MAE). These metrics are widely used to evaluate the error rates for prediction using regression models. Let $\boldsymbol{y_i^{\sim}}$ represent the values of variables for n prediction samples of power consumption, and let $\boldsymbol{y_i}$ represent the observed values. Eqs. 10 to 12 then represent the MSE, RMSE and MAE, respectively.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - y_i^{\sim})^2 \tag{10}$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i^{\sim})^2} \tag{11}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\, y_i - y_i^{\sim} \,| \tag{12}$$

Another statistical metric that measures the forecast performance to check the correctness of the proposed method is

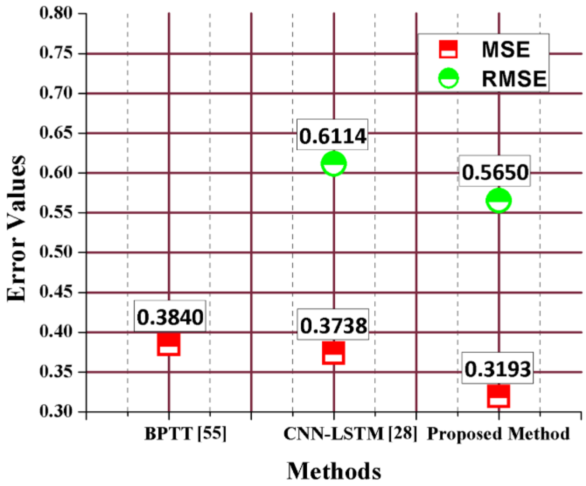

FIGURE 7. Comparative analysis of the proposed method with state-of-the-art techniques on the basis of MSE and RMSE except BPTT where only RMSE is calculated.

Mean Absolute Percentage Error (MAPE), as given in Eq. 13. Further, additional metrics such as bias or Mean Bias Error (MBE), that compute the average of overall forecasting errors which is known as systematic error. The MBE metric is given in Eq. 14. Sometimes, the value obtained for bias or MBE is negative that represents under-forecasting phenomenon while the positive value of MBE shows the over-forecasting. The forecast is greater and less than the actual data in over-forecasting and under-forecasting, respectively. We calculate the MAPE and MBE values for the proposed 10-fold CV method that can be fairly compare with state-of-the-art techniques. However, for holdout method, only the basic metrics such as MSE, RMSE, and MAE are calculated.

$$MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{y_i - y_i^{\sim}}{y_i} \right| \tag{13}$$

$$MBE = \frac{1}{n} \sum_{I=1}^{n} (y_i - y_i^{\sim}) \tag{14}$$

C. EVALUATION USING THE K-FOLD CV METHOD AND COMPARISON WITH DEEP LEARNING METHODS

K-fold CV is a method in which the dataset is partitioned into k subsets, and the hold-out method is repeated k times [56]. For each fold, one subset is used as a test set, and the remaining k-1 subsets are combined to form the training set. This method has several advantages, as each data sample is used in the validation and training sets $k-1$ times. The bias and variance are also significantly reduced, since most of the data are used in fitting. Generally, the literature on machine learning recommends a value for k of 10 [57], so we also chose k =10. We conducted numerous experiments in order to confirm the superiority of the proposed method over other deep learning-based methods such as LSTM and BDLSTM. These methods are summarised in Table 5, and the results are assessed using the five error metrics of MSE, RMSE, MAE, MAPE and MBE. Fig. 5 graphically illustrates the experiments summarised in Table 5, where the loss-to-folds ratio is given. As the k value of the fold is proceeded, the loss value changes. The MSE values obtained for LSTM, BDLSTM and proposed method are 0.3446, 0.3295 and 0.3193, respectively.

TABLE 5. Results of 10-fold cross-validation for various machine learning methods and comparison with the proposed method.

Methods	Evaluation Metrics				
	MSE	RMSE	MAE	MAPE	MBE
LSTM	0.3446	0.5870	0.3667	0.3201	0.03800
BDLSTM	0.3295	0.5740	0.3575	0.3111	0.03571
Proposed Method	**0.3193**	**0.5650**	**0.3469**	**0.2910**	**0.03286**

TABLE 6. Results obtained for the hold-out cross validation method with various deep learning methods.

Methods	Evaluation Metrics		
	MSE	RMSE	MAE
LSTM	0.4171	0.6458	0.4008
BDLSTM	0.3567	0.5972	0.3692
CNN-LSTM	0.4307	0.6562	0.4089
CNN-M-BDLSTM	0.3489	0.5906	0.3730

D. PERFORMANCE EVALUATION USING HOLD-OUT CV METHOD

In our experiments, we also used the hold-out method, a type of CV that separates the data into a training and testing set. During the training process, the model is fitted using a training set; the model function then predicts the output values for the unseen data in the testing set. This method is preferable for residual methods, and takes less time to compute. In this method, 80% of the data are used as the training set, while the remaining 20% are used for testing. We also conducted various experiments on different machine learning models for comparison, such as the LSTM, BDLSTM, CNN-LSTM and CNN-M-BDLSTM models. When using the aforementioned methods, the prediction models were trained with up to 20 epochs. The results obtained for each deep learning model are shown in Table 6. We first used LSTM to check the hold-out method, and obtained a value of 0.4171 for the MSE. We then used BDLSTM, and obtained a value of 0.3567 for MSE. The CNN was then combined with an LSTM and a value of 0.4307 was obtained. Finally, we combined CNN with M-BDLSTM and recorded a value of 0.3489.

E. PREDICTION RESULTS OBTAINED VIA CNN WITH M-BDLSTM NETWORK

This section discusses the results of the proposed method for the prediction of power consumption. A graphical representation of the hourly power consumption data and its prediction for the next hour is shown in Fig. 6. Fig. 6(a) demonstrates that the CNN with M-BDLSTM closely predicts the energy

TABLE 7. Detailed performance analysis of competitive state-of-the-art techniques with our method.

Techniques	Evaluation Metrics				
	MSE	RMSE	MAE	MAPE	MBE
CNN-LSTM [28]	0.3738	0.6114	0.3493	0.3484	---
BPTT [55]	0.3840	---	0.3953	---	---
Seq2Seq [58]	---	0.742	---	---	---
FCBRM [59]	---	0.6663	---	---	---
Proposed Method	**0.3193**	**0.5650**	**0.3469**	**0.2910**	**0.03286**

consumption and other global characteristics of power better than the other deep learning methods given in Table 5. This network also predicts the native characteristics of the power consumption. We also observe that the proposed method organises and models the irregular inclinations of power consumption more effectively. Next, we confirmed that the complex pattern generated by the time series model of power consumption was predicted appropriately, and showed that the proposed method moderated the error at each interval compared to LSTM and BDLSTM. The prediction series for the original input data to BDLSTM and LSTM are given in Fig. 6 (b) and Fig. 6 (c), respectively.

During experiments, each model consumes distinct training time to complete one fold. The proposed model consumes average time of 3 hours to complete one fold and 30 hours for the whole model that is two times faster than LSTM and BDLSTM.

F. COMPARATIVE ANALYSIS OF OUR PROPOSED METHOD

In this section, we evaluate and compare the results of our proposed method in terms of performance with those of competitive state-of-the-art techniques. using the same dataset of household power consumption. For a fair comparison, we use the same time resolution of minutes. The method proposed in [28] used a hybrid approach combining a CNN with LSTM to predict residential power consumption, and obtained values of 0.3738 and 0.6114 for MSE and RMSE, respectively. We also compare our method with one proposed by Kim *et al.* [55], who used an auto-encoder-based deep learning model to forecast energy demand and used a back-propagation through time algorithm to train their forecasting time series model.

They also evaluated their method using MSE, obtaining values of 0.3840 and 0.3953 for the MSE and MAE metrics, respectively. Another study was carried out by Marino *et al.* [58], who proposed a method based on a Deep Neural Network (DNN) and Sequence-to-sequence (Seq2Seq) LSTM architecture to forecast the power energy load, and recorded a value of 0.742 for the RMSE. The method presented in [59] used two stochastic models, a factored conditional restricted Boltzmann machine and a conditional restricted Boltzmann machine for energy forecasting, and obtained a value of 0.6663 for the RMSE. We compared our results with these models and confirmed that our proposed method outperforms the state-of-art techniques. The blank area in Table 7 represents that the mentioned techniques have not used the corresponding metrics for evaluation such as MBE, and MAPE etc., Table 7 summarises the performance of the existing and proposed methods. Further, a comparative analysis based on MSE and RMSE is shown in Fig. 7 that clearly verify the effectiveness of the proposed method.

IV. CONCLUSION

Modelling the consumption and management of power in residential buildings, power sectors, and smart grids is a challenging task, due to noise disturbance and climate conditions. The rate of consumption of power is also increasing markedly. Prediction of this power consumption is now essential in order to improve usage performance, infrastructure, and cooperation between residential energy and smart grids. In this context, we presented a hybrid technique using a CNN and M-BDLSTM for ECP, consisting of three steps. The first step involves efficient pre-processing to verify, cleanse and set the data. As the data are generally scattered before processing, we calculate the mean of the data and convert this into a rolling window sequence. The second step consists of a CNN with an M-BDLSTM network that obtains the input in sequence form and processes it. The M-BDLSTM processes the input sequence in both the forward and backward directions, and concatenates the output predictions. The performance was then evaluated using error metrics. The proposed method outperformed existing methods, achieving the smallest value for the MSE. The proposed method is tested on publicly available dataset of household electric power consumption [54]. In future, we aim to further improve the proposed method by considering different time scales and horizons such as months, years and decades. These horizons can be investigated for energy management among various industries and smart grids to enable the efficient usage of electrical energy. Currently, the proposed model is deployable on PC only, therefore we intend to deploy the prediction model on resource-constrained devices for the Internet of Things (IoT) environment by reducing the cost and computational complexity. Current household data do not contain characteristics of the environment such as the occupants' behaviour and climate conditions, that are influencing factors to be consider. We aim to incorporate such kind of datasets to verify and confirm its impact on the prediction variables that need further research in this area.

REFERENCES

[1] Q. Li, C. Peng, M. Chen, F. Chen, W. Kang, J. M. Guerrero, and D. Abbott, "Networked and distributed control method with optimal power dispatch for islanded microgrids," *IEEE Trans. Ind. Electron.*, vol. 64, no. 1, pp. 493–504, Jan. 2017.

[2] W. Kong, Z. Y. Dong, D. J. Hill, F. Luo, and Y. Xu, "Short-term residential load forecasting based on resident behaviour learning," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1087–1088, Jan. 2017.

[3] P. Moriarty and D. Honnery, "Ecosystem maintenance energy and the need for a green EROI," *Energy Policy*, vol. 131, pp. 229–234, Aug. 2019.

[4] A. Ahmad, M. Hassan, M. Abdullah, H. Rahman, F. Hussin, H. Abdullah, and R. Saidur, "A review on applications of ANN and SVM for building electrical energy consumption forecasting," *Renew. Sustain. Energy Rev.*, vol. 33, pp. 102–109, May 2014.

[5] WorldData. *Energy Consumption in South Korea*. [Online]. Available: https://www.worlddata.info/asia/south-korea/energy-consumption.php, 2019.

[6] M. Zhang and C. Bai, ''Exploring the influencing factors and decoupling state of residential energy consumption in Shandong,'' *J. Cleaner Prod.*, vol. 194, pp. 253–262, Sep. 2018.

[7] A. M. De Livera, R. J. Hyndman, and R. D. Snyder, ''Forecasting time series with complex seasonal patterns using exponential smoothing,'' *J. Amer. Stat. Assoc.*, vol. 106, no. 496, pp. 1513–1527, Dec. 2011.

[8] L. Yang, H. Yan, and J. C. Lam, ''Thermal comfort and building energy consumption implications—A review,'' *Appl. Energy*, vol. 115, pp. 164–173, Feb. 2014.

[9] S. Kim and S. Kim, ''A multi-criteria approach toward discovering killer IoT application in Korea,'' *Technol. Forecasting Social Change*, vol. 102, pp. 143–155, Jan. 2016.

[10] S. Malik and D. Kim, ''Prediction–learning algorithm for efficient energy consumption in smart buildings based on particle regeneration and velocity boost in particle swarm optimization neural networks,'' *Energies*, vol. 11, no. 5, p. 1289, May 2018.

[11] H. Khosravani, M. Castilla, M. Berenguel, A. Ruano, and P. Ferreira, ''A comparison of energy consumption prediction models based on neural networks of a bioclimatic building,'' *Energies*, vol. 9, no. 1, p. 57, Jan. 2016.

[12] N. Fumo and M. Rafe Biswas, ''Regression analysis for prediction of residential energy consumption,'' *Renew. Sustain. Energy Rev.*, vol. 47, pp. 332–343, Jul. 2015.

[13] I. Ullah, R. Ahmad, and D. Kim, ''A prediction mechanism of energy consumption in residential buildings using hidden Markov model,'' *Energies*, vol. 11, no. 2, p. 358, Feb. 2018.

[14] K. Kampouropoulos, J. J. Cardenas, F. Giacometto, and L. Romeral, ''An energy prediction method using adaptive neuro–fuzzy inference system and genetic algorithms,'' in *Proc. IEEE Int. Symp. Ind. Electron.*, May 2013, pp. 1–6.

[15] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, ''Extreme learning machine: Theory and applications,'' *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.

[16] Z. Wang, Y. Wang, R. Zeng, R. S. Srinivasan, and S. Ahrentzen, ''Random forest based hourly building energy prediction,'' *Energy Buildings*, vol. 171, pp. 11–25, Jul. 2018.

[17] L. M. Candanedo, V. Feldheim, and D. Deramaix, ''Data driven prediction models of energy use of appliances in a low-energy house,'' *Energy Buildings*, vol. 140, pp. 81–97, Apr. 2017.

[18] Z. Wang and Y. Ding, ''An occupant-based energy consumption prediction model for office equipment,'' *Energy Buildings*, vol. 109, pp. 12–22, Dec. 2015.

[19] H. Zhong, J. Wang, H. Jia, Y. Mu, and S. Lv, ''Vector field-based support vector regression for building energy consumption prediction,'' *Appl. Energy*, vol. 242, pp. 403–414, May 2019.

[20] H. Cai, S. Shen, Q. Lin, X. Li, and H. Xiao, ''Predicting the energy consumption of residential buildings for regional electricity supply–side and demand–side management,'' *IEEE Access*, vol. 7, pp. 30386–30397, 2019.

[21] M. Fayaz, H. Shah, A. Aseere, W. Mashwani, and A. Shah, ''A framework for prediction of household energy consumption using feed forward back propagation neural network,'' *Technologies*, vol. 7, no. 2, p. 30, Apr. 2019.

[22] H. Liu, J. Su, J. Qi, N. Wang, and C. Li, ''Decentralized voltage and power control of multi–machine power systems with global asymptotic stability,'' *IEEE Access*, vol. 7, pp. 14273–14282, 2019.

[23] C. Tian, C. Li, G. Zhang, and Y. Lv, ''Data driven parallel prediction of building energy consumption using generative adversarial nets,'' *Energy Buildings*, vol. 186, pp. 230–243, Mar. 2019.

[24] A. Almalaq and J. J. Zhang, ''Evolutionary deep learning–based energy consumption prediction for buildings,'' *IEEE Access*, vol. 7, pp. 1520–1531, 2019.

[25] M. Fayaz and D. Kim, ''A prediction methodology of energy consumption based on deep extreme learning machine and comparative analysis in residential buildings,'' *Electronics*, vol. 7, no. 10, p. 222, Sep. 2018.

[26] M. Cui, J. Wang, Y. Wang, R. Diao, and D. Shi, ''Robust time–varying synthesis load modeling in distribution networks considering voltage disturbances,'' *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4438–4450, Nov. 2019.

[27] L. Wen, K. Zhou, S. Yang, and X. Lu, ''Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting,'' *Energy*, vol. 171, pp. 1053–1065, Mar. 2019.

[28] T.-Y. Kim and S.-B. Cho, ''Predicting residential energy consumption using CNN–LSTM neural networks,'' *Energy*, vol. 182, pp. 72–81, Sep. 2019.

[29] K. Li, X. Xie, W. Xue, X. Dai, X. Chen, and X. Yang, ''A hybrid teaching-learning artificial neural network for building electrical energy consumption prediction,'' *Energy Buildings*, vol. 174, pp. 323–334, Sep. 2018.

[30] L. Ruiz, R. Rueda, M. Cuéllar, and M. Pegalajar, ''Energy consumption forecasting based on Elman neural networks with evolutive optimization,'' *Expert Syst. Appl.*, vol. 92, pp. 380–389, Feb. 2018.

[31] M. Sajjad, M. Nasir, F. U. M. Ullah, K. Muhammad, A. K. Sangaiah, and S. W. Baik, ''Raspberry Pi assisted facial expression recognition framework for smart security in law-enforcement services,'' *Inf. Sci.*, vol. 479, pp. 416–431, Apr. 2019.

[32] D. Zhang, Z. Xu, C. Li, R. Yang, M. Shahidehpour, Q. Wu, and M. Yan, ''Economic and sustainability promises of wind energy considering the impacts of climate change and vulnerabilities to extreme conditions,'' *Electr. J.*, vol. 32, no. 6, pp. 7–12, Jul. 2019.

[33] C. Li, L. Zeng, B. Zhou, X. Liu, Q. Wu, D. Zhang, and S. Huang, ''An optimal coordinated method for EVs participating in frequency regulation under different power system operation states,'' *IEEE Access*, vol. 6, pp. 62756–62765, 2018.

[34] S. Khan, K. Muhammad, S. Mumtaz, S. W. Baik, and V. H. C. De Albuquerque, ''Energy–efficient deep CNN for smoke detection in foggy IoT environment,'' *IEEE Internet Things J.*, vol. 6, no. 6, pp. 9237–9245, Dec. 2019.

[35] K. Muhammad, S. Khan, V. Palade, I. Mehmood, and V. H. C. De Albuquerque, ''Edge intelligence–assisted smoke detection in foggy surveillance environments,'' *IEEE Trans. Ind. Informat.*, to be published.

[36] A. Ullah, K. Muhammad, I. U. Haq, and S. W. Baik, ''Action recognition using optimized deep autoencoder and CNN for surveillance data streams of non-stationary environments,'' *Future Gener. Comput. Syst.*, vol. 96, pp. 386–397, Jul. 2019.

[37] A. Ullah, J. Ahmad, K. Muhammad, M. Sajjad, and S. W. Baik, ''Action recognition in video sequences using deep bi–directional LSTM with CNN features,'' *IEEE Access*, vol. 6, pp. 1155–1166, 2018.

[38] F. U. M. Ullah, A. Ullah, K. Muhammad, I. U. Haq, and S. W. Baik, ''Violence detection using spatiotemporal features with 3D convolutional neural network,'' *Sensors*, vol. 19, no. 11, p. 2472, May 2019.

[39] T. Hussain, K. Muhammad, A. Ullah, Z. Cao, S. W. Baik, and V. H. C. De Albuquerque, ''Cloud–assisted multi–view video summarization using CNN and bi–directional LSTM,'' *IEEE Trans. Ind. Inf.*, to be published.

[40] I. U. Haq, K. Muhammad, T. Hussain, S. Kwon, M. Sodanil, S. W. Baik, and M. Y. Lee, ''Movie scene segmentation using object detection and set theory,'' *Int. J. Distrib. Sensor Netw.*, vol. 15, no. 6, Jun. 2019, Art. no. 155014771984527.

[41] F. Kaytez, M. C. Taplamacioglu, E. Cam, and F. Hardalac, ''Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines,'' *Int. J. Electr. Power Energy Syst.*, vol. 67, pp. 431–438, May 2015.

[42] T. Mikolov, M. Karafiát, L. Burget, J. Černocký, and S. Khudanpur, ''Recurrent neural network based language model,'' in *Proc. 11th Annu. Conf. Int. Speech Commun. Assoc.*, 2010.

[43] Y. Bengio, P. Simard, and P. Frasconi, ''Learning long-term dependencies with gradient descent is difficult,'' *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.

[44] S. Hochreiter, Y. Bengio, P. Frasconi, and J. Schmidhuber, ''Gradient flow in recurrent nets: The difficulty of learning long-term dependencies,'' in *A Field Guide to Dynamical Recurrent Neural Networks*. Piscataway, NJ, USA: IEEE Press, 2001.

[45] S. Hochreiter and J. Schmidhuber, ''Long short-term memory,'' *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[46] F. A. Gers, J. Schmidhuber, and F. Cummins, ''Learning to forget: Continual prediction with LSTM,'' in *Proc. 9th Int. Conf. Artif. Neural Netw.*, 1999.

[47] Z. C. Lipton, J. Berkowitz, and C. Elkan, ''A critical review of recurrent neural networks for sequence learning,'' 2015, *arXiv:1506.00019*. [Online]. Available: https://arxiv.org/abs/1506.00019

[48] M. Schuster and K. Paliwal, ''Bidirectional recurrent neural networks,'' *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997.

[49] X. Tang, Y. Dai, T. Wang, and Y. Chen, ''Short-term power load forecasting based on multi-layer bidirectional recurrent neural network,'' *IET Gener., Transmiss. Distrib.*, vol. 13, no. 17, pp. 3847–3854, Sep. 2019.

[50] S. Khan, N. Javaid, A. Chand, A. B. M. Khan, F. Rashid, and I. U. Afridi, ‘‘Electricity Load Forecasting for Each Day of Week Using Deep CNN,’’ in *Proc. Workshops Int. Conf. Adv. Inf. Netw. Appl.*, 2019, pp. 1107–1119.

[51] S. Aimal, N. Javaid, T. Islam, W. Z. Khan, M. Y. Aalsalem, and H. Sajjad, ‘‘An efficient CNN and KNN data analytics for electricity load forecasting in the smart grid,’’ in *Proc. Workshops Int. Conf. Adv. Inf. Netw. Appl.*, 2019, pp. 592–603.

[52] K. He and J. Sun, ‘‘Convolutional neural networks at constrained time cost,’’ in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 5353–5360.

[53] C. Zhou, C. Sun, Z. Liu, and F. Lau, ‘‘A C-LSTM neural network for text classification,’’ 2015, *arXiv:1511.08630*. [Online]. Available: https://arxiv.org/abs/1511.08630

[54] A. B. G. Hebrail. (2012). *Individual Household Electric Power Consumption Data Set*. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+consumption

[55] J.-Y. Kim and S.-B. Cho, ‘‘Electric energy consumption prediction by deep learning with state explainable autoencoder,’’ *Energies*, vol. 12, no. 4, p. 739, Feb. 2019.

[56] J. Schneider. (1997). *Cross Validation*. [Online]. Available: https://www.cs.cmu.edu/~schneide/tut5/node42.html

[57] M. Kuhn and K. Johnson, *Applied Predictive Modeling*, vol. 26. New York, NY, USA: Springer, 2013, doi: 10.1007/978-1-4614-6849-3.

[58] D. L. Marino, K. Amarasinghe, and M. Manic, ‘‘Building energy load forecasting using Deep Neural Networks,’’ in *Proc. 42nd Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2016, pp. 7046–7051.

[59] E. Mocanu, P. H. Nguyen, M. Gibescu, and W. L. Kling, ‘‘Deep learning for estimating building energy consumption,’’ *Sustain. Energy, Grids Netw.*, vol. 6, pp. 91–99, Jun. 2016.

FATH U MIN ULLAH (Student Member, IEEE) received the B.S. degree in computer science from Islamia College, Peshawar, Pakistan, in 2017. He is currently pursuing the M.S. degree with the Intelligent Media Laboratory, Sejong University, South Korea. His research interests include computer vision techniques, violence detection, deep learning, image processing, power energy consumption prediction, smart grid, surveillance video analysis, and facial expression recognition. He has published several articles in reputed peer reviewed international journals such information science and sensors. Moreover, he has expertise in embedded devices such as raspberry pi and so on.

AMIN ULLAH (Student Member, IEEE) received the bachelor’s degree in computer science from the Islamia College Peshawar, Peshawar, Pakistan. He is currently pursuing the M.S. leading to Ph.D. degree with the Intelligent Media Laboratory, Sejong University, South Korea. He has published several articles in reputed peer reviewed international journals and conferences, including IEEE TIE, IEEE IoTJ, IEEE Access, and Elsevier FGCS. His research interests include human actions and activity recognition, sequence learning, image and video analysis, and deep learning for multimedia understanding.

IJAZ UL HAQ (Student Member, IEEE) received the B.S. degree in computer science from the Islamia College at Peshawar, Peshawar, Pakistan. He is currently pursuing the M.S. leading to Ph.D. degree with the Intelligent Media Laboratory, Sejong University, South Korea. His research interests include video summarization, image and video analysis, image hashing, steganography, and deep learning for multimedia understanding.

SEUNGMIN RHO (Member, IEEE) received the M.S. and Ph.D. degrees in computer science from Ajou University, South Korea, in 2003 and 2008, respectively. He visited the Multimedia Systems and Networking Laboratory, The University of Texas at Dallas, from 2003 to 2004. Before, he joined the Computer Sciences Department, Ajou University. From 2008 to 2009, he was a Postdoctoral Research Fellow with the Computer Music Laboratory, School of Computer Science, Carnegie Mellon University. He was a Research Professor with the School of Electrical Engineering, Korea University, from 2009 to 2011. In 2012, he was an Assistant Professor with the Division of Information and Communication, Baekseok University. He is currently a Faculty Member with the Department of Software, Sejong University, South Korea. His current research interests include database, big data analysis, music retrieval, multimedia systems, machine learning, knowledge management, and computational intelligence. He has published 100 articles in refereed journals and conference proceedings in these areas. He has been involved in more than 20 conferences and workshops as various chairs and more than 30 conferences/workshops as a Program Committee Member. He has received a few awards, including Who’s Who in America, Who’s Who in Science and Engineering, and Who’s Who in the World, in 2007 and 2008, respectively. He has edited a number of international journal special issues as a Guest Editor, such as *Multimedia Systems*, *Information Fusion*, *Engineering Applications of Artificial Intelligence*, *New Review of Hypermedia and Multimedia*, *Multimedia Tools and Applications*, *Personal and Ubiquitous Computing*, *Telecommunication Systems*, and *Ad Hoc and Sensor Wireless Networks*.

SUNG WOOK BAIK (Member, IEEE) received the Ph.D. degree in information technology engineering from George Mason University, Fairfax, VA, USA, in 1999. He worked at Datamat Systems Research, Inc., as a Senior Scientist of the Intelligent Systems Group from 1997 to 2002. Since 2002, he has been serving as a Faculty Member with the Department of Digital Contents, Sejong University, Seoul, South Korea. He is currently a Full Professor with the Department of Digital Contents, the Chief of Sejong Industry-Academy Cooperation Foundation, and the Head of the Intelligent Media Laboratory (IM Lab), Sejong University. He has published more than 100 articles in peer-reviewed international journals in the mentioned areas with main target on top venues of these domains including IEEE IoTJ, TSMC-Systems, COMMAG, TIE, TII, Access, Elsevier *Neurocomputing*, FGCS, PRL, Springer MTAP, JOMS, RTIP, and MDPI *Sensors*. He is serving as a Professional Reviewer for several well-reputed journals and conferences, including the IEEE Transactions on Industrial Informatics, the IEEE Transactions on Cybernetics, IEEE Access, and MDPI *Sensors*. His research interests include image processing, pattern recognition, video analytics, big data analysis, multimedia data processing, energy load forecasting, the IoT, the IIoT, and smart cities. His specific areas in image processing include image indexing and retrieval for various applications and in video analytics the major focus is video summarization, action and activity recognition, anomaly detection, and CCTV data analysis. He is involved in several projects including AI-convergence technologies and multiview video data analysis for smart cities, effective energy management methods, experts’ education for industrial unmanned aerial vehicles, and big data models analysis. His work supported by the Korea Institute for Advancement of Technology and the Korea Research Foundation. He has several Korean and an internationally accepted patent with main theme of disaster management, image retrieval, and speaker reliability measurement.