Electrical Energy Prediction in Residential Buildings for Short-Term Horizons Using Hybrid Deep Learning Strategy

Zulfiqar Ahmad Khan, Amin Ullah, Waseem Ullah, Seungmin Rho, Miyoung Lee and Sung Wook Baik *

Sejong University, Seoul 143-747, Korea; mzulfiqar3797@gmail.com (Z.A.K.); qamin3797@gmail.com (A.U.); kwaseem3797@gmail.com (W.U.); smrho@sejong.ac.kr (S.R.); miylee@sejong.ac.kr (M.L.)
* Correspondence: sbaik@sejong.ac.kr

Received: 9 November 2020; Accepted: 29 November 2020; Published: 2 December 2020

Abstract: Smart grid technology based on renewable energy and energy storage systems are attracting considerable attention towards energy crises. Accurate and reliable model for electricity prediction is considered a key factor for a suitable energy management policy. Currently, electricity consumption is rapidly increasing due to the rise in human population and technology development. Therefore, in this study, we established a two-step methodology for residential building load prediction, which comprises two stages: in the first stage, the raw data of electricity consumption are refined for effective training; and the second step includes a hybrid model with the integration of convolutional neural network (CNN) and multilayer bidirectional gated recurrent unit (MB-GRU). The CNN layers are incorporated into the model as a feature extractor, while MB-GRU learns the sequences between electricity consumption data. The proposed model is evaluated using the root mean square error (RMSE), mean square error (MSE), and mean absolute error (MAE) metrics. Finally, our model is assessed over benchmark datasets that exhibited an extensive drop in the error rate in comparison to other techniques. The results indicated that the proposed model reduced errors over the individual household electricity consumption prediction (IHEPC) dataset (i.e., RMSE (5%), MSE (4%), and MAE (4%)), and for the appliances load prediction (AEP) dataset (i.e., RMSE (2%), and MAE (1%)).

Keywords: bidirectional gated recurrent unit; convolutional neural networks; electricity consumption prediction; hybrid deep learning model; residential load prediction

1. Introduction

The electric power industry plays an important role in the economic development of a country, and its decisive operation provides significant societal wellbeing. As reported in [1] the global energy consumption is increasing for sustainable advancement in society; therefore, the effectiveness of electricity consumption prediction needs to be improved [2]. As reported by the World Energy Outlook in 2017, the compound annual growth rate (CAGR) of electricity demand will have a global incremental rise of 1.0% in the period of 2016–2040 [3]. Another report presented in [4] described that residential buildings generally consume 27% of the total energy consumption, whereas buildings in the United States (US) consume 40% of their national energy [5]. Owing to high energy consumption levels in residential buildings, efficient management of their consumption is essential. Therefore, proper planning of energy is vital for energy saving, which is possible through effective energy consumption prediction models.

The electricity prediction strategies for short-term horizons are categorized into four types: very-short, short, medium, and long-term [6,7]. Short and very short-term predictions refer to minutely
ahead predictions up to several days. Medium-term load prediction means one week ahead prediction or up to a year, whereas annual or several years’ ahead prediction is called very long-term prediction. Each electricity prediction horizon has its own applications; however, in this study, we focus on short-term and very-short-term predictions [6,8–10]. The major applications of short-term electric load prediction are power plant’s reliable and secure operation, reliability and economic dispatch, and power system generation scheduling. Short-term load prediction ensures power system security, and it is an essential tool for the determination of the optimal operational state. Reliability and economic dispatch are also important applications of power systems in short-term horizons, wherein the abrupt variation of load demand fluctuates its reliability. This causes a power supply shortage if the load demand is underestimated, which makes it difficult to manage overload conditions and the quality of the overall power supply system. Another application of short-term load prediction is generation scheduling, which can be achieved through accurate load prediction to verify the allocation of operational limitations, generation resources, equipment usage, and environmental constraints. In the literature, several studies have been conducted for short-term load prediction, electricity load forecasting, electricity demand forecasting, electricity storage, and occupant behavior [11–15].

The mainstream electricity load prediction models are mainly grouped into two categories: statistical models and artificial intelligence models [16,17]. Statistical models such as Auto Regressive Integrated Moving Average (ARIMA) [18], linear regression [19], Kalman filtering [20], and clustering [21,22] etc. were used for load prediction in the early days. These models are effective in learning linear data but inadequate to learn the nonlinear complex electricity load. Besides this, the artificial intelligence models can learn nonlinear complex and linear electricity loads, which are further divided into shallow and deep structure methods. Shallow based methods include random forest [23], wavelet neural networks [24], support vector machines (SVMs) [25], artificial neural networks (ANN) [26], and extreme learning machines [27]. Shallow based methods perform well compared to statistical methods but perform poorly in feature mining. Hence, these methods require more features and selects strong features to enhance the prediction accuracy. Obtaining optimal feature extraction is a challenging task for these methods. Further, these methods have insufficient generalization ability over different datasets due to small hypothesis owing to less number of parameters. Deep structure methods have the ability to address the aforementioned concerns of shallow-based methods using multi-layer processing and hierarchical feature learning from electricity historical data. Recently, recurrent neural networks (RNNs) and convolutional neural networks (CNNs) are two powerful architecture proposed in the literature for the analysis of time series data. For instance, Amarasinghe et al. [28] developed a CNN-based methodology for electricity load forecasting and compared their results with a factored restricted Boltzmann machine, sequence-to-sequence long short term memory (LSTM), support vector regressor (SVR), and ANN. Another study presented in [29] proposed a deep CNN network for day-ahead load forecasting and compared the results with an extreme learning machine, ARIMA, CNN, and RNN. Several studies also used RNN models for electricity load prediction, whereas Tokgoz et al. [30] used RNN, gated recurrent unit (GRU), and LSTM models for electricity load prediction in Turkey and extensively decreased the error. Furthermore, the authors of [31] developed an LSTM-based model for periodic energy prediction and compared their results with other models. Another study presented in [32] developed an RNN-based model for medium- and long-term electricity load prediction.

The electricity consumption data is time-series data, which comprises spatial and temporal information. The CNN models perform well for spatial information extraction, but insufficient for temporal information, whereas the RNN models are insufficient for spatial information and can learn temporal information. Therefore, to develop an optimal model for electricity load prediction, hybrid models are introduced in the recent literature. For instance, Kim et al. [33] developed a hybrid model combining CNN with LSTM for short-term load prediction and compared their results with GRU, attention LSTM, LSTM, and bidirectional LSTM. Ullah et al. [34] also developed a hybrid model with a combination of CNN and multi-layer bidirectional LSTM and compared their results with bidirectional
LSTM, LSTM, and CNN-LSTM. Similarly, another study presented in [17] integrated a CNN with an LSTM auto-encoder and compared the final results with LSTM, LSTM autoencoder, and CNN-LSTM. Moreover, Sajjad et al. [35] and Afrasiabi et al. [36] presented the performance of a CNN-GRU based model for electricity forecasting. The performance of hybrid models is quite promising and has achieved state-of-the-art accuracy; however, further improvements are needed for optimal electricity load prediction. Therefore, in the current study, we established a two-step framework for predicting the electricity load that includes data preprocessing and proposed a hybrid model. In the first step, the historical data of electricity is refined to remove abnormalities that are then passed to the next step CNN along with MB-GRU model for learning. To extract the spatial information, we used CNN layers where MB-GRU is used to learn the temporal information. The contributions of the proposed research are summarized below:

• The electricity consumption data are gathered from smart meter sensors, which include missing values, redundant values, outlier values, etc., due to several reasons, such as faults in meter sensors, variable weather conditions, abnormal customer consumption patterns, etc., which need to be refined before training the model. Therefore, in this work, the input raw datasets are refined before training to fill the missing values and remove the outlier values from the dataset. Similarly, the electricity consumption patterns are of very diverse nature where the neural networks are sensitive to it, so a data normalization technique is applied to bring the dataset into a standard range.

• The mainstream methods use solo models for electricity consumption prediction, which are unable to precisely extract spatiotemporal patterns and have high error rates. Therefore, in this study, we proposed a hybrid model with a combination of CNN and MB-GRU that helps to improve the accuracy of electricity consumption prediction.

• The performance of the model was evaluated using the root mean square error (RMSE), mean square error (MSE), and mean absolute error (MAE). The experimental results show that the proposed model extensively decreases the error rate when compared to baseline models.

The main goal of this work is to improve the prediction accuracy for short-term electrical load prediction in residential buildings, which reduces customer consumption and provides economic benefits. The experimental section shows the effectiveness of the proposed method, which ensures the best performance of the proposed method as compared to other baseline models.

The remainder of the paper is arranged as follows: Section 2 provides a detailed explanation of the proposed method; Section 3 includes the experimental results of the proposed method and comparison with other state-of-the-art models. Finally, the manuscript is concluded in Section 4.

2. Proposed Framework

Accurate electricity load prediction is very important for electricity saving and vital economic implications [37]. As reported by [38] a 1% decrease in the error rate of the electricity prediction model can profit 1.6M dollars and can save 10K MW of electricity annually. For accurate load prediction, an appropriate learning methodology is required. The electricity load prediction models are learned from historical data generated from smart meter sensors. However, some times, due to weather conditions, meter faults, etc., they generate some abnormal data that should be refined before training. Therefore, this work represents a two-step framework that includes data preprocessing and the proposed hybrid model. The preprocessing step refines the input raw data of electricity consumption and then passes it to the proposed model to learn it, as demonstrated in Figure 1, where the details of each step are further discussed in the following sections.
The higher the layer count, the deeper the network and can learn tiny representations. A CNN is a particular type of deep neural network that employs alternating layers of convolutions and pooling. Preprocessed input data, and the MB-GRU model is used to learn the sequences between them.

2.1. Data Preprocessing

For better performance of the electricity load prediction models, the training data should be analyzed before training. As previously mentioned, the historical data of electricity consumption include abnormalities that affect the model performance. In this study, we used individual household electricity consumption prediction (IHEPC) and appliances load prediction (AEP) datasets, which consist of missing and outlier values. These abnormalities are removed from the data in the preprocessing step of the proposed framework. For missing values filling, NAN interpolation techniques are used, whereas for outlier values, three sigma rules of thumb [39] are applied. After the outlier reduction and filling missing values, the datasets are normalized using the min–max normalization technique to transform the dataset into a particular range which the neural network can learn easily.

2.2. Proposed CNN and MBGRU Architecture

This work combines a CNN with a multilayered bidirectional GRU (MB-GRU) for short-term electricity load prediction, where the CNN layers are incorporated to extract features from the preprocessed input data, and the MB-GRU model is used to learn the sequences between them. CNN is a neural network architecture that learns in a hierarchical manner whereby each layer learns more and more abstract features. The first layers learn atomic/primitive representations, while the intermediate-level layers learn intermediate abstract representations, and finally, fully connected layers learn high-level patterns. Therefore, the depth of the network is defined by the number of such layers. The higher the layer count, the deeper the network and can learn tiny representations. A CNN is a particular type of deep neural network that employs alternating layers of convolutions and pooling. It contains trainable filter banks per layer. Each individual filter in a filter bank, which is called a kernel and has a fixed receptive field (window) that is scanned over a layer below it, to compute an output feature map. The kernel performs a simple dot product and bias computation as it scans the layer below it and then feeds the result through an activation function, a rectifier, for example, to compute the output map. The output map is then subsampled using sum or max pooling, the latter being more common in order to reduce sensitivity to distortions in the upper layers. This process is alternated up to some point when the features become specific to the problem at hand. Thus, the CNN is a deep neural network for learning increase and more compact features that can later be used for recognition problems. The last few layers in a typical CNN comprise a typical fully connected neural network or support vector machines in order to recognize different combinations.
network or support vector machines in order to recognize different combinations of features from the convolutional layers. The CNN architecture is used in different domains, such as image and video recognition [17,35,40,41], language processing [42,43], electricity load forecasting [44,45], crowd counting [46], etc. In the time series domain, the CNN layers are used to extract spatial information and then pass the output into sequential learning algorithms such as RNN LSTM and GRU.

RNN [47] is a sequence learning architecture with backward connections among hidden layers that include some kind of memory and is extensively used in several domains such as natural language processing [48], time series analysis [49], and speech recognition [50], visual data processing [51–53], etc. The RNN models generate output at each time stamp from the input data, which leads to the vanishing gradient problem. The RNN model forgets the long sequence of electricity data, such as 60-min resolution, which leads to loss of important information.

\[ f_t = \sigma(\hat{\omega} \cdot [h_{t-1}, a_t] + \theta_f) \]  
\[ i_t = \sigma(\hat{\omega}_i \cdot [h_{t-1}, a_t] + \theta_i) \]  
\[ c_t = \tanh(\hat{\omega}_c \cdot [h_{t-1}, a_t] + \theta_c) \]  
\[ C_t = f_t \ast C_{t-1} + i_t \ast c_t \]  
\[ O_t = \sigma(\hat{\omega}_O \cdot [h_{t-1}, a_t] + \theta_O) \]  
\[ h_t = O_t \ast \tanh(C_t) \]  

(1)  
(2)  
(3)  
(4)  
(5)  
(6)

The problem of losing long sequence information is addressed by LSTM using the three-gate mechanism input, output, and forget. The mathematical representation of each gate is shown in Equations (1)–(6). In these equations the output of each gates is represented through “\(i\)”, “\(f\)” and “\(O\)” where “\(O\)” represents the activation function. The weights of the gates are represented through “\(\hat{\omega}\)”, whereas “\(h_{t-1}\)” refers to the output of previous LSTM block and “\(\theta\)” represents the bias of the gates. The LSTM structure is complex and computationally expensive due to these gates’ units and memory cells. To overcome the concern of LSTM, another lightweight architecture is developed called GRU [54] which comprises the reset and update gates. The mathematical representation of GRU gates are shown in Equations (7)–(10) where the update gate examines the earlier cell memory to remain active,

\[ x_t = \sigma(\hat{\omega}_x \cdot [h_{t-1}, a_t] + \theta_x) \]  
\[ r_t = \sigma(\hat{\omega}_r \cdot [h_{t-1}, a_t] + \theta_r) \]  
\[ h_t = \tanh(\hat{\omega}_h \cdot [r_t \ast h_{t-1}, a_t] + \theta_h) \]  
\[ h_t = (1 - x_t) \ast h_{t-1} + x_t \ast h_t \]  

(7)  
(8)  
(9)  
(10)

and the reset gate merges the next cell input sequence with previous cell memory. In this study, we used MB-GRU, which processes the sequence of input data in both backward and forward directions [55]. The bidirectional RNN models perform better in several domains such as classification, summarization [56], and load forecasting [57]. Therefore, in this study, we incorporate bidirectional GRU layers that contain both backward and forward layers, where the output sequence of the foreword layer is iteratively calculated through input in the positive sequence. The output of the backward layer is calculated through the reverse of the input.

The electricity consumption patterns include spatial and temporal features. Some researchers deployed a solo model that is insufficient to extract both types of features at a time. Therefore, in this work, we established a hybrid model that combines CNN with MB-GRU, as shown in Figure 1b. The proposed hybrid model includes an input layer, CNN layers, and bidirectional GRU layers. Two CNN layers are incorporated after several experiments over different layers and different parameters. Finally, we select filters of 8 and 4 for the first and second CNN layers with a kernel size.
of 3.1, respectively and used ReLU as an activation function in these layers. After convolutional layers, two bidirectional GRU layers are incorporated to learn the temporal information of the electricity historical data. Finally, the fully connected layers are integrated for the final output prediction.

3. Results and Discussion

In this section, we provide a detailed description of the dataset, evaluation metrics, and experimentation over the IHEPC and AEP datasets, and compare them with other baseline models. The model was trained over a GeFore GTX 2060 GPU with 64 GB RAM using the Keras framework with backend TensorFlow.

3.1. Datasets

The model’s performance is assessed on two benchmark datasets, AEP and IHEPC [58,59]. The AEP dataset was recorded in 4.5 months in a residential house in 10 min resolution. This dataset comprises 29 various parameters of weather information (wind speed, humidity, dew point, temperature, and pressure), light, and appliance energy consumption, as presented in Table 1. The data samples were collected from both indoor and outdoor environments through a wireless sensor network. The outdoor data are collected from nearby airport. The building includes 9 indoor and 1 outdoor temperature sensors, 9 humidity sensors in which 7 are integrated with indoor environment and one is in outdoor environment. The outdoor pressure, visibility, temperature, humidity, and dew point are recorded nearby airport region. The IHEPC dataset includes 9 parameters which are; date, time, voltage, global-active-power (GAP), intensity, global-reactive-power (GRP) and three sub-metering as shown in Table 2. The dataset was recorded in a residential house in France during 2006 and 2010 for one-minute resolution.

Table 1. AEP dataset variables, a short description and its units.

| S.no | Data Features | Units |
|------|---------------|-------|
| 1    | Appliances: total energy consumption by appliances. | Wh    |
| 2    | Light: total energy consumption by lights. | Wh    |
| 3    | T1: demonstrate the temperature of kitchen. | C     |
| 4    | RH1: demonstrate the humidity in kitchen. | %     |
| 5    | T2: demonstrate the temperature of living room. | C     |
| 6    | RH2: demonstrate the humidity of living room. | %     |
| 7    | T3: demonstrate the temperature of laundry room. | C     |
| 8    | RH3: demonstrate the humidity of laundry room. | %     |
| 9    | T4: demonstrate the temperature of office room. | C     |
| 10   | RH4: demonstrate the humidity of office room. | %     |
| 11   | T5: demonstrate the temperature of bathroom. | C     |
| 12   | RH5: demonstrate the humidity of bathroom. | %     |
| 13   | RH6: demonstrate the outside temperature of building. | C     |
| 14   | RH6: demonstrate the outside humidity of building. | %     |
| 15   | T7: demonstrate the temperature of ironing room. | C     |
| 16   | RH7: demonstrate the humidity in ironing room. | %     |
| 17   | T8: demonstrate the temperature of teenager room. | C     |
| 18   | RH8: demonstrate the humidity of teenager room. | %     |
| 19   | T9: demonstrate the temperature of parent room. | C     |
| 20   | RH9: demonstrate the humidity of parent room. | %     |
| 21   | To: demonstrate the outside temperature which are collected from Chievres Weather Station (CWS). | C     |
| 22   | Pressure: outside pressure which are collected from CWS. | Mm Hg |
| 23   | Rho: demonstrate the outside humidity from CWS. | %     |
| 24   | Wind speed: outside wind speed which are collected from CWS. | m/s   |
| 25   | Visibility: outside visibility from CWS. | Km    |
| 26   | Tdewpoint: outside Tdewpoint from CWS. | C     |
3.2. Metrics of Evaluation

To evaluate the performance of the model, we used RMSE, MSE, and MAE metrics. The mathematical representation of these metrics is depicted in Equations (11)–(13). RMSE calculates the difference between all predicted data points and the actual data point, then compute the mean of these square errors and finally calculate the square root of the mean values. The MSE calculates the mean disparity between the actual and model output values. The MAE calculates the mean absolute difference between the actual and predicted values.

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{11}
\]

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{12}
\]

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{13}
\]

3.3. Experimentations over IHEPC, AEP Dataset and Comparison with other Models

In this section, we evaluate the performance comparison of the proposed model with existing models for short-term load prediction (one hour ahead) over the IHEPC and AEP datasets. For the IHEPC dataset, the proposed model achieved 0.42, 0.18, and 0.29 RMSE, MSE, and MAE, respectively. The performance prediction over the test data is displayed in Figure 2a. A comparison of the proposed model over IHEPC dataset for short-term load prediction with other baseline models is shown in Figure 3. For more detail the performance of the proposed model is compared with [1,17,33–35,60,61] in the short-term horizon. In [60] the authors used deep learning methodology for residential load prediction and obtained 0.79 RMSE and 0.59 MAE, whereas [33] used a CNN-LSTM hybrid network for short-term residential load prediction and achieved 0.59, 0.35, and 0.33 RMSE, MSE, and MAE, respectively. 0.47, 0.19 and 0.31 RMSE, MSE and MAE was reported in [17] whereas [34] reports 0.56, 0.31 and 0.34 values for these metrics. In [61] the authors achieved 0.38 MSE and 0.39 MAE, whereas [1] reported 0.66 RMSE. Another strategy presented in [35] attained 0.47, 0.22, and 0.33 RMSE, MSE, and MAE, respectively. Among these results, the proposed model achieved the lowest error rate for short-term electric load prediction.

| S.no | Data Features          | Units  |
|------|------------------------|--------|
| 1    | Date                   | dd/mm/yyyy |
| 2    | Time                   | hh: mm: ss |
| 3    | GAP                    | Kw     |
| 4    | GRP                    | Kw     |
| 5    | Voltage                | V      |
| 6    | Global intensity       | Amp    |
| 7    | Submetering-1          | W      |
| 8    | Submetering-2          | W      |
| 9    | Submetering-3          | W      |
Furthermore, the effectiveness of the proposed model is evaluated over the AEP dataset for a short-term horizon. For the AEP dataset, the proposed model attained 0.31, 0.10, 0.33 \text{RMSE}, \text{MSE}, \text{and MAE}, respectively, whereas the prediction results are shown in Figure 2b. Similarly, the effectiveness of the proposed model over the AEP dataset is also compared with other baseline models, as shown in Figure 4. For instance, the results are compared with [34,62–64]. For further details, Ref. [62] achieved 0.59 \text{RMSE} and 0.26 \text{MSE}, Ref. [63] achieved 0.35 \text{RMSE} and 0.66 \text{MSE}, and [64] achieved a 0.59 \text{RMSE}. In [35], authors achieved 0.31, 0.09, and 0.24 scores for \text{RMSE}, \text{MSE}, and \text{MAE}, respectively.
and MAE, respectively. Compared to these models, the proposed model performed better in reducing the RMSE and MAE error rate, while only the results of Sajjad et al. [35] in terms of MSE is better than the proposed model in the short-term horizon.

![Figure 4. Performance contrast of the proposed model with the other baseline models over AEP dataset.](image)

4. Conclusions

In this study, we established a two-step methodology for short-term load prediction. In the first step, we performed data preprocessing over raw data to refine it for training. The refinement of the data is important because the historical data of energy consumption is generated from smart meter sensors, which include abnormalities such as outliers, missing values etc. These abnormalities from the raw data are extracted in this step, and finally, the normalization technique is applied to transform the data into a specific range. The second step is the hybrid model, which is a combination of CNN and multilayered bi-directional GRU (MB-GRU). The CNN layers are incorporated to extract important features from the refined data, while the MB-GRU layers are used to learn the temporal information of electricity consumption data. The proposed methodology is tested over two challenging datasets and achieves better performance when compared to other methods, as demonstrated in the results section.

**Author Contributions:** Conceptualization, Z.A.K.; methodology, Z.A.K.; software, Z.A.K.; validation, W.U. and A.U.; formal analysis, M.Y.L.; investigation, A.U.; resources, S.W.B.; data curation, Z.A.K.; writing—original draft preparation, Z.A.K.; writing—review and editing, Z.A.K., W.U. and A.U.; visualization, W.U.; supervision, S.W.B.; project administration, M.L.; funding acquisition, S.W.B. and S.R. All authors have read and agreed to the published version of the manuscript.

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

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. Mocanu, E.; Nguyen, P.H.; Gibescu, M.; Kling, W.W. Deep Learning for Estimating Building Energy Consumption. *Sustain. Energy Grids Netw.* 2016, 6, 91–99. [CrossRef]
2. Guo, Z.; Zhou, K.; Zhang, C.; Lu, X.; Chen, W.; Yang, S. Residential Electricity Consumption Behavior: Influencing Factors, Related Theories and Intervention Strategies. *Renew. Sustain. Energy Rev.* 2018, 81, 399–412. [CrossRef]
3. IEA. *International Energy Outlook*; IEA: Paris, France, 2014; Volume 18.
4. Nejat, P.; Jomehzadeh, F.; Taheri, M.M.; Gohari, M.; Majid, M.Z.A. A Global Review of Energy Consumption, CO 2 Emissions and Policy in the Residential Sector (with an Overview of the Top Ten CO 2 Emitting Countries). *Renew. Sustain. Energy Rev.* 2015, 43, 843–862. [CrossRef]
5. Amarasinghe, K.; Wijayasekara, D.; Carey, H.; Manic, M.; He, D.; Chen, W.-P. Artificial Neural Networks Based Thermal Energy Storage Control for Buildings. In Proceedings of the 41st Annual Conference of the IEEE Industrial Electronics Society, Yokohama, Japan, 9–12 November 2015; pp. 005421–005426.

6. Yang, A.; Li, W.; Yang, X. Short-Term Electricity Load Forecasting Based on Feature Selection and Least Squares Support Vector Machines. Knowl. Based Syst. 2019, 163, 159–173. [CrossRef]

7. Koprina, I.; Rana, M.; Agelidis, V.G. Correlation and Instance Based Feature Selection for Electricity Load Forecasting. Knowl. Based Syst. 2015, 82, 29–40. [CrossRef]

8. Khwaja, A.S.; Anpalagan, A.; Naeem, M.; Venkatesh, B. Joint Bagged-Boosted Artificial Neural Networks: Using Ensemble Machine Learning to Improve Short-Term Electricity Load Forecasting. Electr. Power Syst. Res. 2020, 179, 106080. [CrossRef]

9. Heydari, A.; Nezhad, M.M.; Pirshayan, E.; Garcia, D.A.; Keynia, F.; De Santoli, L. Short-Term Electricity Price and Load Forecasting in Isolated Power Grids Based on Composite Neural Network and Gravitational Search Optimization Algorithm. Appl. Energy 2020, 277, 115503. [CrossRef]

10. Zhang, C.; Li, J.; Zhao, Y.; Li, T.; Chen, Q.; Zhang, X. A Hybrid Deep Learning-Based Method for Short-Term Building Energy Load Prediction Combined with an Interpretation Process. Energy Build. 2020, 225, 110301. [CrossRef]

11. Naspi, F.; Arnesano, M.; Zampetti, L.; Stazi, F.; Revel, G.M.; D’Orazio, M. Experimental Study on occupants’ Interaction with Windows and Lights in Mediterranean Offices during the Non-Heating Season. Build. Environ. 2018, 127, 221–238. [CrossRef]

12. Stazi, F. Thermal Inertia in Energy Efficient Building Envelopes; Elsevier: Amsterdam, The Netherlands, 2017.

13. Rupp, R.F.; Ghisi, E. Assessing Window Area and Potential for Electricity Savings by Using Daylighting and Hybrid Ventilation in Office Buildings in Southern Brazil. Simulation 2017, 93, 935–949. [CrossRef]

14. Pereira, P.F.; Ramos, N.M.M. Influence of Occupant Behaviour on the State of Charge of a Storage Battery in a Nearly-Zero Energy Building. In E3S Web of Conferences; EDP Sciences: Paris, France, 2020; Volume 172, p. 16010.

15. Bot, K.; Ramos, N.M.; Almeida, R.M.; Pereira, P.F.; Monteiro, C. Energy Performance of Buildings With on-site Energy Generation and Storage—An Integrated Assessment Using Dynamic Simulation. J. Build. Eng. 2019, 24, 100569. [CrossRef]

16. Han, T.; Muhammad, K.; Hussain, T.; Lloret, J.; Baik, S.W. An Efficient Deep Learning Framework for Intelligent Energy Management in IoT Networks. IEEE Internet Things J. 2020, 99, 1. [CrossRef]

17. Khan, Z.A.; Hussain, T.; Ullah, A.; Rho, S.; Lee, M.; Baik, S.W. Towards Efficient Electricity Forecasting in Residential and Commercial Buildings: A Novel Hybrid CNN with a LSTM-AE Based Framework. Sensors 2020, 20, 1399. [CrossRef]

18. Wei, L.; Zhen-Gang, Z. Based on Time Sequence of ARIMA Model in the Application of Short-Term Electricity Load Forecasting. In Proceedings of the International Conference on Research Challenges in Computer Science, Shanghai, China, 28–29 December 2009; pp. 11–14.

19. Hong, T.; Gui, M.; Baran, M.E.; Willis, H.L. Modeling and Forecasting Hourly Electric Load by Multiple Linear Regression with Interactions. In Proceedings of the IEEE PES General Meeting, Minneapolis, MN, USA, 25–29 July 2010; pp. 1–8.

20. Al-Hamadi, H.; Soliman, S. Fuzzy Short-Term Electric Load Forecasting Using Kalman Filter. IEEE Proc. Gen. Transm. Distrib. 2006, 153, 217–227. [CrossRef]

21. Ullah, A.; Haydarov, K.; Haq, I.U.; Muhammad, S.; Rho, S.; Lee, M.Y.; Baik, S.W. Deep Learning Assisted Buildings Energy Consumption Profiling Using Smart Meter Data. Sensors 2020, 20, 873. [CrossRef]

22. Lu, Y.; Tian, Z.; Peng, P.; Niu, J.; Li, W.; Zhang, H. GMM Clustering for Heating Load Patterns in-depth Identification and Prediction Model Accuracy Improvement of District Heating System. Energy Build. 2019, 190, 49–60. [CrossRef]

23. Lahouar, A.; Slama, J.B.H. Day-Ahead Load Forecast Using Random Forest and Expert Input Selection. Energy Convers. Manage. 2015, 103, 1040–1051. [CrossRef]

24. Chen, Y.; Luh, P.B.; Guan, C.; Zhao, Y.; Michel, L.D.; Coolbeth, M.A.; Friedland, S.; Rourke, S.J. Short-Term Load Forecasting: Similar Day-Based Wavelet Neural networks. IEEE Trans. Power Syst. 2009, 25, 322–330. [CrossRef]

25. Wang, Y.; Xia, Q.; Kang, C. Secondary Forecasting Based on Deviation Analysis for Short-Term Load Forecasting. IEEE Trans. Power Syst. 2011, 26, 500–507. [CrossRef]
26. Tsekouras, G.; Hatziargyriou, N.; Dialynas, E. An Optimized Adaptive Neural Network for Annual Midterm Energy Forecasting. *IEEE Trans. Power Syst.* 2006, 21, 385–391. [CrossRef]
27. Li, S.; Wang, P.; Goel, L. A Novel Wavelet-Based Ensemble Method for Short-Term Load Forecasting With Hybrid Neural Networks and Feature Selection. *IEEE Trans. Power Syst.* 2015, 31, 1788–1798. [CrossRef]
28. Amarasinghe, K.; Marino, D.L.; Manic, M. Deep Neural Networks for Energy Load Forecasting. In Proceedings of the 26th International Symposium on Industrial Electronics (ISIE), Edinburgh, UK, 19–21 June 2017; pp. 1483–1488.
29. Khan, S.; Javaid, N.; Chand, A.; Khan, A.B.M.; Rashid, F.; Afridi, I.U. Electricity Load Forecasting for Each Day of Week Using Deep CNN. In *Advances in Intelligent Systems and Computing*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 1107–1119.
30. Tsekoz, A.; Unal, G. A RNN Based Time Series Approach for Forecasting Turkish Electricity Load. In Proceedings of the 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, 2–5 May 2018; pp. 1–4.
31. Wang, J.Q.; Du, Y.; Wang, J. LSTM Based Long-Term Energy Consumption Prediction with Periodicity. *Energy* 2020, 197, 117197. [CrossRef]
32. Rahman, A.; Srikumar, V.; Smith, A.D. Predicting Electricity Consumption for Commercial and Residential Buildings Using Deep Recurrent Neural Networks. *Appl. Energy* 2018, 212, 372–385. [CrossRef]
33. Kim, T.-Y.; Cho, S.-B. Predicting Residential Energy Consumption Using CNN-LSTM Neural Networks. *Energy* 2019, 182, 72–81. [CrossRef]
34. Ullah, F.U.M.; Ullah, A.; Haq, I.U.; Rho, S.; Baik, S.W. Short-Term Prediction of Residential Power Energy Consumption via CNN and Multi-Layer Bi-Directional LSTM Networks. *IEEE Access* 2020, 8, 123369–123380. [CrossRef]
35. Sajjad, M.; Khan, Z.A.; Ullah, A.; Hussain, T.; Ullah, W.; Lee, M.Y.; Baik, S.W. A Novel CNN-GRU-Based Hybrid Approach for Short-Term Residential Load Forecasting. *IEEE Access* 2020, 8, 143759–143768. [CrossRef]
36. Afrasiabi, M.; Mohammadi, M.; Rastegar, M.; Stankovic, L.; Afrasiabi, S.; Khazaie, M. Deep-Based Conditional Probability Density Function Forecasting of Residential Loads. *IEEE Trans. Smart Grid* 2020, 11, 3646–3657. [CrossRef]
37. Bunn, D. Forecasting Loads and Prices in Competitive Power Markets. *Proc. IEEE* 2000, 88, 163–169. [CrossRef]
38. Hobbs, B.F.; Jitprapaikulsarn, S.; Konda, S.; Chankong, V.; Loparo, K.A.; Maratukulam, D.J. Analysis of the Value for Unit Commitment of Improved Load Forecasts. *IEEE Trans. Power Syst.* 1999, 14, 1342–1348. [CrossRef]
39. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly Detection: A survey. *ACM Comput. Surv.* 2009, 41, 1–58. [CrossRef]
40. Sajjad, M.; Zahir, S.; Ullah, A.; Akhtar, Z.; Muhammad, K. Human Behavior Understanding in Big Multimedia Data Using CNN Based Facial Expression Recognition. *Mob. Netw. Appl.* 2019, 25, 1611–1621. [CrossRef]
41. Haq, I.U.; Ullah, A.; Muhammad, K.; Lee, M.Y.; Baik, S.W. Personalized Movie Summarization Using Deep CNN-Assisted Facial Expression Recognition. *Complexity* 2019, 2019, 1–10. [CrossRef]
42. Young, T.; Hazarika, D.; Poria, S.; Cambria, E. Recent Trends in Deep Learning Based Natural Language Processing [Review Article]. *IEEE Comput. Intell. Mag.* 2018, 13, 55–75. [CrossRef]
43. Mustaqeem; Kwon, S. MLT-DNet: Speech Emotion Recognition Using 1D Dilated CNN Based on Multi-Learning Trick Approach. *Expert Syst. Appl.* 2020, 114177. [CrossRef]
44. Chen, K.; Chen, K.; Wang, Q.; He, Z.; Hu, J.; He, J. Short-Term Load Forecasting with Deep Residual Networks. *IEEE Trans. Smart Grid* 2018, 10, 3943–3952. [CrossRef]
45. Wang, H.; Yi, H.; Peng, J.; Wang, G.; Liu, Y.; Jiang, H.; Liu, W. Deterministic and Probabilistic Forecasting of Photovoltaic Power Based on Deep Convolutional Neural Network. *Energy Convers. Manag.* 2017, 153, 409–422. [CrossRef]
46. Khan, N.; Ullah, A.; Haq, I.U.; Menon, V.G.; Baik, S.W. SD-Net: Understanding Overcrowded Scenes in Real-Time via an Efficient Dilated Convolutional Neural Network. *J. Real Time Image Process.* 2020, 1–15. [CrossRef]
47. Jozefowicz, R.; Zaremba, W.; Sutskever, I. An Empirical Exploration of Recurrent Network architectures. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 2342–2350.

48. Yin, W.; Kann, K.; Yu, M.; Schütze, H.J.A.P.A. Comparative Study of cnn and rnn for Natural Language processing. arXiv 2017, arXiv:1702.01923.

49. Liu, Y.; Gong, C.; Yang, L.; Chen, Y. DSTP-RNN: A Dual-Stage Two-Phase Attention-Based Recurrent Neural Network for Long-Term and Multivariate Time Series Prediction. Expert Syst. Appl. 2020, 143, 113082. [CrossRef]

50. Guo, J.; Tiwari, G.; Droppo, J.; Van Segbroeck, M.; Huang, C.-W.; Stolcke, A.; Maas, R. Efficient Minimum Word Error Rate Training of RNN-Transducer for End-to-End Speech Recognition. Interspeech 2020. [CrossRef]

51. Ullah, A.; Muhammad, K.; Hussain, T.; Baik, S.W. Conflux LSTMs Network: A Novel Approach for Multi-View Action Recognition. Neurocomputing 2020.

52. Hussain, T.; Muhammad, K.; Ullah, A.; Lee, M.; Baik, S.W. Deep LSTM-Based Sequence Learning Approaches for Action and Activity Recognition. In Deep Learning in Computer Vision; Informa UK Limited: Colchester, UK, 2020; pp. 127–150.

53. Ullah, W.; Ullah, A.; Haq, I.U.; Muhammad, K.; Sajjad, M.; Baik, S.W. CNN Features with Bi-Directional LSTM for Real-Time Anomaly Detection in Surveillance Networks. Multimed. Tools Appl. 2020, 1–17. [CrossRef]

54. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv 2014, arXiv:1412.3555.

55. Schuster, M.; Paliwal, K. Bidirectional Recurrent Neural Networks. IEEE Trans. Signal Process. 1997, 45, 2673–2681. [CrossRef]

56. Tang, X.-L.; Dai, Y.; Wang, T.; Chen, Y. Short-Term Power Load Forecasting Based on Multi-Layer Bidirectional Recurrent Neural Network. IET Gener. Transm. Distrib. 2019, 13, 3847–3854. [CrossRef]

57. Rajabi, R.; Estebsari, A. Deep Learning Based Forecasting of Individual Residential Loads Using Recurrence Plots. In IEEE Milan PowerTech; IEEE: New York, NY, USA, 2019; pp. 1–5.

58. UCI. Individual Household Electric Power Consumption Data Set. Available online: https://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption (accessed on 15 October 2020).

59. Bandic, L.; Kevric, J. Near Zero-Energy Home Prediction of Appliances Energy Consumption Using the Reduced Set of Features and Random Decision Tree Algorithms. In Advances in Intelligent Systems and Computing; Springer: Berlin/Heidelberg, Germany, 2018; pp. 146–154.