An Efficient Deep Learning Framework for Intelligent Energy Management in IoT Networks

[1]Tao Han, *Member, IEEE*, [1]Khan Muhammad, *Member, IEEE*, Tanveer Hussain, *Student Member*, Jaime Lloret, *Senior Member, IEEE*, Sung Wook Baik, *Member, IEEE*,

***Abstract*— Green energy management is an economical solution for better energy usage, but the employed literature lacks focusing on the potentials of edge intelligence in controllable Internet of Things (IoT). Therefore, in this article, we focus on the requirements of todays' smart grids, homes, and industries to propose a deep learning based framework for intelligent energy management. We predict future energy consumption for short intervals of time as well as provide an efficient way of communication between energy distributers and consumers. The key contributions include edge devices based real-time energy management via common cloud-based data supervising server, optimal normalization technique selection, and a novel sequence learning based energy forecasting mechanism with reduced time complexity and lowest error rates. In the proposed framework, edge devices relate to a common cloud server in an IoT network that communicates with the associated smart grids to effectively continue the energy demand and response phenomenon. We apply several preprocessing techniques to deal with diverse nature of electricity data, followed by an efficient decision-making algorithm for short-term forecasting and implement it over dependable resource-constrained devices. We perform extensive experiments and witness 0.15 and 3.77 units reduced MSE and RMSE for residential and commercial datasets, respectively.**

***Index Terms*— Energy management, energy forecasting, GRU, machine learning, LSTM, dependable IoT, smart grids, smart homes/industries, edge computing.**

I. INTRODUCTION

Energy management at smart grids via automated techniques for future load forecasting is an interesting area of research. Smart grids are the secure and trust-worthy locations to distribute the electric energy among diverse sets of consumers such as smart homes and industries. The electric energy retails chain includes production at power plants, distribution at smart grids, and consumption at residential [1] or commercial buildings and industrial sectors [2]. The amount of energy produced in power plants that is distributed at grids is entirely influenced by its usage at consumer side. Majority of the consumers are non-experts of energy demands from electric grids, resulting financial loss and futile energy expenditure. Similarly, the producers want to minimize the cost and obtain an optimized level of energy generation, farming the need of appropriate scheduling and management strategies.

A proper planning for energy production and consumption ensures its purposeful usage at industries/household and a balanced amount of energy generation at power plants. The channel holding the energy communication stability between producer and consumer is smart grid that is responsible for the equilibrium state of energy for both parties [3]. Energy forecasting methods are significantly helpful in this regard that predict the future energy of a consumer and demands accordingly from the grids. Miss-prediction of energy leads to additional costs and its wastage. A loss of 10 million pounds per year is reported with an increase of 1% forecasting error in the United Kingdom in 1984 for a residential building [4]. Therefore, precise energy demand forecasting methods are required for optimal future decisions. The energy forecasting methods are in abundance with applications to household and industrial zones. The representative methods that are particularly related to the presented work are discussed in the subsequent paragraphs, while the detailed literature is covered in Section II.

The individual load forecasting systems are deployable in many daily life applications such as day-ahead residential forecasting assists in appropriate energy demands from smart grids [5]. The computationally intelligent techniques involving load forecasting play a vital role in reducing the energy crisis and contributes to the environmental greenery. Most of these methods consist of deep learning based sequential learning mechanisms such as long short-term memory (LSTM), which is the most popular in energy forecasting related methods. LSTM is a type of recurrent neural network (RNN) that is widely used in many computer vision domains such as video analytics for sequence and series learning tasks [6]. Despite the usage of LSTMs, hybrid approaches incorporating fuzzy neural inference systems with genetic algorithms are part of energy forecasting related literature. Different from the aforementioned strategies, T.-Y. Kim and S.-B. Cho [7] introduced the usage of spatial and temporal features assimilated together for effective housing energy consumption prediction. The authors proved the supremacy of convolutional neural networks (CNNs) to extract the representative features of different variables that affect the

[1]The authors contributed significantly and considered as co-first authors.

Manuscript received March 19, 2020; Accepted: XXX, Published: XXXX. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019M3F2A1073179). This work has been partially supported by the "Ministerio de Economía y Competitividad" in the "Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia, Subprograma Estatal de Generación de Conocimiento" within the project under Grant TIN2017-84802-C2-1-P. This work has also been partially supported by European Union through the ERANETMED (Euromediterranean Cooperation through ERANET joint activities and beyond) project ERANETMED3-227 SMARTWATIR. This paper was recommended by Associate Editor XYZ.
(*Corresponding author: Khan Muhammad and Jaime Lloret*)

T. Han is with DGUT-CNAM Institute Dongguan University of Technology, Dongguan, China (e-mail: hant@dgut.edu.cn).

Khan Muhammad is with the Department of Software, Sejong University, Seoul 143-747, South Korea. (Email: khan.muhammad@ieee.org)

Tanveer Hussain and Sung Wook Baik are with Intelligent Media Laboratory, Digital Contents Research Institute, Sejong University, Seoul 143-747, South Korea (Email: tanveer445@ieee.org and sbaik3797p@gmail.com)

J. Lloret is with the Universitat Politecnica de Valencia, 46022 València, Spain (e-mail: jlloret@dcom.upv.es).

energy consumption prediction. Furthermore, these representative features with CNNs degrade the error rates over individual household power consumption dataset.

A thorough study of the employed load forecasting related literature leaves several open challenges for future research. The most prominent and challenging task while presenting a novel energy prediction technique is achieving exactness in the forecasting accuracy. Another big challenge that is inadequately covered in the employed literature is execution of the implemented algorithm over the edge nodes that leads to fruitful communication between interconnected devices in an IoT network for energy utilization. Recently, resource-constrained devices in IoT environments have shown high-level of potentials in video analytics [8], healthcare [9], and many other domains [10]. In continuation to these challenges, the reduced time complexity of an energy forecasting method, particularly while dealing the problem of short-term load forecasting is also a primary concern. Furthermore, the cloud [11] and fog computing [12, 13] paradigms are scarcely utilized in energy forecasting literature, which are trustworthy platforms for efficient Big Data analysis and instant decision making, such as anomalous energy demand prediction. Therefore, to handle these issues efficiently and effectively in controllable IoT networks by using deep learning strategies, we propose a novel edge-intelligence based energy forecasting framework for smart grids energy management with the following summarized contributions:

- We handle energy demand fluctuations via dependable edge intelligence-based novel and adaptable framework to bring the energy producers and consumers to a common platform for effective communication based on future predictions of our employed algorithm.

- We present an infrastructure to deploy resource-constrained controllable devices at variable consumer locations (smart homes or industries), that are connected through IoT network with cloud supervising server to upload their current demands and inform about the future requirements. Smart grids respond to the domestic and industrial requests received from cloud server and transmits the specific amount of energy, ensuring smooth energy management. Cloud server filters out each demand to report about the anomalous energy demands from consumers. It has a bonus of energy forecasting data storage that can be used for further in-depth analysis.

- Based on our extensive experiments, we prove our framework to be as a paradigm for future edge-intelligence based energy forecasting methods. The initial experiments include normalization technique selection, choosing optimal sequential model, where we demonstrate the performance of our framework relative to each model. We analyze the execution time of different flavors of the series-learning models to gauge between the running time and preciseness of a model.

The rest of the paper has three major sections. Section II explains the state-of-the-art methods for intelligent load forecasting. The proposed methodology and functionalities of our framework are given in Section III. The experimentation details and performance evaluation are explained in Section IV. The overall research is concluded in Section V with some future research directions.

II. Related work

This section has two major sub-sections; (1) statistical methods, and (2) deep learning based strategies. Energy load forecasting related literature is very old and can be studied in detail from a survey [14] that covers research articles from 1956 to 2013. Similarly, a recent survey is presented by Fallah et al. [15] with coverage of 52 papers in the range of 2001~2019. The energy forecasting methods during the given tenure [16, 17] lack focusing on the usage of resource-constrained devices, which are emerging due to their computational capabilities and instant decision support system. The subsequent sections discuss these methods in a classified format i.e., statistical and deep learning based load forecasting methods.

A. Statistical approaches towards load forecasting

Statistical methods such as set theories [18] etc. are widely used for many applications, such as energy forecasting and are observed in comparatively old related literature [19]. The major techniques include clustering [20], support vector regression (SVR) [21], extreme learning machine (ELM) [22], etc. The center of focus for majority of the forecasting methods is short-term load forecasting (STLF). For instance, Ceperic et al. utilized SVR machines to predict the future load for short-term duration [23]. In this paper, authors introduced two significant improvements over the existing SVR based forecasting techniques. The first advancement is the mechanism for generation of model inputs and the second one is its subsequent model input selection by utilizing feature selection algorithms. In this work, authors employed particle swarm global optimization to optimize the SVR hyper-parameters which in turn reduces the operator interaction. This research methodology is tested over two load forecasting datasets and a fair comparison with state-of-the-art indicates their improved accuracy. In another followed research for STLF, Li et al. forecasted energy by wavelet transform and evolutionary ELM [22]. The presented strategy is not entirely dependent on ELM, rather it is a hybrid strategy of ELM and a modified artificial bee colony algorithm that forecasts for 1 to 24 hours ahead. The artificial bee colony algorithm is used to support the ELM in selection of best parameters from given input weights. The authors achieved new state-of-the-art results on electric utility data from ISO New England and North America.

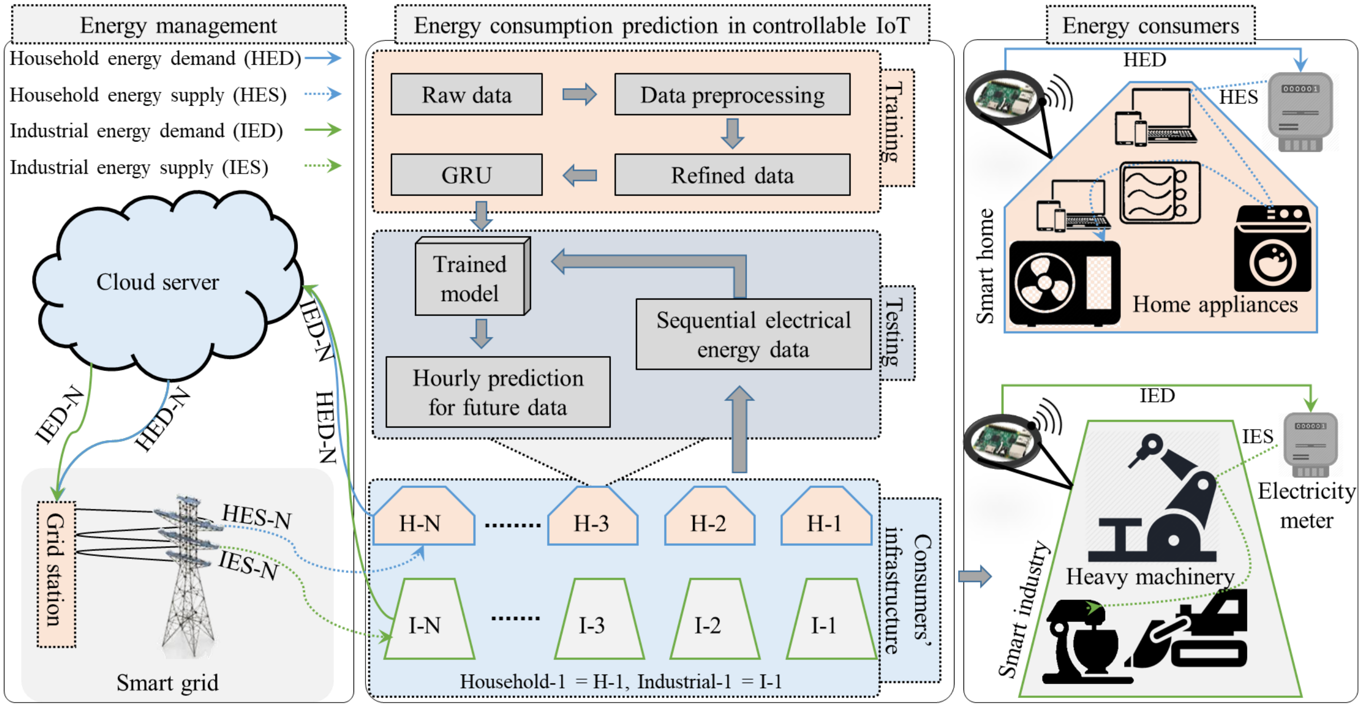


Figure 1: The proposed framework for intelligent energy management using dependable and controllable IoT with energy distribution, depletion, and management.

B. *Deep learning based methods for future load prediction*

With the emergence of deep learning in computer vision, IoT [24], security, healthcare [25], etc., scientists also utilized it for energy forecasting [26, 27] to achieve better and precise prediction results. Mainstream deep learning based methods in energy forecasting related literature focus on prediction for residential buildings, as Kong et al. proposed STLF using resident behavior learning and LSTM [28]. They mainly focused on handling the variable behavior of residential loads that hinder the precise prediction results. Another followed research [29] presented a hybrid technique for energy forecasting of residential buildings, where they incorporated deep learning and genetic algorithms with LSTM to propose an optimized objective function with hidden neurons for energy forecasting. Their method is tested over residential and commercial buildings data for VSTLF prediction, and the results are dominant over existing conventional prediction models. Wu et al. utilized multiple kernel learning based transfer regression method for load forecasting and performed experiments over residential buildings data to show the large margin of decreased error rate [30]. Similarly, a recent research [7] utilized CNN and LSTM and [31] implemented ensemble structures via wavelet neural networks for STLF. The deep learning based literature for energy forecasting is dense with major focus on sequential data processing techniques such as RNN and LSTMs. Till date, the sequential learning models are not transformed to the edge nodes with significant accuracy. Therefore, to handle this problem, we present an energy forecasting framework that is functional over resource-constrained devices. The explanatory details about our framework are given in Section III.

III. Efficient Multi-Layer GRU for Load Forecasting

The overall framework is given in **Figure 1**, where two major tiers and the energy consumers scenario at industrial and residential sectors are separately described. First tier depicts the energy management with household and industrial demand and supply. The resources (such as windmill, solar plants, etc.) provide energy to grid stations, where it is distributed among several types of consumers, primarily residential and industrial zones. The energy management tier is entirely responsible for energy consumption prediction and its appropriate management, where a cloud server is involved as a third-party communicator between consumers and smart grids. The cloud server contains demands from household and industries that are stored, analyzed, and forwarded to the grid station for energy supply to the respective consumer. The energy consumption prediction tier has a central role in our framework, where the consumer parties are equipped with a resource-constrained device for future energy prediction. Energy production resources and their related details are out of the scope of this paper and we assume the grid station to receive enough energy from the given resources.

A. *Energy management via controllable IoT devices*

A grid is a secure location to distribute the electrical energy among consumers with varied attributes such as level of consumption. A smarter grid with appropriate energy management (distribution) mechanism saves energy wastage and its extra depletion. Traditional grids openly supply energy to the demanding customers, without any information about their usage, climate changes, and many other situations that yield in poor utilization of energy. On the other hand, a smart grid keeps track of the energy demands and distributes it accordingly. But most of the times, grids show poor performance as they are overloaded or most prominently the grids do not preserve energy demands related data. Therefore, provide no mechanism to detect anomalous energy demand from residential or commercial sector. This issue is tackled in our framework through an intermediate cloud analysis concept, where the demands from consumers undergo certain analysis steps before they are passed on to the smart grids.

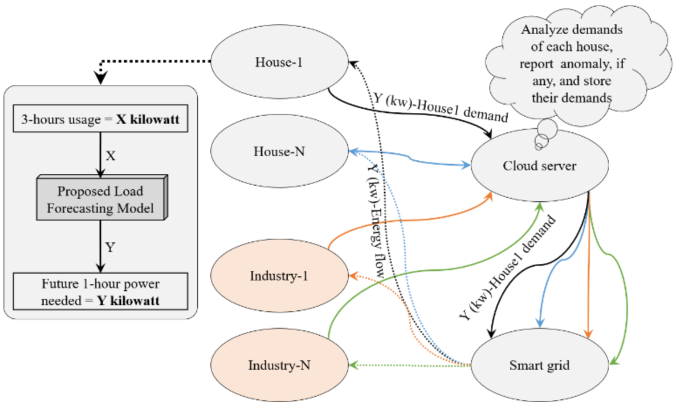


Figure 2: Sample scenario to portray the proposed energy management system using dependable resource-constrained devices.

A sample scenario of energy management using our proposed infrastructure is portrayed in **Figure 2**, where the future load forecasting advanced to its demand transmission and energy acquisition is performed for "House-1". The figure has different colors for distinction of demands of each house and industry. In **Figure 2**, solid line shows energy demand and the dotted ones represent the energy supply from the smart grid, respective to each color for a different location. The energy data usage in minutes for "House-1" is given as an input to our proposed trained model that outputs the energy usage for the future 1-hour. House-1 has trained forecasting model embedded over the resource-constrained device. It gives the input data for 3-hours (X kilowatt) and the trained model predicts future 1-hour usage, termed as "Y". House-1 transmits the request to cloud server which saves it and analyze the demand with previous history for abnormality check, and optimally transmits it to the smart grid. The abnormality may refer to sudden fluctuation in demand from residential building or an industry. Smart grid responds to the request and supplies Y- kilowatt energy to House-1. This cycle continues for all the houses and industries and rotates smoothly due to the fast processing over cloud server.

B. *Energy consumption prediction*

The technical contributions of our framework are the future energy prediction using a resource-constrained device with reduced error rate and optimized computation. There are several steps involved in achieving the final trained model that is functional in real-world scenarios. The first step is preprocessing raw data of an existing dataset, followed by our novel sequential learning mechanism to obtain the optimum trained model, as explained below.

1) *Data preprocessing*

Electric energy data contain several parameters such as date, time, active and reactive power, voltage, etc. that are involved in data recording via smart meters. The smart meter acts as a hub to connect the wires of different appliances or machineries in a single main board. Normally, the data is collected for a month or year, where it has several issues such as redundancy, missing values, long ranged parameters, etc. These errors are caused due to defects in measuring device, climate change, metering problems, and individuals' mistakes. Thus, the electric energy data need cleansing and data normalization techniques for better refinement and appropriate results.

In our framework, we apply several preprocessing techniques to purify the data for training purposes. Firstly, we remove the missing values and extract the purposeful data. Second, we perform outlier detection, prior to normalization method. It has a key advantage of ignoring the exceptional odd digits that may affect the range of normalization values and drag the parameters toward maximal or minimal range. The next important preprocessing step is normalization, where we applied several techniques before preceding to the optimal "standard transform selection" for final experiments. These normalization techniques include minmax scalar, standard scalar, maxabs scalar, quantile transform, and power transformer. The transition effect of data after normalization is visualized for the residential parameters in **Figure 3**, where the data in range of 0 to 250 is normalized between -2.5 and 3.5. The majority of the parameter values in normalized data lie between -1 and 1, therefore, it can play a significant role in precise model training. Finally, we convert the original datasets (residential and commercial) into shorter intervals because we are dealing with short-term load forecasting. The preprocessing techniques over raw format of data for both the datasets results in enhanced prediction performance.

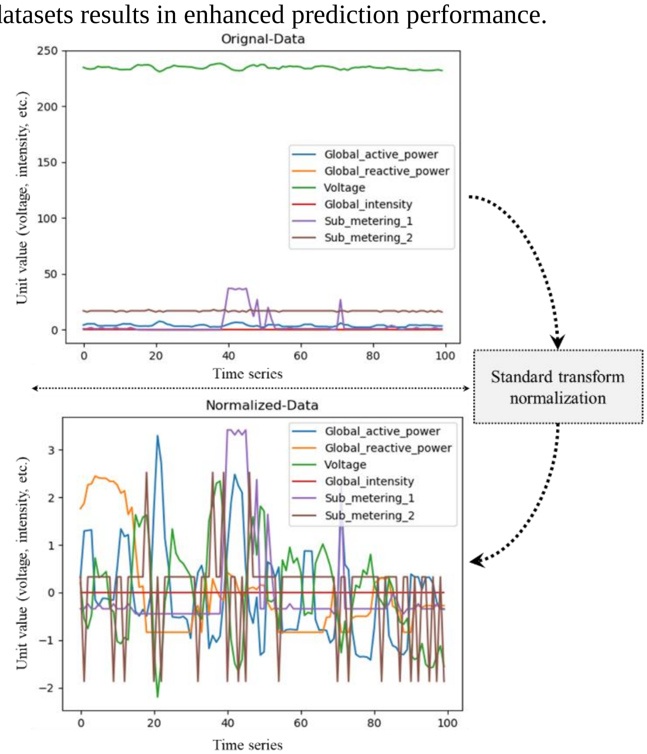


Figure 3: Visualization of residential dataset parameters; before and after applying the optimal normalization technique.

2) *Proposed sequential load forecasting model*

The trending sequential learning neural networks used in the employed energy forecasting literature are RNNs and

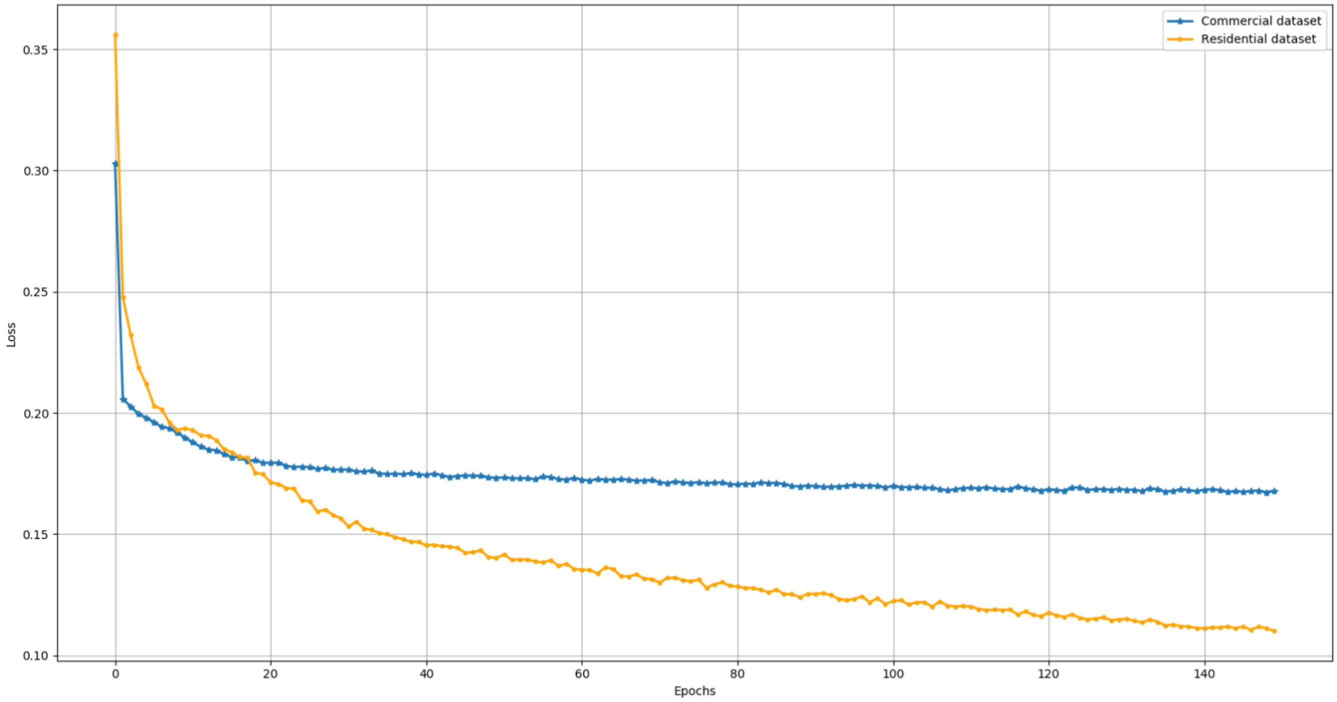


Figure 4: Training loss of our proposed multi-layered GRU over commercial and residential dataset. The training loss over commercial dataset is better as compared to residential due to the parameters difference between these datasets.

$$\mathrm{i_t} = \sigma[\mathrm{w_i}[\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{b_i}] \tag{1}$$

$$\mathrm{f_t} = \sigma[\mathrm{w_f}[\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{b_f}] \tag{2}$$

$$\mathrm{o_t} = \sigma[\mathrm{w_0}[\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{b_0}] \tag{3}$$

LSTMs. Traditional neural networks consider only a single input, while RNNs [32], in contrast, take input at multiple time steps and analyze the series of patterns. The RNNs take input and generate output at each time stamp, therefore, it encounters vanishing gradient problem i.e., forgetting the effect of a longer sequence. The RNNs always face hard times while carrying information from earlier time stamps in long-lasting sequential information. For instance, processing a long sequence of energy raw data will lead in losing some important information from the initial sequences. This problem is solved by LSTMs, which has several gates (input, forget, and output gates) to learn long-term sequential information as shown mathematically in **Eq. 1** to **3.** In these equations, “i_t”, “f_t”, “o_t” are input, forget, and output gates, respectively. “σ” refers to the sigmoid function, which is used to coerce the output between 0 and 1. “w_i”, “w_f”, and “w_0” are the weights of the corresponding gates, “h_{t-1}” indicates the output of the previous LSTM block at varied timestamp (t), “x_t” shows the input at the ongoing timestamp. Finally, “b_i”, “b_f”, and “b_0” are the biased terms for the respective gates i.e., input, forget, and output gates, correspondingly. The structure of the LSTMs is more complex and yields in huge processing complexity due to the presence of gated recurrent units and memory cell, working together to achieve final output. Another effective yet efficient solution to this problem is gated recurrent neural network (GRU) [33], that contains only two gates; reset and update gate with an activation unit.

To simplify the mathematics behind the GRU, suppose an update gate “$\mathrm{U_t}$” at time duration “$\mathrm{t_d}$”. When any input “i” is fed into the network with time “$\mathrm{i_t}$”, it is then multiplied with its own weights, given as “$\mathrm{W_1}$” and the same process continues for “$\mathrm{i_{t-1}}$” that is the previous unit and is multiplied by its own weight “$\mathrm{W_2}$”. A sigmoid activation function is applied on their resultant sum to acquire the output value of the update gate between 0 and 1, as given in **Eq. 4**.

$$U_t = \sigma[(W_1 \times i_t) + (W_2 \times i_{t-1})] \tag{4}$$

Following this, consider a reset gate “$\mathrm{R_t}$”, the formula to compute its value is given in **Eq. 5**. It is used to decide how much of the previous information to forget.

$$R_t = \sigma[(W_1 \times i_t) + (W_2 \times i_{t-1})] \tag{5}$$

Now, to store the reset gate information, introduce a memory content “M_r'” which has information related to the past and has the following (**Eq. 6**) tangent function corresponding to the weights.

$$M_r' = tanh\,[(W_1 \times i_t) + (R_t \odot i_{t-1})] \tag{6}$$

The element-wise product between the reset gate “$\mathrm{R_t}$” and “$\mathrm{W_2}$” determines the information to be removed from the previous time stamps. The final memory at current time stamp is calculated using element-wise multiplication and sum operation, as illustrated in **Eq. 7**.

$$M_r = U_t \odot i_{t-1} + (1 - U_t) \odot M_r') \tag{7}$$

The simple structure of GRU makes it implementable in real-time over resource-constrained devices such as Raspberry-Pi.

Although some research studies [33] advocate the superiority of LSTMs for specific problems but in our framework, the multi-layered GRU dominates LSTM in terms of accuracy and computational complexity, as evident from experimental results given in Section IV. The proposed model has two stacked layers of GRU that help better learning of sequential data. In our architecture, we use 0.2 dropout after each layer of GRU. The detailed explanation of sequential learning mechanism (memory cells and gates) and its mathematical computation is out of the scope of this paper and can be deeply studied from the referred research works [6, 34]. After the stacked GRU layers, we pass its output to a dense layer for final sequential data prediction. The number of epochs used for both residential and commercial datasets are 150 and its learning is shown in **Figure 4**.

IV. Experimental Results and Discussion

We perform extensive experiments including comparison with state-of-the-art on two datasets and time complexity analysis for personal computers (PC) and resource-constrained devices. We utilize two datasets: Individual household electric power consumption [35] and Commercial dataset [36] for comparison and the results are convincing for our multi-layered GRU, as compared to recent methods in energy forecasting related literature. In this section, first, we explain the evaluation metrics used in this research work. Second, we explain the datasets utilized for experiments and provide discussion about the dominant performance of our framework. Finally, we evaluate our models' size and execution time over resource-constrained devices and PCs, as explained in the subsequent sections.

$$\text{MSE} = \frac{1}{n}\sum_{1}^{n}(\mathbf{y}-\hat{\mathrm{y}})^2 \tag{8}$$

$$MAE = \frac{1}{n}\sum_{1}^{n}|\boldsymbol{y}-\hat{\mathrm{y}}| \tag{9}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{1}^{n}(y-\hat{y})^2} \tag{10}$$

$$\text{MAPE} = \frac{100\%}{n}\sum_{t-1}^{n}\left|\frac{A_t-F_t}{A_t}\right| \tag{11}$$

A. Evaluation metrics

For the performance evaluation, mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) are used in our experimental results. **Eq. 8** to **11** demonstrates the mathematical formulation of these metrics. The first used metric is MSE that measures average of the squares of errors i.e., it is considered as the mean squared difference between the predicted and the actual values, as given in **Eq. 8.** Secondly, we compute MAE that is the average magnitude of the prediction errors without considering their directions. In other words, it is the average of the absolute differences between a models' prediction and its actual values for all instances in the testing set. **Eq. 9** shows the mathematical formula for MAE computation. RMSE is the standard deviation of prediction errors and is a commonly used metric in climatology, forecasting, and regression analysis used to verify the experimental models and is determined in **Eq. 10**. The last metric namely MAPE is a measure of prediction accuracy of a forecasting method such as time series prediction. This metric express accuracy in percentage, as depicted mathematically in **Eq. 11**.

B. Performance comparison with state-of-the-art methods

We compare the performance of the proposed method on competitive benchmarks using individual household electric power consumption and commercial dataset. The comparison with recent methods over residential and commercial dataset is explained in the coming subsections, where the supremacy of our proposed model is described in detail.

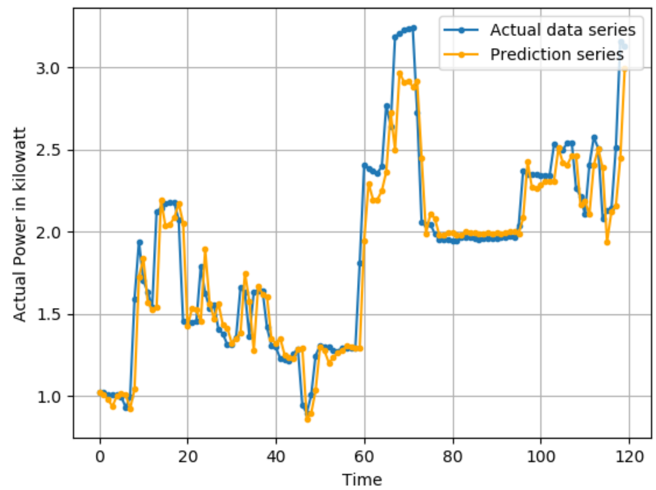


Figure 5: Visualization of our proposed GRU based trained model when compared to original values present in residential household prediction dataset. The difference between real and predicted power is very narrow, thus, the better performance of the proposed model for future load prediction is clearly observable.

1) Evaluation over Residential dataset

The actual data and our predicted results for residential dataset are plotted in **Figure 5** and the comparative graph is illustrated in **Figure 6**, where the better performance of our trained model compared to existing methods is observable over the residential load forecasting dataset. On this dataset (UCI dataset) [35], our method achieved the lowest error score compared to all the recent methods under consideration. For instance, Kim et al. proposed a novel energy load prediction methodology based on deep neural and CNN-LSTM based network and achieved 0.37, 0.34, 0.61 and 34.84 error rate for MSE, MAE, RMSE, and MAPE, respectively [7]. Another autoencoder based network introduced in [37] attained 0.21 unit MSE and 0.25 value for MAE. A followed research by Wu et al. reduced the MAPE error rate for the same dataset up to 73.07 (between 1 and 100, non-normalized) using MKL regression [30], that is normalized between 0 and 1 in **Figure 6**. In contrast to these methods, the proposed GRU model achieved the lowest error rates of 0.17, 0.19, and 0.22, for MSE, MAE, and RMSE, respectively. Similarly, the MAPE of the proposed model is 60, that is normalized to the range of 0 and 1 and is plotted in **Figure 6** against recent state-of-the-art methods. Besides the best and accurate performance, our proposed method has lower computational complexity that is discussed in Section IV. The ground truth values graph when

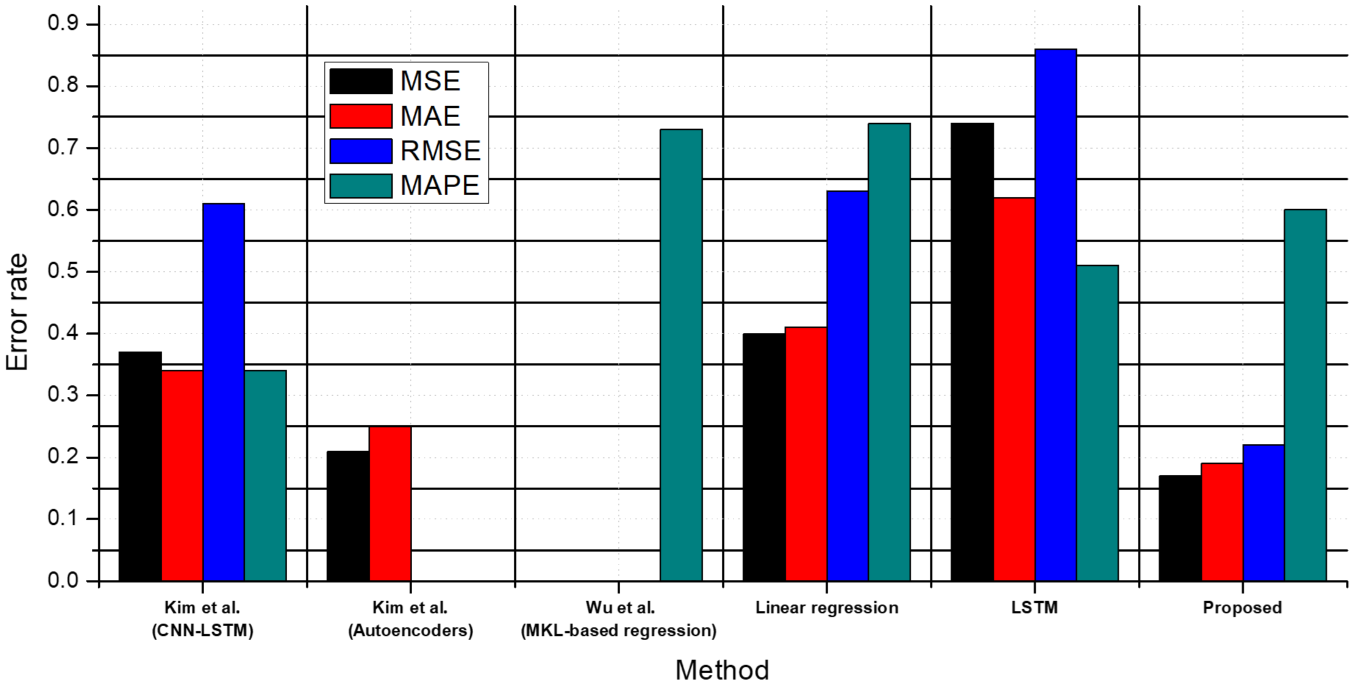


Figure 6: Comparison of the proposed model with recent existing methods and LSTM sequential forecasting over residential dataset, where our trained multi-layered GRU dominates Kim et al. (CNN-LSTM) [7], Kim et al. (Autoencoders) [37], and Wu et al. (multi-kernal learning based regression) [30] for all the used metrics. MAPE of our proposed model lags behind Kim et al. (CNN-LSTM) [7], which has higher computational complexity of CNN and LSTMs.

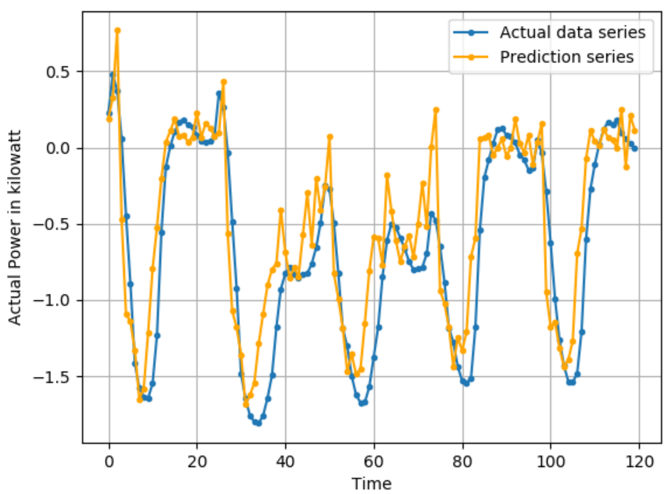


Figure 7: Visual representation of actual data and prediction output results using commercial dataset, where the data is plotted for time series in minutes against the actual power in kilowatt.

compared to the predicted power by the proposed model is given in **Figure 5** with minor observable variations between both the values (ground truth and prediction), indicating effective real-world deployment of the proposed model.

2) Evaluation over Commercial dataset

The proposed framework has high-level of adoptability for both industrial and residential buildings, and for approval of this claim, we also made experiments over a well-known commercial dataset, "PJM hourly energy consumption dataset [36]". It is collected by a regional transmission organization in United States, known as PJM Interconnection LLC (PJM). PJM is a part of Easter Interconnection grid, which is responsible for energy supply to 14 different regions including Delaware, Illinois, Indiana, etc. The data given in this dataset are hourly and measured in megawatts, where it has coverage of the aforementioned regions and is collected between 2006 and 2018. The proposed model prediction results on this dataset against the test data ground truth values are visualized in **Figure 7**, where a slight gap is observable in time duration of 40 to 80 minutes. The rest of the values are highly overlapping, indicating the higher accuracy of the proposed model.

TABLE I: COMPARATIVE ANALYSIS OF THE PROPOSED MULTI-LAYERED GRU WITH GAO ET AL. [35], MUJEEB ET AL. [36], AND VARIOUS FLAVORS OF CONVENTIONAL METHODS PRESENTED BY [36]. IN THE ABOVE FIGURE, NARX IS NONLINEAR AUTOREGRESSIVE NETWORK WITH EXOGENOUS INPUTS, DE-ELM REFERS TO DIFFERENTIAL EVOLUTION ELM, RELM STANDS FOR RECURRENT ELM, DE-RELM REFERS TO DE RECURRENT ELM, AND ESAENARX INDICATES EFFICIENT SPARSE AUTOENCODER NONLINEAR AUTOREGRESSIVE NETWORK WITH EXOGENOUS.

Method	RMSE
ELM [38]	21.2
NARX [38]	9.26
DE-ELM [38]	9.18
RELM [38]	9.04
CEANN (Gao et al.) [39]	8.96
DE-RELM (Mujeeb et al.) [38]	5.24
ESAENARX (Mujeeb et al.) [38]	3.86
Proposed	**0.09**

After an extensive research, we compared our results with two recently published energy forecasting methods; [39] and various flavors offered by Mujeeb et al. [38]. The overall comparison is given in **Table I**, where the lowest error rates are reported by our multi-layered GRU based energy forecasting model. There are several types of data available in

PJM dataset, where we experimented over the already used sequences by [39] and [38] in their methods. In **Table I**, it is illustrated that the proposed model achieved 0.09 RMSE value over commercial [36] dataset, that is the lowest error rate when compared to recent energy forecasting methods using this dataset. Therefore, it is evident from experiments, as reported in **Table I** and **Figure 7** for actual and predicted output data, that our proposed model is malleable and can be utilized for both household and industrial sectors in real-world scenarios.

C. Time complexity analysis

Efficient time complexity of a trained model is a difficult task to achieve along with higher accuracy, particularly, when a model is implemented over resource-constrained devices. Therefore, we carry out a detailed time complexity analysis with major focus on the model size and its execution time, while considering the proposed GRU based approach as well other possibilities. Since, the employed energy forecasting literature lacks focusing on resource-restricted devices, therefore, to present a fair comparison, we analyze the execution time on both, Raspberry-Pi and PC. The tested PC for experiments has Intel(R) Core (TM) i7-7700 CPU (3.60Hz) processor with 16 GB RAM windows 10 64-bit, Python version 3.6.4, Tensorflow version 1.12.0, and Keras version 2.2.4. The Raspberry-Pi used for experiments is ARM Cortex A53 processor, with Raspbian operating system. The possible details related to time complexity analysis are given in **Table II**, which advocates that for PC and a resource constrained-device the best performance in terms of model size and execution time is shown by our multi-layered GRU. The closest match after GRU is LSTM-based forecasting model, where it has 779.6 KB model size and running time for 2-hours prediction is 6.43 seconds. The model size for CNN and Bi-directional LSTM is very huge i.e., 20336 KB and as compared to all the flavors of LSTM, the lowest execution time is 6.38 seconds and 591 KB model size. We implemented the given sequential forecasting models and computed their time complexity. The future time is predicted for coming 2-hours, where the proposed model consumes minimum time among all the given options and has the least model size with accurate results. The best performance in **Table II** is given as bold, where the time analysis proves that our proposed model fits the requirements of smart grids and can transform the forecasting problem into the edge.

TABLE II: TIME COMPLEXITY ANALYSIS OF OUR PROPOSED MODEL WHEN COMPARED TO EXISTING SEQUENTIAL LEARNING BASED ENERGY FORECASTING APPROACHES.

Method	Execution time (secs)		Model size (KBs)
	Raspberry-Pi	PC	
LSTM	N/A	16.92	779.6
CNN-LSTM		29.43	790.83
Bi-directional LSTM		19.34	1695.34
CNN-Bi-directional LSTM		59.27	2391.34
Proposed	**20.36**	**6.38**	**591**

V. CONCLUSIVE REMARKS AND FUTURE LINEATION

The influence of IoT devices for various problems is increasing on a daily basis with numerous solutions to real-world tasks. These devices are mostly used in computer vision and image processing problems for intelligent surveillance and activity recognition. The future energy prediction and its appropriate management using IoT devices is rarely studied, particularly the deep learning and its related concepts are not inferenced to the edge. In our research, we applied lightweight computationally intelligent techniques, functional over resource-constrained devices for future energy prediction, that yields in its effective management.

To this end, we investigated controllable IoT devices for energy load forecasting and presented a functional algorithm over the edge nodes in smart homes/industries. In the proposed framework, a controllable resource-constrained device is equipped with our pre-trained model for short-term load forecasting. The obtained model is trained using existing datasets via multi-layered GRU that has an efficient and accurate output prediction results. The dependable resource-constrained device predicts the future energy usage which is demanded from smart grid using the cloud server as a channel of communication. Smart grid supplies the demanded energy to that specific residential building or industry, obtained as a request from the cloud server. Thus, through our users' friendly framework, energy management has become very efficient and effective and is feasible for installation at smart homes/industries.

Besides the edge intelligence using dependable IoT, in future, the resource-constrained devices can be interconnected together in an IoT network for mutual energy sharing to fulfil each others' demand and save energy resources. Similarly, we intend to integrate sequential learning with fuzzy logics for effective real-time energy forecasting methods. Further, we aspire to study efficient set theory concepts integrated with effective CNNs using weighted fusion schemes and implementing cloud and fog computing for highly accurate and quick output predictions for weekly and monthly forecasting.

References

[1] Y. Wang, Q. Chen, M. Sun, C. Kang, and Q. Xia, "An ensemble forecasting method for the aggregated load with subprofiles," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3906-3908, 2018.

[2] Y.-F. Zhang and H.-D. Chiang, "Enhanced ELITE-Load: A Novel CMPSOATT Methodology Constructing Short-Term Load Forecasting Model for Industrial Applications," *IEEE Transactions on Industrial Informatics*, 2019.

[3] P. Zhuang and H. Liang, "Hierarchical and decentralized stochastic energy management for smart distribution systems with high BESS penetration," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6516-6527, 2019.

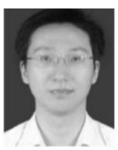
[4] G. Černe, D. Dovžan, and I. J. I. T. o. I. E. Škrjanc, "Short-term load forecasting by separating daily profiles and using a single fuzzy model across the entire domain," vol. 65, no. 9, pp. 7406-7415, 2018.

[5] Y. Huang *et al.*, "LoadCNN: A Efficient Green Deep Learning Model for Day-ahead Individual Resident Load Forecasting," 2019.

[6] 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 Transactions on Industrial Informatics,* 2019.
[7] T.-Y. Kim and S.-B. Cho, "Predicting Residential Energy Consumption using CNN-LSTM Neural Networks," *Energy,* 2019.
[8] T. Hussain, K. Muhammad, J. Del Ser, S. W. Baik, and V. H. C. de Albuquerque, "Intelligent Embedded Vision for Summarization of Multi-View Videos in IIoT," *IEEE Transactions on Industrial Informatics*, 2019.
[9] T. Hussain, K. Muhammad, S. Khan, A. Ullah, M. Y. Lee, and S. W. Baik, "Intelligent baby behavior monitoring using embedded vision in IoT for smart healthcare centers," *Journal of Artificial Intelligence and Systems. J. Artif. Intell. Syst*, vol. 1, no. 15, p. 2019, 2019.
[10] K. Muhammad, S. Khan, V. Palade, I. Mehmood, and V. H. C. De Albuquerque, "Edge Intelligence-Assisted Smoke Detection in Foggy Surveillance Environments," *IEEE Transactions on Industrial Informatics,* 2019.
[11] C. Zhu, V. C. Leung, K. Wang, L. T. Yang, and Y. Zhang, "Multi-method data delivery for green sensor-cloud," *IEEE Communications Magazine,* vol. 55, no. 5, pp. 176-182, 2017.
[12] C. Tang, X. Wei, C. Zhu, W. Chen, and J. J. Rodrigues, "Towards smart parking based on fog computing," *IEEE Access,* vol. 6, pp. 70172-70185, 2018.
[13] G. Jia, G. Han, A. Li, and J. Du, "SSL: Smart street lamp based on fog computing for smarter cities," *IEEE Transactions on Industrial Informatics,* vol. 14, no. 11, pp. 4995-5004, 2018.
[14] L. Hernandez *et al.*, "A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings," *IEEE Communications Surveys & Tutorials,* vol. 16, no. 3, pp. 1460-1495, 2014.
[15] S. Fallah, R. Deo, M. Shojafar, M. Conti, and S. Shamshirband, "Computational intelligence approaches for energy load forecasting in smart energy management grids: state of the art, future challenges, and research directions," *Energies*, vol. 11, no. 3, p. 596, 2018.
[16] Z. Shi, N. Yao, and F. Zhang, "Scheduling feasibility of energy management in micro-grids based on significant moment analysis," in *Cyber-Physical Systems*: Elsevier, 2017, pp. 431-449.
[17] J. Praß *et al.*, "Smart energy and grid: novel approaches for the efficient generation, storage, and usage of energy in the smart home and the smart grid linkup," *Smart Cities: Foundations, Principles, and Applications,* pp. 575-604, 2017.
[18] I. U. Haq *et al.*, "Movie scene segmentation using object detection and set theory," *International Journal of Distributed Sensor Networks*, vol. 15, no. 6, p. 1550147719845277, 2019.
[19] D. Jiang, P. Zhang, Z. Lv, and H. Song, "Energy-efficient multi-constraint routing algorithm with load balancing for smart city applications," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1437-1447, 2016.
[20] F. L. Quilumba, W.-J. Lee, H. Huang, D. Y. Wang, and R. L. Szabados, "Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 911-918, 2014.
[21] W.-C. Hong, "Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm," *Energy*, vol. 36, no. 9, pp. 5568-5578, 2011.
[22] S. Li, P. Wang, and L. Goel, "Short-term load forecasting by wavelet transform and evolutionary extreme learning machine," *Electric Power Systems Research*, vol. 122, pp. 96-103, 2015.
[23] E. Ceperic, V. Ceperic, and A. Baric, "A strategy for short-term load forecasting by support vector regression machines," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4356-4364, 2013.
[24] G. Dartmann, H. Song, and A. Schmeink, *Big data analytics for cyber-physical systems: machine learning for the internet of things*. Elsevier, 2019.
[25] W. Yuan, C. Li, D. Guan, G. Han, and A. M. Khattak, "Socialized healthcare service recommendation using deep learning," *Neural Computing and Applications*, vol. 30, no. 7, pp. 2071-2082, 2018.
[26] L. Hernández *et al.*, "Improved short-term load forecasting based on two-stage predictions with artificial neural networks in a microgrid environment," *Energies,* vol. 6, no. 9, pp. 4489-4507, 2013.
[27] L. Hernández *et al.*, "Artificial neural network for short-term load forecasting in distribution systems," *Energies,* vol. 7, no. 3, pp. 1576-1598, 2014.
[28] W. Kong, Z. Y. Dong, D. J. Hill, F. Luo, and Y. Xu, "Short-term residential load forecasting based on resident behaviour learning," *IEEE Transactions on Power Systems,* vol. 33, no. 1, pp. 1087-1088, 2017.
[29] A. Almalaq and J. J. Zhang, "Evolutionary Deep Learning-Based Energy Consumption Prediction for Buildings," *IEEE Access,* vol. 7, pp. 1520-1531, 2018.
[30] D. Wu, B. Wang, D. Precup, and B. Boulet, "Multiple Kernel Learning based Transfer Regression for Electric Load Forecasting," *IEEE Transactions on Smart Grid,* 2019.
[31] G. T. Ribeiro, V. C. Mariani, and L. dos Santos Coelho, "Enhanced ensemble structures using wavelet neural networks applied to short-term load forecasting," *Engineering Applications of Artificial Intelligence,* vol. 82, pp. 272-281, 2019.
[32] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of recurrent network architectures," in *International conference on machine learning*, 2015, pp. 2342-2350.
[33] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural

networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.
[34] A. Ullah, K. Muhammad, J. Del Ser, S. W. Baik, and V. Albuquerque, "Activity recognition using temporal optical flow convolutional features and multi-layer LSTM," *IEEE Transactions on Industrial Electronics*, 2018.
[35] A. B. Georges Hebrail "Individual household electric power consumption Data Set " *https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+consumption*, 2012.
[36] R. Mulla. (Accessed on 29 December 2019). Available: https://www.kaggle.com/robikscube/hourly-energy-consumption
[37] 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, 2019.
[38] S. Mujeeb and N. Javaid, "ESAENARX and DE-RELM: Novel schemes for big data predictive analytics of electricity load and price," *Sustainable Cities and Society*, vol. 51, p. 101642, 2019.
[39] W. Gao, A. Darvishan, M. Toghani, M. Mohammadi, O. Abedinia, and N. Ghadimi, "Different states of multi-block based forecast engine for price and load prediction," *International Journal of Electrical Power & Energy Systems*, vol. 104, pp. 423-435, 2019.



Tao Han received the M.E. degree in communication and information system from the South China University of Technology, Guangzhou, China, in 2006, and the Ph.D. degree in communication and information systems from Sun Yat-Sen University, Guangzhou, in 2011. He is currently a Lecturer with the DGUTCNAM Institute, Dongguan University of Technology, Dongguan, China. He has several years of experience in research, academia, and teaching. His research interests include wearable sensing technology, biomedical signal sensing and processing, body sensor networks, bio feedback, biometric security, Internet of Things, and multimodal medical data fusion

Khan Muhammad (S’16, M’18) is an Assistant Professor at Department of Software and lead researcher of Intelligent Media Laboratory (IM Lab), Sejong University, Seoul, South Korea. His research interests include medical image analysis (brain MRI, diagnostic hysteroscopy and wireless capsule endoscopy), information security (steganography, encryption, watermarking and image hashing), video summarization (single-view and multi-view), multimedia, computer vision, IoT and smart cities, and video analytics. He has published over 100 papers in peer-reviewed international journals and conferences in these research areas with target venues as IEEE COMMAG, NETWORK, TII, TIE, TSMC-Systems, IoTJ, Access, TSC, Elsevier INS, Neurocomputing, ASOC, PRL, FGCS, COMCOM, COMIND, JPDC, PMC, BSPC, CAEE, Springer NCAA, MTAP, JOMS, and RTIP, etc. He is serving as a professional reviewer for over 70 well-reputed journals and conferences including IEEE Transactions on Image Processing, IEEE Transactions on Neural Networks and Learning Systems, Elsevier Future Generation Computer Systems, Elsevier Neurocomputing, Springer Mobile Networks and Applications, Springer Nonlinear Dynamics, Springer Medical & Biological Engineering & Computing, Springer Journal of Super Computing.

Tanveer Hussain (S’19) acknowledged his degree of Bachelor's in Computer Science from Islamia College Peshawar, Peshawar, Pakistan with Gold Medal distinction. Currently, he is enrolled in joint Master and Ph.D. program at Sejong University, Seoul, Republic of Korea and serving as a Research Assistant at Intelligent Media Laboratory (IM Lab). His major research domains are features extraction (learned and low-level features), video analytics, image processing, pattern recognition, medical image analysis, multimedia data retrieval, deep learning for multimedia data understanding, single/multi-view video summarization, IoT, IIoT, and resource constrained programming. He has published several journal articles in these areas in reputed journals including IEEE TII, IoTJ, Elsevier PRL, and Wiley IJDSN. He is a student member of IEEE and providing professional review services in various reputed journals such as IEEE TII, Cybernetics, IoTJ, Elsevier PRL. For further activities and implementations, visit: https://github.com/tanveer-hussain

Jaime Lloret is an associate professor at Politechnic University of Valencia, Spain. He was Internet Technical Committee Chair during 2014–2015 and is the current Chair of IEEE 1907.1. He is the director of the Research Institute IGIC and head of the Innovation Group EITA-CURTE. He is co-Editor-in-Chief of Ad Hoc and Sensor Wireless Networks and Editor-in-Chief of Network Protocols and Algorithms. He has been General Chair of 36 international workshops and conferences.

Sung Wook Baik (M’09, SM’2015) received the B.S degree in computer science from Seoul National University, Seoul, Korea, in 1987, the M.S. degree in computer science from Northern Illinois University, Dekalb, in 1992, and the Ph.D. degree in information technology engineering from George Mason University, Far fax, VA, in 1999. He worked at Datamat Systems Research Inc. as a senior scientist of the Intelligent Systems Group from 1997 to 2002. In 2002, he joined the faculty of the College of Electronics and Information Engineering, Sejong University, Seoul, Korea, where he is currently a Full Professor and the Chief of Sejong Industry-Academy Cooperation Foundation. He is also the head of Intelligent Media Laboratory (IM Lab) at Sejong University. His research interests include computer vision, multimedia, pattern recognition, machine learning, data mining, virtual reality, and computer games.