Journal of Network Intelligence ©2021 ISSN 2414-8105 (Online)
Taiwan Ubiquitous Information Volume **6**, Number **1**, February **2021**

Short-Term Household Load Forecasting Model Based on Variational Mode Decomposition and Gated Recurrent Unit with Attention Mechanism

Fu Lin

College of Computer Science and Mathematics
Fujian University of Technology
Fuzhou,350118, China
1468978192@qq.com

Shi-Jian Liu

College of Computer Science and Mathematics
Fujian University of Technology
Fuzhou,350118, China
liusj2003@126.com

Han-Chieh Chao

Department of Electrical Engineering
National Dong Hwa University
Hualien, Taiwan
hcc@niu.edu.tw

Jeng-Shyang Pan*

Fujian Provincial Key Laboratory of Big Data Mining and Applications
Fujian University of Technology
Fuzhou 350118, China
College of Computer Science and Engineering
Shandong University of Science and Technology
Qingdao 266590, China
*Correponding Author:jengshyangpan@gmail.com

Received August 2020; revised December 2020

Abstract. *The household load is an important part of the load. Accurate short-term household load forecasting is of great help to issues such as power price formulation, demand response, or power transmission. This paper proposed a forecasting model based on Variational Mode Decomposition (VMD) and Gated Recurrent Unit (GRU) combined with self-attention mechanism (SAM) and attention mechanism (AM) for short-term household load forecasting. First, we use VMD to decompose household electricity load data, extract sub-sequences with strong periodic information, and reorganize the original data sets with sub-sequences and data pre-processed. The new data sets contain more abundant information. Secondly, GRU combined with SAM and AM forecasting model for learning and training to obtain results. Finally, RMSE and MAPE evaluation functions are used for method evaluation to verify the superiority of the proposed scheme. The RMSE and MAPE of the proposed method are 0.193 kW and 15.567%, respectively, which are more favorable than traditional GRU.*
Keywords: Variational Mode Decomposition, Gated Recurrent Unit, Self-attention mechanism, Attention mechanism

---

1. **Introduction.** The production, transmission, and consumption of electric energy are carried out at the same time. To ensure the safety and economy of power transmission and distribution, as well as the high-quality demand for power, reasonable planning of transmission lines and prediction of power loss on the demand side have become critical steps. With the increasing popularity of smart meters, each household is equipped with smart meters, which provides a large amount of reliable data for short-term household load forecasting and further promotes its development.

There are many methods for short-term load forecasting, from classic time series analysis methods to current deep neural networks [1–3]. The time series analysis methods include autoregressive integrated moving average (ARIMA) [4], random forest [5], k-nearest neighbor (KNN) [6–8], etc. However, with the rise of big data, neural networks have become one of the most popular methods for analyzing big data, and great results have been achieved in image processing, biomedicine, and text processing. In [9], it believed that user behavior will lead to a substantial increase in the utilization rate of some home appliances and there is a certain connection. The user load is analyzed and processed by a clustering method based on the density of noise, and finally, LSTM is used to predict [10].

In order to further improve forecasting accuracy, many scholars have proposed hybrid prediction models. One is to combine different forecasting models and construct a new model combining the advantages of each model. In [11], it combined with LSTM and XGBoost forward load prediction experiment, both of them weighted the prediction results through the reciprocal error method and obtained more accurate prediction results. The reference [12] uses the combination of ARIMA and ANN forecasting models for prediction, and good results are also obtained. The other is to build a forecasting model for the decomposed sub-sequences by decomposing the original signal and superimposing the prediction results of each sub-sequence to obtain the final prediction value [13,14]. In [15], VMD is used to process the original power signal, the sub-signals are decomposed and the external factors obtained through correlation analysis are combined to form the input vector of each sub-LSTM. Moreover, Bayesian Optimization Algorithm(BOA) [16–19] was used to optimize the forecasting model for the number of nodes of LSTM.

At present, the neural network model is one of the mainstream methods for short-term household load forecasting. Given the volatility and randomness of household load, the decomposition analysis of power data is helpful to improve the prediction accuracy. The attention mechanism assigns different weights to the state values of the hidden layers of each neural network so that hidden layers can better express important information. In the attention mechanism, the query comes from outside, while in the self-attention mechanism, the query comes from inside, so the self-attention mechanism works better when dealing with long sequences [20]. Aiming at the vanishing gradient problem of the traditional recurrent neural network (RNN) [21], GRU is proposed to solve this problem [22]. To further improve the accuracy of short-term household load forecasting, this paper proposes a VMD-GRU-SAM-GRU-AM short-term household load forecasting method. This method uses VMD to decompose the original power data sequence into a series of subsequences with limited bandwidth and reconstructs the sub-sequences according to the sample entropy (SE). The reconstructed sub-sequences and the data set after data preprocessing are reorganized to obtain a new dataset, and finally, the GRU-SAM-GRU-AM forecasting model is used for prediction. Through the actual dataset experiment, the forecasting method proposed in this paper has a good prediction effect in the short-term household load forecasting.

2. **VMD-GRU-SAM-GRU-AM forecasting method.** The forecasting method of VMD-GRU- SAM-GRU-AM in this paper is shown in Figure 1. The process of this

method: First, the dataset used in this article is processed for missing values, and the time interval is converted into 1h. Second, the household load in the dataset is selected for VMD decomposition, and the decomposed sub-signals are reconstructed by SE. Third, the reconstructed sub-signals are added to the pre-processed data set to form a new dataset. Fourth, for the training of GRU-SAM-GRU-AM, the SAM weights the information extracted by the GRU as the processing information of the next layer of GRU, AM weights the output of GRU, and the degree of influence of different information on the prediction result is related to the weight. Finally, output the prediction result.

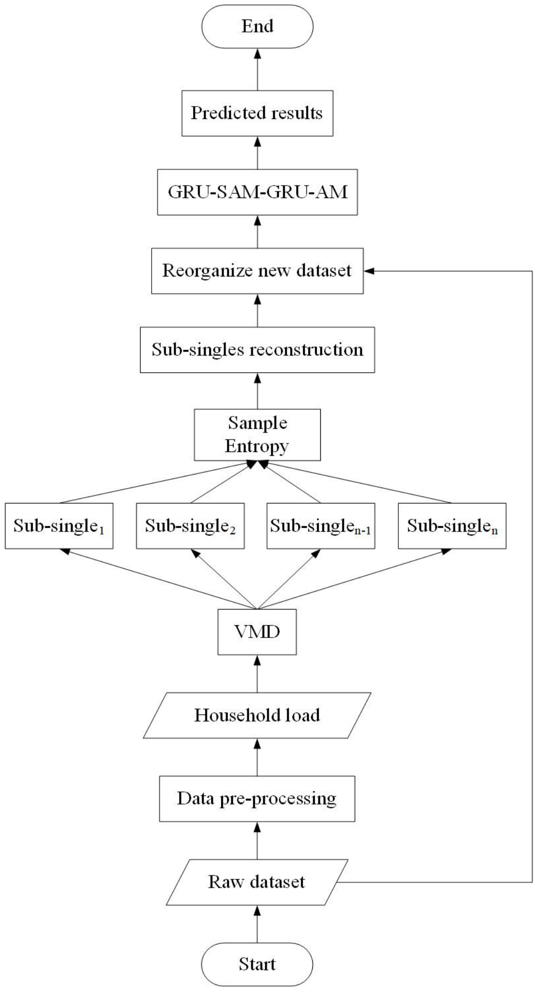

FIGURE 1. Overall architecture

2.1. **VMD.** VMD is a new non-recursive, adaptive decomposition estimation method proposed by Dragomiretskiy and Zosso [23] in 2014. VMD decomposes the original timing signal and gets k sub-signals, whose modal functions $u_k(\mathrm{t})$ have different center frequencies. Wiener filtering and alternating direction multiplier method are used to update the modes to minimize the sum of the estimated bandwidth of each mode.

Construction of variational mode decomposition problem:

Hilbert transformation was performed for each modal component $u_k(\mathrm{t})$ to obtain the corresponding unilateral spectrum.

$$\left[\delta\left(t\right)+\frac{j}{\pi t}\right]*u_k\left(t\right) \tag{1}$$

To adjust the frequency spectrum corresponding to each mode to the corresponding base frequency band, for each mode function $u_k(\mathrm{t})$, it is aliased by the exponent term $\mathrm{e}^{-j\omega kt}$ of its corresponding center frequency ω_k to obtain Eq.(2).

$$\left[(\delta(t) + \frac{j}{\pi t}) * u_k(t)\right] * e^{-j\omega_k t} \tag{2}$$

Construct a constrained variational model and estimate the bandwidth by Gaussian smoothing, that is, calculate the square L^2 norm of the analytical signal gradient to generate a constrained variational problem:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \{\sum\limits_{\mathrm{k}} \left\|\partial_{\mathrm{t}} \left\| \left[\left[(\delta(t) + \frac{j}{\pi t}) * u_k(t)\right] * e^{-j\omega_k t}\right] \right\|^2\right\} \\ s.t \sum\limits_{k} u_k = f \end{cases} \tag{3}$$

The solution of Variational Modal Model:

Transform the constrained problem into a non-constrained problem: introduce the quadratic penalty factors α and Lagrange multiplication operator $\lambda(\mathrm{t})$. where α guarantees the accuracy of the signal reconstruction, $\lambda(\mathrm{t})$ maintains the strictness of the constraint conditions, and the augmented Lagrangian expression is as follows:

$$\begin{aligned} [L(\{u_k\}, \{\omega_k\}, \{\lambda\}) = \alpha \sum_k \left\| \partial_{\mathrm{t}} \left\| \left[\left[(\delta(t) + \frac{j}{\pi t}) * u_k(t)\right] * e^{-j\omega_k t}\right] \right\| \right\|^2 \\ + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \end{aligned} \tag{4}$$

Using the alternating direction multiplier method (ADMM), alternately updating u_k^{n+1}, $\omega_k^{n+1}$and λ^{n+1}, the following formula can be obtained:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum\limits_i \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \tag{5}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_i(\omega)|^2 \mathrm{d}\omega}{\int_0^\infty |\hat{u}_i(\omega)|^2 \mathrm{d}\omega} \tag{6}$$

$$\lambda^{n+1} \leftarrow \lambda^n + \tau \left(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega)\right) \tag{7}$$

2.2. **GRU.** The GRU contains two gates: reset and update gate [24]. Formula (8) and (9) represent the working principle of the reset gate. The hidden layer state h_{t-1} at time *t-1* and the input x_t at time t produce the gating signal r_t under the action of the activation function sigmoid, r_t determines the degree of information retention in h_{t-1}, and the result is spliced with x_t, and then obtain the new state $\mathrm{h}_t^{'}$ containing x_t information through the activation function tanh. The formulas for the update gate are (10) and (11), similarly, the gating signal z_t is generated after h_{t-1} and x_t, the range of z_t is 0-1, and the size of z_t indicates how much information is memorized. In formula (11)(1-z_t) $\times$ h_{t-1} represents the forgetting of unimportant information in the hidden layer state at time *t-1*, and z_t $\times$ $\mathrm{h}_t^{'}$ represents the selective memory of the information at time t. Therefore, h_t represents the hidden layer state at the current time t.

$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right) \tag{8}$$

$$h_t^{'} = \tanh\left(W \cdot [r_t \times h_{t-1}, x_t]\right) \tag{9}$$

$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right) \tag{10}$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times h_t^{'} \tag{11}$$

The structure diagram of GRU is shown in Figure 2. Where $\times$ represents Hadamard Product, σ represents the activation function sigmoid, and 1- represents that the data propagated forward on the link is 1-z_t.

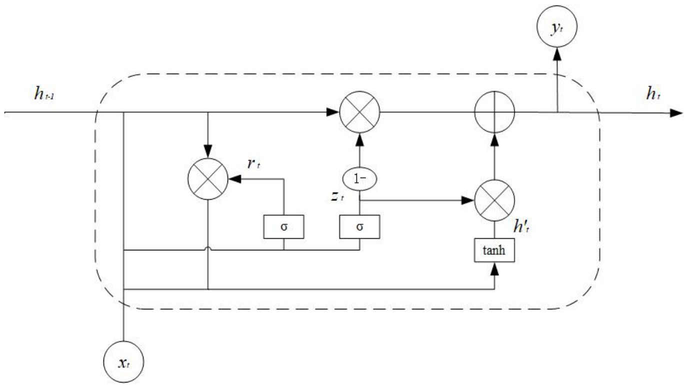

FIGURE 2. GRU architecture

2.3. **SAM and AM.** The attention mechanism is a brain signal processing mechanism that simulates the human visual reception signal. When people observe or process something, they will focus on the most distinctive features of the thing, and ignore other features that are not obvious to a certain extent. Use this similar feature in the attention mechanism algorithm to assign different weights to each different feature in the input sequence data, and optimize these weights through training to achieve the purpose of highlighting key features, thereby improving the processing efficiency of the prediction model.

The new attention mechanism proposed by the Google team in a paper entitled "Attention Is All You Need" in 2017 is called self-attention mechanism (SAM) [25]. In this paper, SAM is used to connect two GRU layers. Because SAM itself is insensitive to timing data, a GRU layer is designed to encode input data to obtain context information. The introduction of SAM highlights the important information in the context information, and then enters the next layer of GRU for processing. About the calculation process of the SAM layer:

Assuming that the feature information set obtained by the first-level GRU can be expressed as $H^1 = \{h_1^1, h_2^1, ..., h_{n-1}^1, h_n^1\}$ as the input of the self-attention layer, where n represents the time step, the calculation process is as follows:

$$M = \tanh\left(W_1 H^1\right) \tag{12}$$

$$A = soft\max(W_2 M) \tag{13}$$

$$S = A^T H \tag{14}$$

where W_1 and W_2 are training matrices, M is the similarity value using the perceptron similarity calculation formula, A is the weight matrix, $S = \{s_1, s_2, ..., s_{n-1}, s_n\}$ as the input of the second-level GRU.

The information is $H^2 = \{h_1^2, h_2^2, ..., h_{n-1}^2, h_n^2\}$ after being processed by the second-level GRU. H^2 is used as the input of the AM layer. AM integrates the output information of the second-level GRU according to the corresponding weights. The calculation process is as follows:

$$e_i = W_3 H^2(dot) \tag{15}$$

$$a_i = \frac{\exp(e_i)}{\sum_{k=1}^{t} \exp(e_k)} \tag{16}$$

$$r_t = \sum_{i=1}^{t} a_i \times h_i^2 \tag{17}$$

where e^i represents the feature information contained in the i-th GRU hidden layer state h_2^i, W_3 represents the weight matrix, a^i represents the weight corresponding to the i-th feature information, and rt represents the information output by the attention at time t.

Figure 3 shows the architecture of GRU-SAM-GRU-AM.

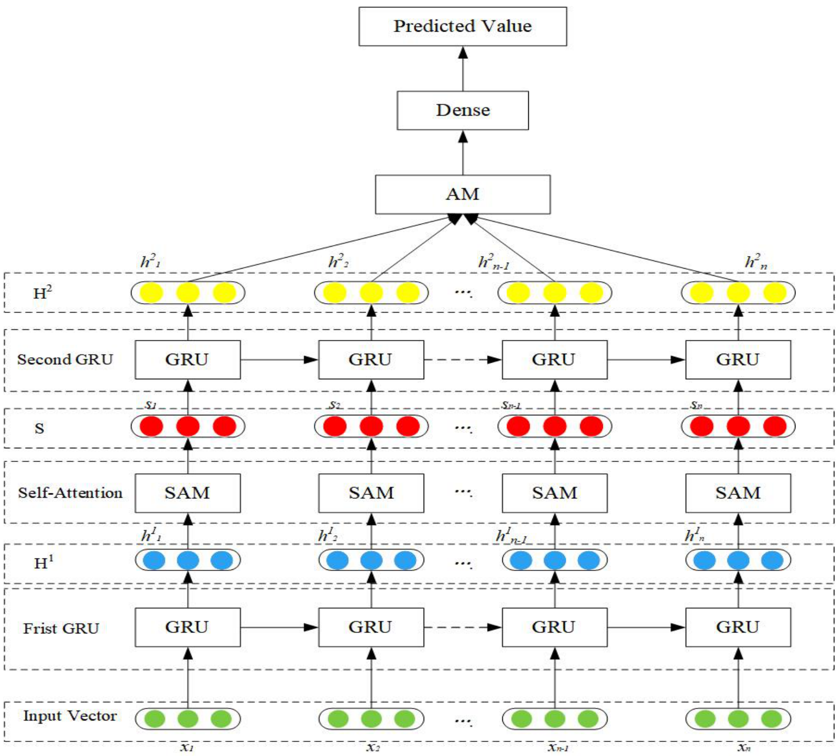

FIGURE 3. GRU-SAM-GRU-AM architecture

3. **Simulation.** In this paper, The dataset comes from the individual household electric power consumption (IHEPC) dataset in the UCI database [26]. The dataset contains 2 075 259 data, the sampling frequency is 1 minute, and the sampling time of the dataset is 47 months. Mean Absolute Percentage Error(MAPE) and Root Mean Square Error(RMSE) are used as evaluation functions. The formula is shown below.

$$MAPE = \frac{1}{m}\sum_{i=1}^{n}\left|\frac{y_i - {y_i}^*}{y_i}\right| \times 100\% \tag{18}$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{n}(y_i - y_i^*)^2} \tag{19}$$

In the formula, m represents the number of samples, y_i represents the actual value of the i-th sample, and y_i^* represents the predicted value of the i-th sample.

3.1. **Model parameters.** Parameter setting mainly includes the prediction model and comparison prediction method. The comparison methods in this article include ANN and SVR. The parameter setting table is shown in Table 1.

Table 1. Model parameter

Model	Parameter	
GRU	GRU layer	2
	GRU node	64
	Batch size	128
	Dropout	0.2
	Dense activation function	Relu
	Dense node	1
	Epoch	150
	Optimizer	Adam
	Validation split	0.2
ANN	Dense node	64
	Dense activation function	Sigmoid
	Regularizer L^2	0.001
	Epoch	200
	Batch size	128
	Validation split	0.2

The VMD parameters are set as penalty parameter α=1000; the number of decompositions of the IHEPC data set is selected according to experience to k=4; the initial center frequency ω=0; the convergence criterion γ=10^{-7}. The sample entropy of the IHEPC data set is shown in Table 2 below. It can be seen that the sample entropy of the IHEPC data set is similar in Intrinsic Mode Function 3(IMF3) and IMF4. According to the principle of superposition of sample entropy in [27], IMF3 and IMF4 can be added to reconstruct a new IMF, which decomposes three variables as input variables, and combines the IHEPC dataset to form a new data set for prediction.

3.2. **Experiment analysis.** Through experiments, we get experimental results that prove that the experimental method of VMD-GRU-SAM-GRU-AM can achieve better experimental results. First of all, through the comparison of RMSE and MAPE in Figure

Table 2. Sample entropy of the IHEPC dataset

Dataset	IMF1	IMF2	IMF3	IMF4
IHEPC	0.206	1.464	1.938	2.047

4, it can be seen that there is a large gap between GRU and VMD-GRU, GRU-SAM-GRU-AM, and VMD-GRU-SAM-GRU-AM. the RMSE of VMD-GRU 0.309 lower than GRU and 19.133 lower than MAPE. It can be obtained that the use of VMD can greatly improve the prediction effect. The reason is that the sub-signals decomposed by VMD will greatly help the training of the forecasting model. Secondly, the GRU-SAM-GRU-AM forecasting model in Figure 4 has better prediction results than the traditional GRU. The MAPE of VMD-GRU-SAM-GRU-AM 0.853 lower than VMD-GRU. Because the SAM and AM add the concept of weight to the information processing process, the processing capability of the data information is better than the traditional GRU.

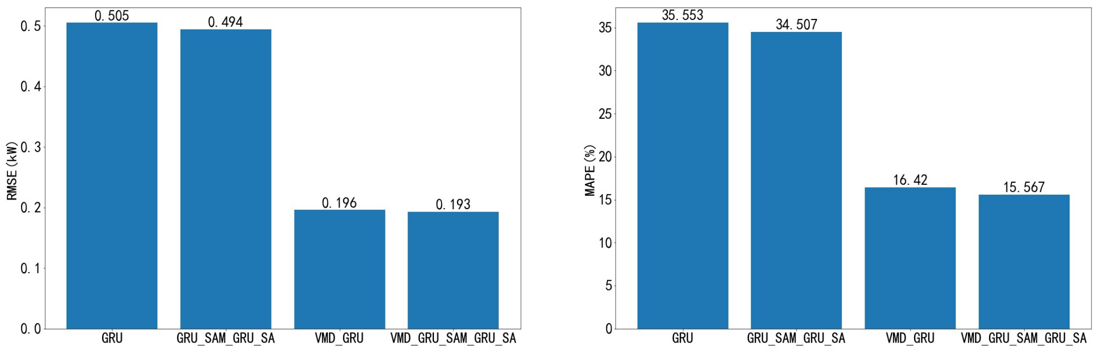

Figure 4. RMSE and MAPE experimental results

In order to further verify the superiority of the method, this paper also compared with ANN and SVR. The comparative experimental results are shown in Table 3 below. ANN has the least ideal effect in the case of big data, while SVR also has significantly lower performance than GRU in the face of data with tens of thousands of levels and severe fluctuations. In the case of big data, ANN has the most unsatisfactory effect. SVR is also dealing with tens of thousands of levels and violently fluctuating data, and its performance is significantly lower than GRU.

Table 3. Experimental results

Method	RMSE(kW)	MAPE(%)
VMD-ANN	0.257	27.498
VMD-SVR	0.212	21.556
VMD-GRU	0.196	16.420
VMD-GRU-SAM-GRU-AM	0.193	15.567

The part of the comparative experimental prediction curve is shown in Figure 5. In the trough of 1470-1480, the predicted line of VMD-GRU-SAM-GRU-AM is closer to the actual value line. The fitting degree of the VMD-ANN curve and actual value curve is poor compared with other curves.

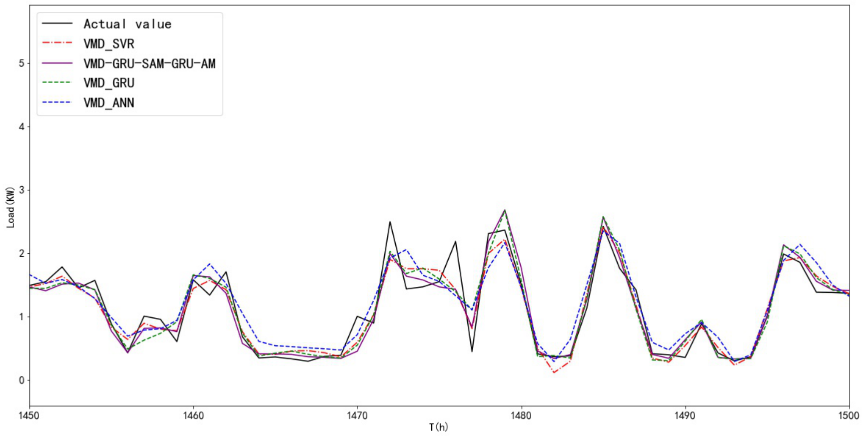

FIGURE 5. Comparison of prediction curves

4. **Conclusions.** This paper proposes the VMD-GRU-SAM-GRU-AM method for household load, and the validity of the method is verified by experiment. The main conclusions are as follows: 1) Data processing by VMD decomposition algorithm can greatly improve the prediction accuracy, and 2) GRU-SAM-GRU-AM is more accurate than traditional GRU prediction accuracy. The proposed approach may be further improved by adopting the meta-heuristic optimization algorithms [28–32].

Acknowledgment. This work is supported by the Collaborative Education Program of Industry and Education of Ministry of Education (No. 201901052008), Scientific Research Project of Fujian Education Department (No. JAT190069, No. JK2017029) and Scientific Research Project of Fujian University of Technology (No. XF-X19017)

REFERENCES

[1] L.S. Liu, D.W. He, Y. Ma, Y. Ma, J. Huang, J.N. Li, J.X. Yao, A Novel License Plate Location Method Based on Deep Learning, *Journal of Network Intelligence*, vol. 5, no. 3, pp. 93-101, 2020.

[2] K.W. Huang, C.C. Lin, Y.M. Lee, Z.X. Wu, A Deep Learning and Image Recognition System for Image Recognition, *Data Science and Pattern Recognition*, vol. 3, no. 2, pp. 1-11, 2019.

[3] Y.F. Li, J.B. Li, J.S. Pan, Hyperspectral Image Recognition Using SVM Combined Deep Learning, *Journal of Internet Technology*, vol. 20, no. 3, pp. 851-859, 2019.

[4] A.J. Conejo, M.A. Plazas, R. Espinola, A.B. Molina, Day-Ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models, *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1035-1042, 2005.

[5] N.T. Huang, G.B. Lu, D.G. Xu, A Permutation Importance-Based Feature Selection Method for Short-Term Electricity Load Forecasting Using Random Forest, *Energies*, vol. 9, no. 10, 2016.

[6] J. Deshmukh, U. Bhosle, A Study of Mammogram Classification using AdaBoost with Decision Tree, KNN, SVM and Hybrid SVM-KNN as Component Classifiers, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 9, no. 3, pp. 548-557, 2018.

[7] Y.L. Qiao, J.S. Pan, S.H. Sun, Improved K Nearest Neighbors Classification Algorithm, *ACTA Electronica Sinica*, vol. 33, no. 6, pp. 1146-1149, 2005.

[8] J.S. Pan, Y.L. Qiao, S.H. Sun, A Fast K Nearest Neighbors Classification Algorithm, *IEICE Transactions on Fundamentals of Electronics, Communication and Computer Sciences*, vol. E87A, no. 3, pp. 961-963, 2004.

[9] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, Y. Zhang, Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network, *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841-851, 2019.

[10] H.C. Chao, F. Lin, J.S. Pan, W.C. Chien, C.F. Lai, Power Load Forecasting Based on VMD and Attention-LSTM, *DSIT 2020: Proceedings of the 3rd International Conference on Data Science and Information Technology*, pp. 19-25, 2020.
[11] C. Li, Z.Y. Chen, J.B. Liu, D.P. Li, X.Y. Gao, F.C. Di, L.X. Li, X.H. Ji, Power Load Forecasting Based on the Combined Model of LSTM and XGBoost. *PRAI '19: Proceedings of the 2019 the International Conference on Pattern Recognition and Artificial Intelligence*, pp. 46-51, 2009.
[12] A.A. El desouky, M.M. Elkateb, Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA, *IEE Proceedings - Generation, Transmission and Distribution*, vol. 147, no. 4, pp. 213-217, 2000.
[13] J.L. Zhang, Y.M. Wei, D.Z. Li, Z.F. Tan, J.H. Zhou, Short term electricity load forecasting using a hybrid model, *Energy*, pp. 774-781, 2018.
[14] N. Liu, Q.F. Tang, J.H. Zhang, W. Fan, J. Liu, A hybrid forecasting model with parameter optimization for short-term load forecasting of micro-grids, *Applied Energy*, pp. 336-345, 2014.
[15] F.F. He, J.Z. Zhou, Z.K. Feng, G.B. Liu, Y.Q. Yang, A hybrid short-term load forecasting model based on variational mode decomposition and long short-term memory networks considering relevant factors with Bayesian optimization algorithm, *Applied Energy*, pp. 103-116, 2019.
[16] M. Pelikan, D.E. Goldberg, E. Cant-paz, BOA: The Bayesian optimization algorithm, *GECCO'99: Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation*, vol. 1, pp. 525-532, 1999.
[17] H. Wang, T.W. Dai, K. Liu, X.X. Ru, Y.L. Lou, Node Confidence Calculation Method Based on D-separation of Bayesian Network, *Journal of Network Intelligence*, vol. 5, no. 4, pp. 166-178, 2020.
[18] F.Q. Zhang, G.Y Ding, Z.J. Mao, L. Xu, X.Y. Zheng, Bayesian Network for Motivation Classification in Creative Computation, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 8, no. 4, pp. 888-902, 2017.
[19] C.H. Hsieh, C.M. Kuo, Y.S. Hsieh, Bayesian-Based Probabilistic Architecture for Image Categorization Using Macro- and Micro-Sense Visual Vocabulary, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 9, no. 6, pp. 1628-1638, 2018.
[20] M. Yan, J.Z. Wang, J. Li, K. Zhang, Z.M Yang, Traffic scene semantic segmentation using self-attention mechanism and bi-directional GRU to correlate context, *Neurocomputing*, vol. 386, pp. 293-304, 2019.
[21] X.S. Zhao, R. Zhang, J.L. Wu, P.C. Chang, A Deep Recurrent Neural Network for Air Quality Classification, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 9, no. 2, pp. 346-354, 2018.
[22] E.Arianyan, H.Taheri, S. Sharifian, Novel energy and SLA efficient resource management heuristics for consolidation of virtual machines in cloud data centers, *Computers Electrical Engineering*, vol. 47, pp. 222-240, 2015.
[23] K. Dragomiretskiy, D. Zosso, Variational Mode Decomposition, *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531-544, 2014.
[24] R. Fu, Z. Zhang, L. Li, Using LSTM and GRU neural network methods for traffic flow prediction, *2016 31st Youth Academic Annual Conference of Chinese Association of Automation*, pp. 324-328, 2016.
[25] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, Attention is all you need, *Advances in neural information processing systems*, pp. 5998-6008, 2017.
[26] UCI Machine Learning Repository, *[Online]Available:http://archive.ics.uci.edu/ml/index.php*, 2017.
[27] Q.L. Wu, H.X Lin, Daily urban air quality index forecasting based on variational mode decomposition, sample entropy and LSTM neural network, *Sustainable Cities and Society*, vol. 50, 101657, 2019.
[28] Z.Y. Meng, J.S. Pan, L.P. Kong, Parameters with Adaptive Learning Mechanism (PALM) for the enhancement of Differential Evolution, *Knowledge-Based Syst*, vol. 141, pp. 92-112, 2018.
[29] Y.P Liao, X.K. Lv, SINR Improvement in Space-time Adaptive Processing Systems Based on Meta-heuristic Algorithms, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 9, no. 6, pp. 1479-1488, 2018.
[30] S.C. Chu, H.C. Huang, J.F. Roddick, J.S. Pan, Overview of Algorithms for Swarm Intelligence, *In: Jedrzejowicz P., Nguyen N.T., Hoang K. (eds), Berlin/Heidelberg, Springer, LNCS, 6922*, pp. 28-41, 2011.
[31] J.S. Pan, X.P. Wang, S.C. Chu, T.T. Nguyen, A Multi-group Grasshopper Optimisation Algorithm for Application in Capacitated Vehicle Routing Problem, *Data Science and Pattern Recognition*, vol. 4, no. 1, pp. 41-56, 2020.

[32] B.Q. Jiang, J.S. Pan, A Parallel Quasi-Affine Transformation Evolution Algorithm for Global Optimization, *Journal of Network Intelligence*, vol. 4, no. 2, pp. 30-46, 2019.