Proc. of the International Conference on Electrical, Computer and Energy Technologies (ICECET 2022)
20-22 June 2022, Prague-Czech Republic

Comparison of non-intrusive load monitoring supervised methods using harmonics as feature

Omar Al-Khadher
Faculty of Engineering, Universiti Pertahanan Nasional Malaysia
Kuala Lumpur, Malaysia
o.khadher@hotmail.com

Azharudin Mukhtaruddin
Faculty of Engineering, Universiti Pertahanan Nasional Malaysia
Kuala Lumpur, Malaysia
azharudin@upnm.edu.my

Fakroul Ridzuan Hashim
Faculty of Engineering, Universiti Pertahanan Nasional Malaysia
Kuala Lumpur, Malaysia
fakroul@upnm.edu.my

Muhammad Mokhzaini Azizan
Faculty of Engineering and Built Environment, Universiti Sains Islam Malaysia
Negeri Sembilan, Malaysia
mokhzainiazizan@usim.edu.my

Husin Mamat
School of Aerospace Engineering, Universiti Sains Malaysia
Penang, Malaysia
hussinm@usm.my

Mohamed Mani
School of Engineering, Universiti Tun Hussein Onn Malaysia
Johor, Malaysia
fallatahmani@gamil.com

***Abstract*— Non-intrusive Load Monitoring (NILM), also known as energy disaggregation, is a useful technique for analyzing energy consumption data, monitored from a single-point source such as a smart meter. In this paper, a three-phase induction motor was designed in SIMULINK to be used for NILM system based on current waveforms and odd-numbered harmonics up to the ninth harmonic. Supervised learning classifiers were proposed including decision tree, KNN, NN, Ensemble, and SVM algorithms to classify the loads with high accuracy. In comparison, results show that the decision tree classifier can classify the loads efficiently for the most loads. Although the ensemble showed a high accuracy but still needs more time for training due to the complexity of the model. Additionally, the more samples obtained the more accuracy of classification, but a high sampling rate has more cost and analysis it takes more time for training.**

Keywords—Non-intrusive load monitoring (NILM), induction motor (IM), harmonics, feature extraction, load classification, SIMLUINK

I. INTRODUCTION

It is estimated that the building sector accounts for over 40% of total non-renewable energy consumption, 40% of greenhouse gas emissions, and 70% of total power consumption in developed nations [1]. In tandem with the fast expansion of the economy, the demand for energy among the public continues to rise. Electricity conservation therefore contributes directly and significantly to sustainable development in both economy and environment, with significant market demand for technologies and equipment that facilitate the realization of efficient and energy-saving electricity consumption [2].

It is possible to conduct load monitoring and identification in an intrusive or non-intrusive manner. The methods of intrusive load monitoring depend on the installation of many specialized sensors on each device to gather data on its electricity use. It is difficult to construct a sensor-distributed measurement network in reality, despite the high precision of identification provided by direct measuring [3]. In addition to being known as energy disaggregation, non-intrusive load monitoring (NILM) is a valuable approach for assessing energy consumption data that is collected from a single source, such as a smart meter. This is because the technology may be simply implemented into existing structures. When used in conjunction with signal processing and machine learning methods, the NILM is capable of extracting individual load profiles from aggregate signals [4].

Moradzadeh et al. [5] introduced a method for monitoring energy consumption based on power features and SVM. SVM classifier shows excellent correlation and good overlap between the target data and the output of the designed SVM. Yaqian et al. [6] presented an approach for SVM based on harmonic characteristics. The comparative experiment verifies that the harmonic features can play a supplementary role when the active power and reactive power features are confused, which improves the classification accuracy.

Literature [7] improved the identification accuracy and the detection of appliances based on k-nearest neighbor (k-NN) classification algorithm and the V-I trajectory. Guohua et al. presented a RF algorithm and using a harmonic analysis based on Fourier Transform to achieve the feature extraction [8]. Yassine et al. [9] introduced an effective NILM system based

978-1-6654-7087-2/22/$31.00 ©2022 IEEE

on a novel multi-descriptor fusion with dimensionality reduction and Decision Bagging Tree classifier.

This paper [10] proposes a network structure with residual unit added to the traditional convolutional neural network, to reduce the probability of gradient explosion. Hasbi et al. [11] presented a method for monitoring energy usage based on CNN from V-I trajectory image features. . CNN classifier shows an efficient performance for appliance classification. Laura et al. [12] proposes a LSTM-based neural network applied to the load identification issue in NILM, which can efficiently recognize appliances, based on power consumption measurements coming from a single point at low sampling rates.

Researchers concentrate on the residential buildings more than commercial buildings due to the large number types of devices as well the ability to classify loads. commercial buildings consume high power for a limiting time per day as well lack of interesting in commercial building energy management studies. In this paper, we propose induction motors classification based on current waveforms and harmonics. Using several classifiers to compare and obtain the best for the loads of commercial buildings.

II. TYPES OF LOADS

Because of its simple and sturdy design, a three-phase induction motor (IM) is extensively utilized in almost any industry or in commercial structures. A three-phase IM is a machine that operates on a single excitation. The electromagnetic induction law of Faraday is the basis of the induction motor's operation [13]. In this paper, we design DC voltage source inverter (VSI) supply to three models of induction motor with different specifications. Moreover, using VSI to generate the odd-numbered harmonics in the motors and use them as features in NILM system. The following parameters are 400 V, 50 Hz, IM1 (4KW), IM2 (7.5KW), IM3(37KW). Fig. 1. illustrates one of the models of three-phase IM.

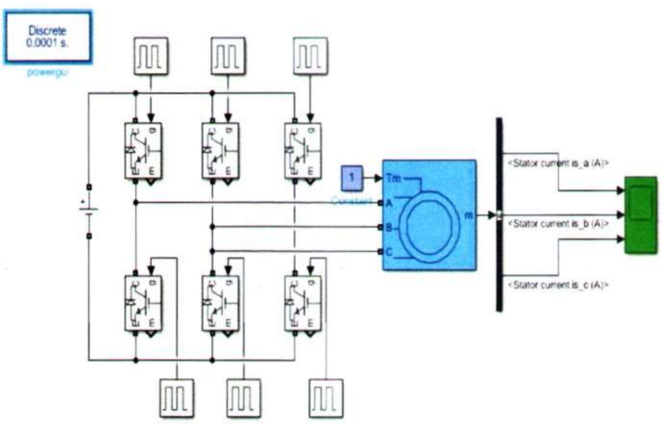


Fig. 2. Three-phase induction motor model.

The high-power consumption in the commercial buildings comes from the large motors due to the operating on the weekdays for a limiting time which using these loads are essential. In the next section, the NILM system will be discussed in detail such as sampling frequency, features extraction, and how the loads can be classified.

III. CASE STUDY

Fig. 2. demonstrates the proposed NILM system in this study. The data was collected by using SIMULINK to design the induction motor. obtaining the data was for 1 min at 1 kHz sampling frequency. FFT analysis and scope block in SIMULINK can be used to obtain the features of the loads including current waveforms and harmonics. In this paper, the collected data comes from three loads with three phases (Motor1_P1, Motor1_P2, Motor1_P3, Motor2_P1, Motor2_P2, Motor2_P3, Motor3_P1, Motor3_P2, Motor3_P3) to be classified each load based on the selected features.

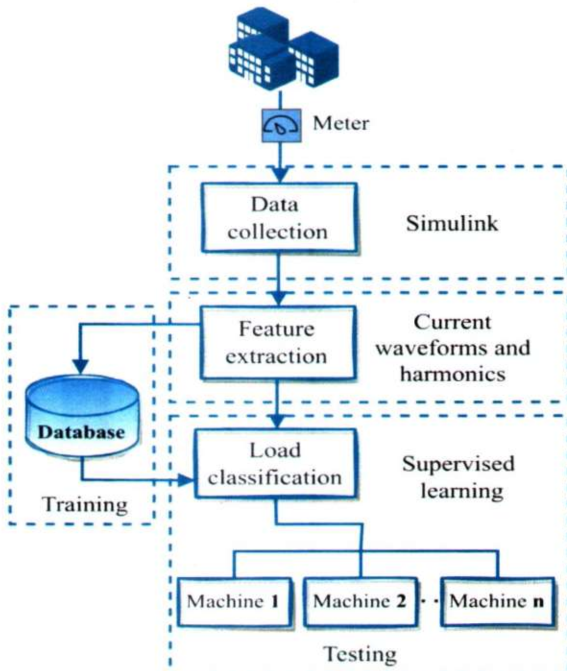


Fig. 1. The proposed NILM system to the IMs.

IV. FEATURES EXTRACTION

In the commercial buildings, all the loads need to be running at the same time which makes it a challenge for the NILM system to classify the loads based on three phases. Load events are characterized primarily by their active and reactive power. Almost all appliances have a minor overlap in their power ranges, which is distinguishable solely by the difference between active and reactive power in their operation. The fact that the power ranges of certain appliances have a significant overlap makes it easy to get confused only on the basis of the difference between active power and reactive power [6]. The interesting thing is that each load has unique current waveforms and harmonics with different specifications which leads to various current levels between the loads.

In this paper, the fundamental wave amplitude of current and harmonics have been obtained by FFT analysis in SIMULINK

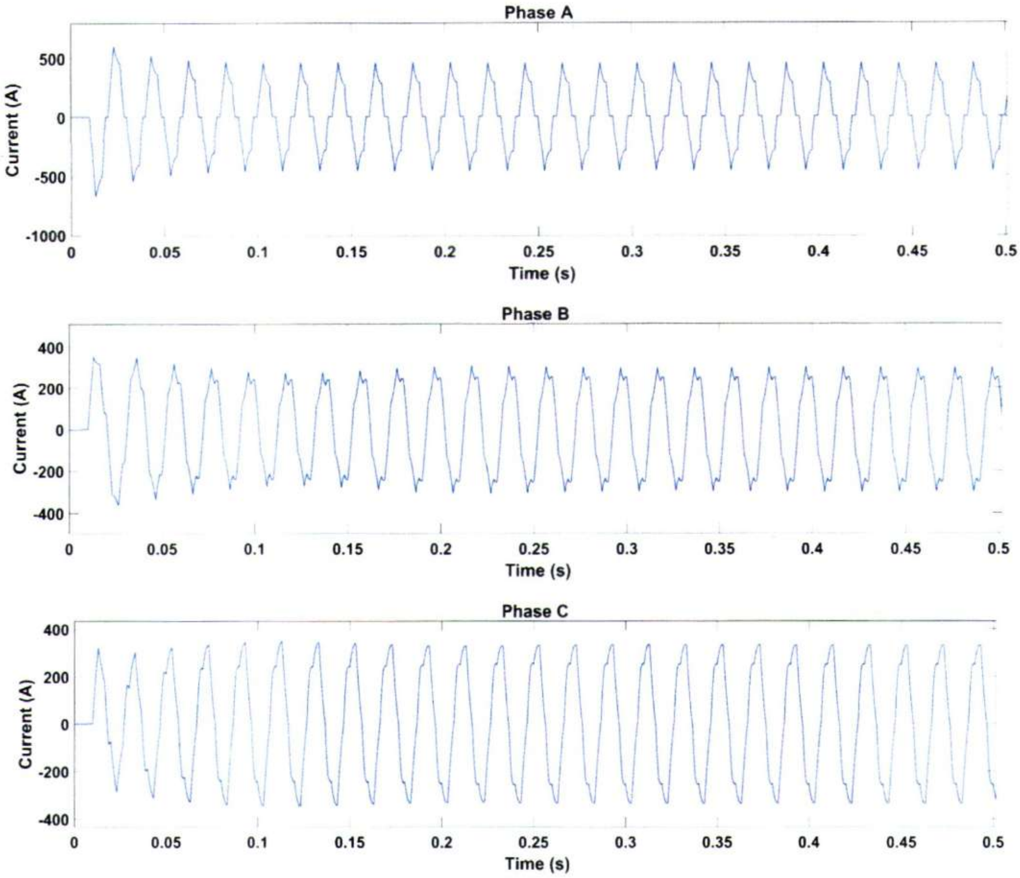


Fig. 3. The total current waveform of induction motors for each phase.

as well extracting the features of the collected data based on odd-numbered harmonics up to 9th harmonic and event harmonics will be neglected as shown in Fig. 4. Choosing 1 cycle of the current waveform to analyze the harmonics order of loads.

Fig. 3 shows the signatures of the motors and how the current can be different with similar loads but with diverse specifications. After extracting the features, the next stage of NILM is the classification which each motor will be labeled after collecting the total current consumption.

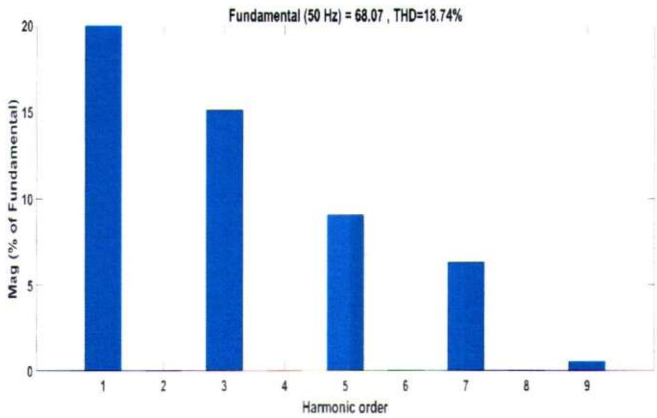


Fig. 4. Current harmonics Phase A for induction motor (10 HP, 7.5kW).

V. LOAD CLASSIFICATION

Due to diverse types of loads presenting constant patterns, numerous states and non-linear electrical behavior, accuracy and flexibility of load identification methods are required. Therefore, several techniques have been developed to improve the efficacy of the designed methods, which can be classified in function of the used machine learning algorithm. Generally, machine learning techniques can be grouped as supervised and unsupervised methods, as well as event or non-event-based methodologies [14].

Supervised techniques prove that can achieve the high accuracy of classification due to the labeled data in the database and the classifier need to be trained to obtain the true class. In this paper, we present supervised methods to classify the loads based on SVM, KNN, NN, Ensemble, Decision Tree. The collected data in the classification should be split into training data and test data to demonstrate the unseen data to the classifier and how should identify them with high accuracy. In the proposed system, 60% of the data were considered for the training dataset and 40% for the testing dataset.

The kernel function of SVM is quadratic and an automatic kernel scale. Moreover, based on KNN the number of neighbors is 10 with a euclidean distance metric and equal distance weight. For NN model, the number of fully connected layers is 1 as well the first layer size is 25 with a 1000 iteration limit. The ensemble method is AdaBoost, and the learner type is a decision tree with 20 splits and 0.1 of the learning rate. However, the decision tree has 100 splits and uses Gini's diversity index of split criterion.

VI. Results and Discussion

A. *Evaluation metries of load classification*

The accuracy assessment metrics are widely used in the evaluation of the process of classification. It is mostly useful for the analysis of false assumptions, which is generally applicable to unused datasets. According to Table 1, the parameters for the confusion matrix are derived from True Positive (TP) and True Negative (TN) values for the purpose of computing the performance assessment. When the appliances are turned on and off, the TP and False Negative (FN) are taken into consideration [15].

TABLE I. CONFUSION MATRIX

	Predicted Positive	Predicted Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

Sensitivity and recall parameters are terms used to describe the performance result derived from actual positive rate measurements, whereas precision refers to the positive projected value. In order to determine the accuracy and recall value, use formulas (1) and (2) [15].

$$recall = \frac{TP}{TP+FN} \tag{1}$$

$$precision = \frac{TP}{TP+FP} \tag{2}$$

The assessment criteria are differentiated by a performance evaluation indicator that takes into account classification accuracy, as shown by the accuracy pointers and the F-measure that are computed as given below [15]:

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

$$F_{measure} = 2 * \frac{precision*recall}{precision+recall} \tag{4}$$

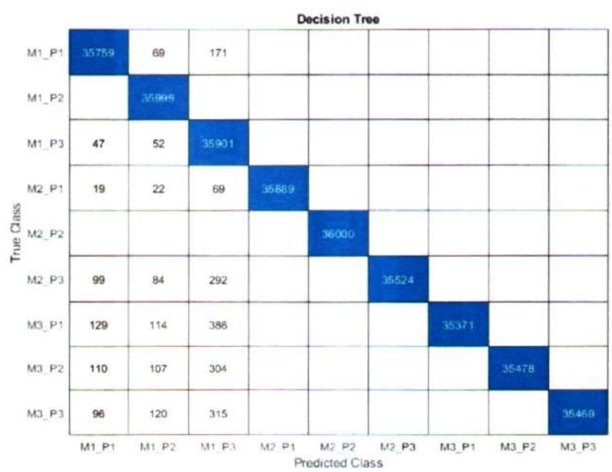
B. *Result analysis*

In this paper, the characteristics were extracted based on steady-state features and using FFT analysis to obtain the odd-numbered harmonics up to the ninth harmonic and select five classifiers to identify the loads by KNN, SVM, NN, ensemble, decision tree. Fig 5 shows the confusion matrix to calculates precision, recall, accuracy and $F_{measure}$ which to be summarized in Table 2. As shown in Fig 5 (a) the FN accounts at 240 and FP shows 2365 that the classifier has missed predicting the classes correctly. However, Fig 5 (b) and (c) illustrate a huge number of misprediction points for FN are 974 and 2393, and FP are 3825 and 2593. Fig 5 (d) demonstrates the lowest misprediction points with 1164 for FN and 1052 for FP, while Fig 5 (e) shows only FP with 4785. Ensemble proves that be able to classify with less misprediction.

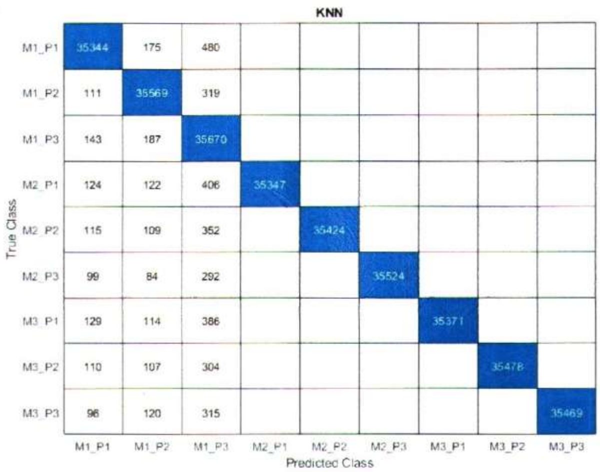
TABLE II. RESULTES OF CLASSIFICATION BASED ON SUPERVISED LEARNING.

Classifier	recall	precision	accuracy	F-measure
Decision Tree	99.2%	99.21%	99.2%	99.2%
KNN	98.53%	98.58%	98.5%	98.5%
NN	98.48%	98.5%	98.4%	98.5%
Ensemble	99.32%	99.35%	99.3%	99.3%
SVM	98.53%	98.7%	98.5%	98.6%

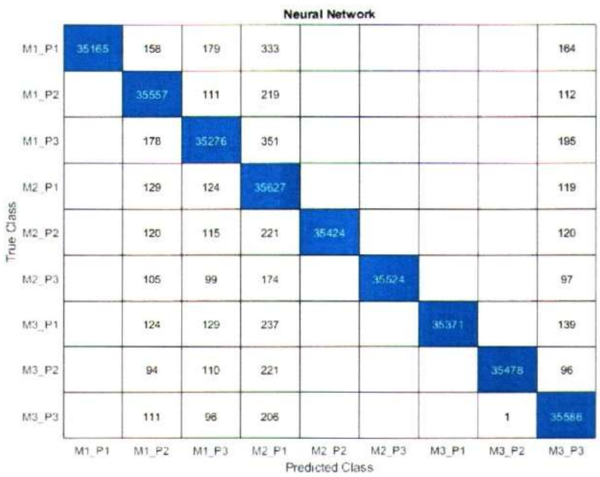
The observation illustrates as shown in Table 2, the techniques of load classification have an excellent accuracy for the three motors having three phases with various specifications. The results show that the accuracy rate of classifiers can reach 98.4% to 99.3% as well the accuracy rate is 88.3% and above for each load. Ensemble (Boosted trees) can efficiently classify the loads with an accuracy rate of 99.3%. In comparison, the decision tree has the lowest time training following that KNN classifier, as well the longest time training goes to SVM. Although the ensemble showed a high accuracy but still needs more time for training due to the complexity of the model. Additionally, the more samples obtained the more accuracy of classification, but a high sampling rate has more cost and analysis it takes more time for training. Moreover, selecting harmonics as features need to set a high sampling rate (kHz) to obtain high order harmonics.



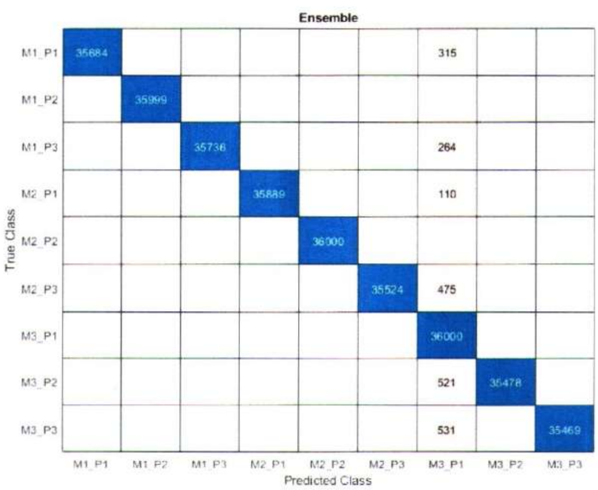
(a) Testing result by decision tree classifier.



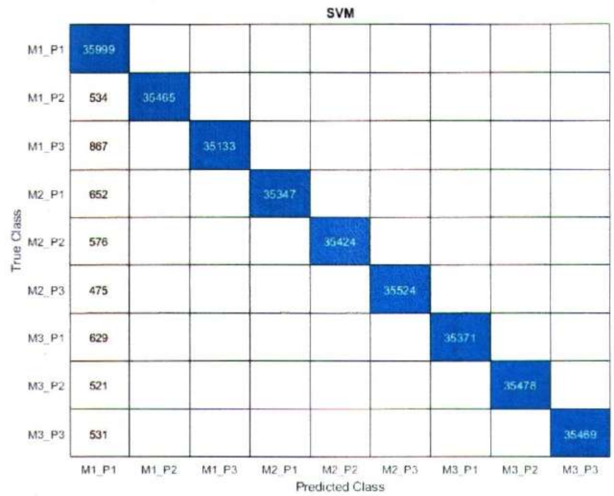
(b) Testing result by KNN classifier.



(c) Testing result by NN classifier.



(d) Testing result by Ensemble classifier.



(e) Testing result by SVM classifier.

Fig. 5. Classification of loads based on superivsed learning.

VII. CONCLUSION

This paper has designed three-phase IM to be used in the NILM system for load classification based on supervised learning. three motors with different specifications that be used mostly in commercial buildings and consume huge power more than the small loads. Using current waveforms and harmonics as features for the proposed NILM system as well as KNN, NN, SVM, decision tree, ensemble algorithms for load classification. The algorithms of NILM system proved that can easily identify the loads with an accuracy rate reach 99.3% based on boosted tree classifier. IMs (Type-I) can easily be classified due to the single-state pattern, the various power of the loads, and the lack of appliances number.

A high sampling rate (kHz) is required to collect high-order harmonics, and the more samples that are obtained, the more accurate classification will be. However, high sampling has a higher cost and takes longer for training. The model needs to add several different appliances and modify the design of the IM to make it more real such as adding some noise to the motor or adding several unseen appliances' data. However, the collected data was 60 seconds for the classification only which also need to obtain huge data but with a low sampling rate.

ACKNOWLEDGMENT

This study was made possible with the Fundamental Research Grant Scheme (FRGS/1/2020/TK0/UPNM/03/3) made available by the Ministry of Higher Education Malaysia.

REFERENCES

[1] J. Aguilar, A. Garces-Jimenez, M. D. R-Moreno, and R. García, "A systematic literature review on the use of artificial intelligence in energy self-management in smart buildings," *Renew. Sustain. Energy Rev.*, vol. 151, p. 111530, 2021, doi: 10.1016/j.rser.2021.111530.

[2] and J. D. Tao, Peng, XiAng Liu, Yangrui Zhang, Chong Li, "Multi-level non-intrusive load identification based on k-NN," *2019 IEEE 3rd Conf. Energy Internet Energy Syst. Integr.*, pp. 1905–1910, 2019, doi: 10.1109/EI247390.2019.9061896.

[3] D. Yang, S. Member, X. Gao, S. Member, and L. Kong, "An Event-Driven Convolutional Neural Architecture for Non-Intrusive Load Monitoring of Residential Appliance," *IEEE Trans. Consum. Electron.*, vol. 66, no. 2, pp. 173–182, 2020, doi: 10.1109/TCE.2020.2977964.

[4] A. Faustine and L. Pereira, "Improved appliance classification in non-intrusive load monitoring usingweighted recurrence graph and convolutional neural networks," *Energies*, vol. 13, no. 13, 2020, doi: 10.3390/en13133374.

[5] A. Moradzadeh, S. Zeinal-Kheiri, B. Mohammadi-Ivatloo, M. Abapour, and A. Anvari-Moghaddam, "Support vector machine-assisted improvement residential load disaggregation," *2020 28th Iran. Conf. Electr. Eng. ICEE 2020*, pp. 1–6, 2020, doi: 10.1109/ICEE50131.2020.9260869.

[6] and X. L. Li, Yaqian, Yuquan Yang, Kai Sima, Boyang Li, Tong Sun, "Non-intrusive load monitoring based on harmonic characteristics," *Procedia Comput. Sci.*, vol. 183, pp. 776–782, 2021, doi: 10.1016/j.procs.2021.02.128.

[7] H. T. Kelati, Amleset, Hossam Gaber, Juha Plosila, "Implementation of non-intrusive appliances load monitoring (NIALM) on k-nearest neighbors (k-NN) classifier," *AIMS Electron. Electr. Eng.*, vol. 4, no. 3, pp. 326–344, 2020, doi: 10.3934/ElectrEng.2020.3.326.

[8] and X. Y. Guohua, Wu, Yuan Diping, Yin Jiyao, Zeng Wenhua, Deng Peng, "Research on non-intrusive load monitoring based on random forest algorithm," *2020 4th Int. Conf. Smart Grid Smart Cities*, pp. 1–5, 2020, doi: 10.1109/ICSGSC50906.2020.9248565.

[9] A. A. Himeur, Yassine, Abdullah Alsalemi, Faycal Bensaali, "Effective non-intrusive load monitoring of buildings based on a novel multi-descriptor fusion with dimensionality reduction," *Appl. Energy*, vol. 279, p. 115872, 2020, doi: 10.1016/j.apenergy.2020.115872.

[10] Y. Li, B. Yin, P. Wang, and R. Zhang, "Non-intrusive load monitoring based on convolutional neural network mixed residual unit," *J. Phys. Conf. Ser.*, vol. 1176, no. 2, pp. 1–9, 2019, doi: 10.1088/1742-6596/1176/2/022052.

[11] F. I. H. and W. A. H. A. Shiddieqy, "Plug-load classification based on CNN from V-I trajectory image using STM32," *2021 Int. Symp. Electron. Smart Devices*, pp. 1–5, 2021, doi: 10.1109/ISESD53023.2021.9501919.

[12] Á. H. and R. N. L. de Diego-Otón, D. Fuentes-Jimenez, "Recurrent LSTM architecture for appliance identification in non-intrusive load monitoring," *2021 IEEE Int. Instrum. Meas. Technol. Conf.*, pp. 1–6, 2021, doi: 10.1109/i2mtc50364.2021.9460046.

[13] D. Kumar, "Performance analysis of three-phase induction motor with AC direct and VFD," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 331, no. 1, 2018, doi: 10.1088/1757-899X/331/1/012025.

[14] A. M. F. Jaramillo, Andres F. Moreno, David M. Laverty, D. John Morrow, Jesús Martinez del Rincon, "Load modelling and non-intrusive load monitoring to integrate distributed energy resources in low and medium voltage networks," *Renew. Energy*, vol. 179, pp. 445–466, 2021, doi: 10.1016/j.renene.2021.07.056.

[15] A. Kelati, H. Gaber, J. Plosila, and H. Tenhunen, "Implementation of non-intrusive appliances load monitoring (NIALM) on k-nearest neighbors (k-NN) classifier," *AIMS Elect*, vol. 4, no. 3, pp. 326–344, 2020, doi: 10.3934/ElectrEng.2020.3.326.