JJEE Volume 9 | Number 2 | June 2023 | Pages 228-240

Jordan Journal of Electrical Engineering

ISSN (print): 2409-9600, ISSN (online): 2409-9619

Homepage: jjee.ttu.edu.jo

Occupancy Detection in a Building Using Hybrid Models

Lohit Kumar Sahoo[1], Mitali Ray[2], Bibhu Prasad Ganthia[3*]

[1] Electrical and Electronics Engineering Department, Trident Academy of Technology, Bhubaneswar, Odisha, India
[2] Skill Development and Technical Education Department, Cuttack, Odisha, India
[3] Electrical Engineering Department, Indira Gandhi Institute of Technology, Sarang, Odisha, India
E-mail: jb.bibhu@gmail.com

Received: November 23, 2022 *Revised: January 7, 2023* *Accepted: January 11, 2023*

Abstract – Buildings consume over 40% of the world's total energy supply, and their occupancy is increasingly recognized as a major performance indicator due to its effect on a building's energy costs and occupant satisfaction. In this paper, a hybrid model is created to estimate future loads of a building with high efficiency and accuracy. The proposed model is composed of two - connected in a cascade - artificial neural networks, where the outcomes of the first network are fed into the second one, which in its turn performs the load forecasts. A pre-existing dataset is used to verify the proposed model and to test a variety of training set sizes. Analysis of the results is executed by taking six pair of combinations separately for both open door and closed door fault cases. In this analysis, cascaded back propagation and Elman back propagation method - among the rest of the analyzed methods – is found to give the best accuracy, i.e, 97.2% - 97.9%, which indicates that the suggested hybrid technique is more accurate than the existing non-hybrid methods.

Keywords – Occupancy detection; Prediction model; Hybrid neural network structure; Accuracy.

Abbreviations

ANN	Artificial neural network	OD	Open Door
NN	Neural network	CD	Closed Door
CFBP	cascaded forward back propagation	GA	Genetic algorithm
FF	Feed forward	SVM	Support vector machine
EBP	Elman back propagation	k-NN	K nearest neighbours
HVAC	Heating and ventilating air conditioning		

1. INTRODUCTION

The World Watch Institute reports that buildings use over 40% of global energy annually [1]. As a result, good building load management is critical for effective energy use and reduced energy consumption. Since the number of people in the space is a major consideration for managing the loads, occupancy sensors are installed. Because of this, occupancy rates have a big impact on how energy management is done.

Several studies and experts agree that accurate occupancy data is crucial to improving building energy staging and, hence, lowering energy consumption in buildings [2, 3]. In [4], the authors offer a heating and ventilating air conditioning (HVAC) system in which energy consumption is cut by 10-15% of manual switching control based on the number of people

* Corresponding author
Article's DOI: 10.5455/jjee.204-1669195686

present in the office building. Using a control ventilation approach in HVAC system [5, 6] can reduce energy consumption to 55% by taking human activity within the building into account.

A building's energy consumption is affected by a wide variety of factors [7, 8], such as the building's design, materials, equipment efficiency, environmental conditions, and occupancy rates. A proper examination of all the aspects impacting the energy consumption of buildings from the pre-construction phase to the post-construction phase is necessary for the construction of an efficient smart building. However, throughout the phase of building use, most of the elements cannot be readily managed or controlled by humans. Nonetheless, occupancy is a factor that can impact energy usage and is also within human control [9, 10]. The number of people living in a building has been determined using a dynamic method [11-15]. Occupancy detection in an office building was conducted in [11], where the presence accuracy of occupants was determined to be 95.8% and the number of occupants' identification accuracy was found to be 80.6%.

Occupancy detection in buildings is another research area where artificial neural networks (ANNs) have been applied. In [16], CO_2 data was used in conjunction with feed forward neural networks to determine the number of people present. The outcomes show an average accuracy of 70%. With the help of neural networks, the authors of [17, 18] were able to detect the number of people present with an accuracy of 75%. Pattern recognition of power consumption data for occupancy detection is examined in [19], where support vector machine (SVM) is pitted against kernels with linear, polynomial, and radial basis functions. In this case, the information has been gathered in a research facility, and the results show an accuracy of 55.37–79.12%.

Pattern matching in CO_2 concentration [20] was used to assess the impact of occupancy detection on indoor air quality, and the optimum methodology was found to have an accuracy of 82% true positives and a false positive rate of 22.5%. Using the relationship between CO_2 concentration, temperature, fresh air intake, and door opening and closing cases, the total number of individuals in the room may be calculated [21]. The accuracy ranges between 82.2% and 88.8%.

Some research has looked at unsupervised approaches for occupancy detection as a solution to the challenge of gathering a practical dataset since the algorithms can train without being fed information about whether or not the buildings are occupied [22-26]. In this case, smart metres were used to tally room usage. Therefore, this is another area that future research should focus on improving in order to remove obstacles to the widespread use of occupancy detection techniques.

The outcomes of artificial neural networks are more precise and efficient, and they may be used to solve issues of high complexity. Smart building efficiency may also be greatly improved by paying attention to occupancy levels. Because of this, hybrid models have been used to conduct an effective assessment of a building's energy use while taking occupant status into account. To solve this problem, hybrid models were developed by mixing any two of the three ANN networks suggested. The key highlights of this work are:

- Implementation of hybrid networks for occupancy detection in a building. two different cases, like an open door and a closed door.
- The impact on system performance in two different cases: open door and closed door.

- Performance comparison of the proposed approach with some of the other approaches considered in the literature.

This paper is organised as follows: section 2 describes the proposed models and used dataset; section 3 describes the methodology for hybrid model performance calculation; section 4 analyses the results; and sections 5 and 6 discuss comparative analysis and conclusion.

2. THE PROPOSED MODELS AND THE USED DATASET

In this paper, data is analyzed by considering three ANN methods: cascaded forward back propagation (CFBP) NN, feed forward (FF) NN, and Elman back propagation (EBP) NN, each having three layered 20-20-20-1 neurons with the Levenberg-Marquardt training algorithm. Because it is more efficient in small and medium-sized networks and patterns, the Levenberg-Marquardt training algorithm is used here. The transfer function used in each method has been set to tansig, tansig, tansig, and purelin, with a performance goal of 10e^{-5}. This combination is considered here because, after several tests, this combination of neurons and transfer function is giving better results. All these proposed training models are shown in Figs. 1 to 3, respectively. During training of the dataset, it has been observed that the cascaded forward back propagation network took 332 iterations and 491 s, the feed forward network took 498 iterations and 370 s, and the elman back propagation network took 415 iterations and 6154 s to train the network.

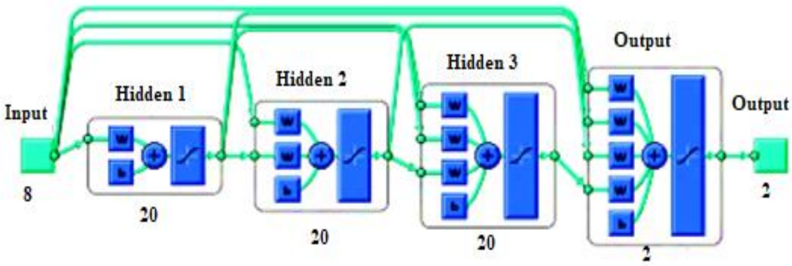


Fig. 1. The developed cascaded forward back propagation network.

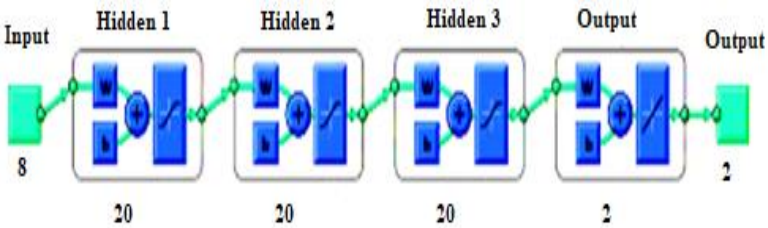


Fig. 2. The developed feed forward back propagation network.

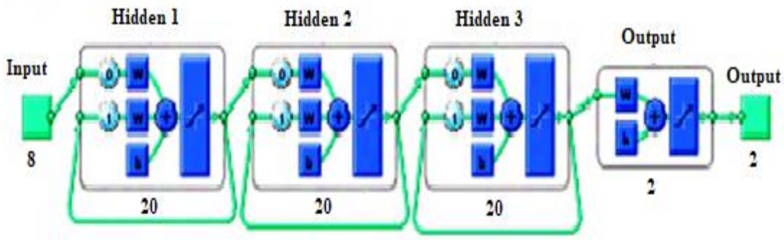


Fig. 3. The developed Elman back propagation network.

A multivariate time-series data set was used in this case. Seven numbers of attributes were considered in 20560 instances. Experimental data of temperature [°C], humidity, relative humidity [kg water-vapor/kg-air], light [Lux], and CO_2 [ppm] are collected from the sensor readings for a binary (1, 0) calculation of the room occupancy as an output by taking the exact date and time [year-month-day hour:minute:second].

The performance is evaluated after training and testing the dataset. Ground-truth occupancy was obtained from time-stamped pictures that were taken every minute. To know the dependency between multiple variables at a time, a correlation matrix is generally used. Before any data analysis, it is very important to find out the dependency or correlation between each variable and the others. The correlation of the used dataset is given in Fig. 4. Here, dark blue indicates a positive correlation. The colour intensity and circle size are proportional to the correlation coefficient. On the right side of the correlogram, the correlation coefficients and their respective colours are given.

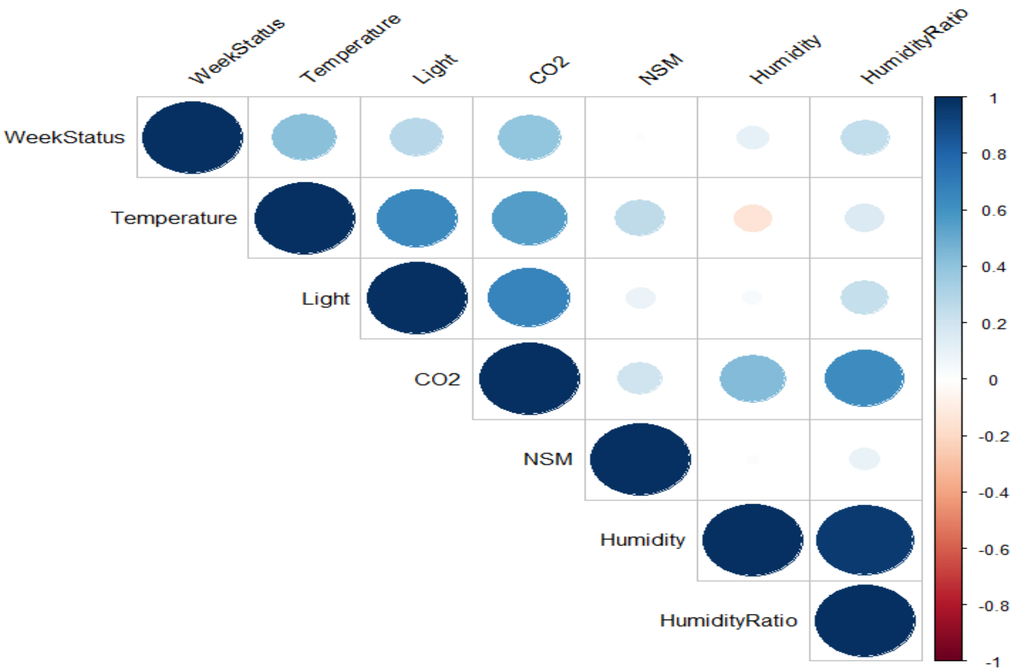


Fig. 4. Correlation of the used dataset.

After finding out the correlation between the taken dateset, the networks are combined in cascaded manner, as shown in Fig. 5.

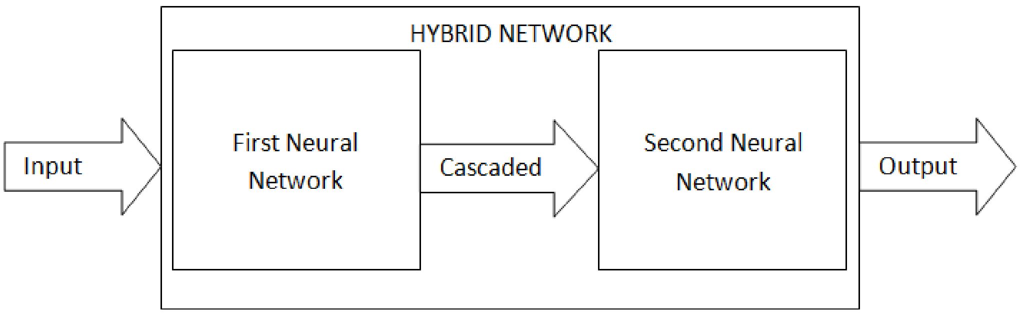


Fig. 5. Block diagram of the hybrid network.

3. METHODOLOGY FOR HYBRID MODEL PERFORMANCE CALCULATION

Here, any two of the proposed methods are assembled in a back-to-back manner. The output of one model is fed into the second model, and then the performance is evaluated in terms of confusion matrices. The hybrid structure is shown in Fig. 6.

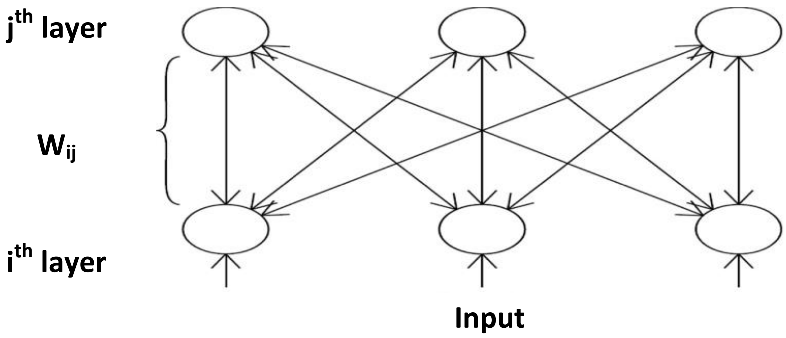


Fig. 6. Structure of hybrid network.

Flowchart of the algorithm is depicted in Fig. 7. It has the following steps:

- Step1: selection of input dataset
- Step2: training of dataset in first neural network
- Step3: if training is successful then go to next step. Otherwise go to step2
- Step4: training of dataset with other selected network
- Step5: if training is successful go to next step. Otherwise go to step2
- Step6: testing of data
- Step7: prediction of output

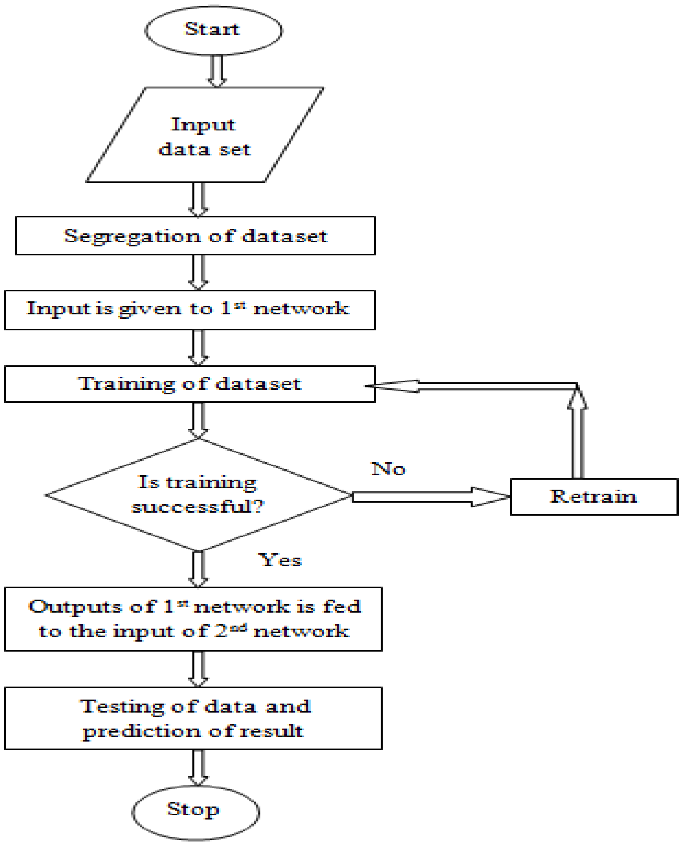


Fig. 7. Flowchart of the methodology for hybrid model performance calculation.

3.1. Training of Data

The sensor output data are created as input variables and fed to the network. The output of the network is set to either "0" for not occupied or "1" for occupied. The performance of the models is described by confusion matrices on a set of test data for which the actual values are known. The significance of different colors in a confusion matrix is shown in Fig. 8 where the Y-axis represents the actual value and the X-axis represents the predicted value and the precision can be expressed as follows [27]:

$$Accuracy = (TN + TP) / (TN + FN + TP + FP) \quad (1)$$

		Predicted condition	
	Total population = P + N	Positive (PP)	Negative (PN)
Actual condition	Positive (P)	True positive (TP)	False negative (FN)
	Negative (N)	False positive (FP)	True negative (TN)

Fig. 8. Confusion matrix [28].

Figs. 9 to 11 depict the performance plots of all proposed networks. It is observed that the performance of a cascaded forward back propagation network with performance goal of $10e^{-5}$ is achieved at 332 epochs. At 498 epochs, a feed-forward backward propagation network with a performance goal of $10e^{-5}$ is achieved. Similarly, for the Elman back propagation network, a performance goal of $10e^{-5}$ is achieved at 415 epochs.

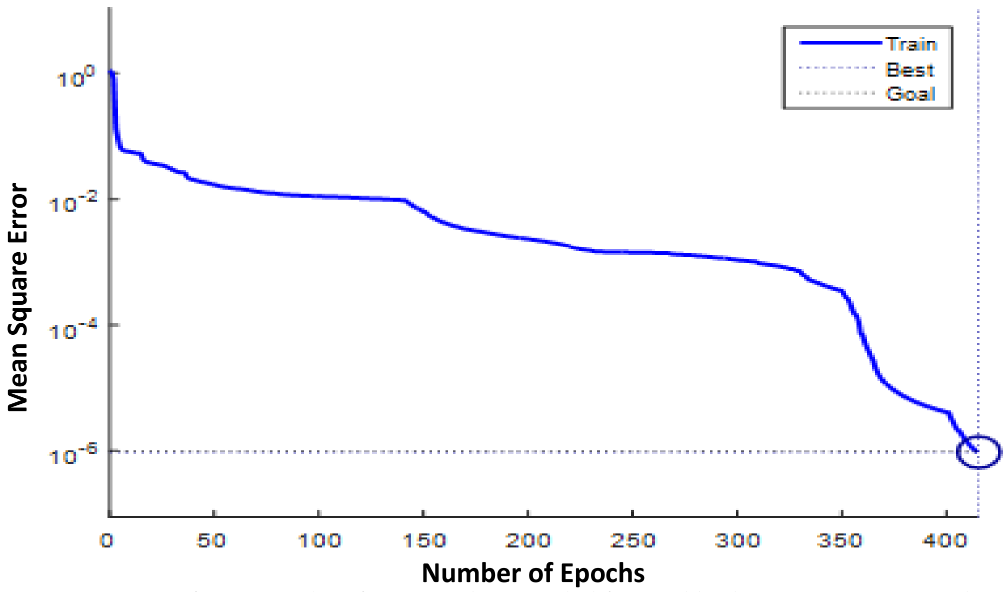


Fig. 9. Performance plot of training the cascaded forward back propagation network.

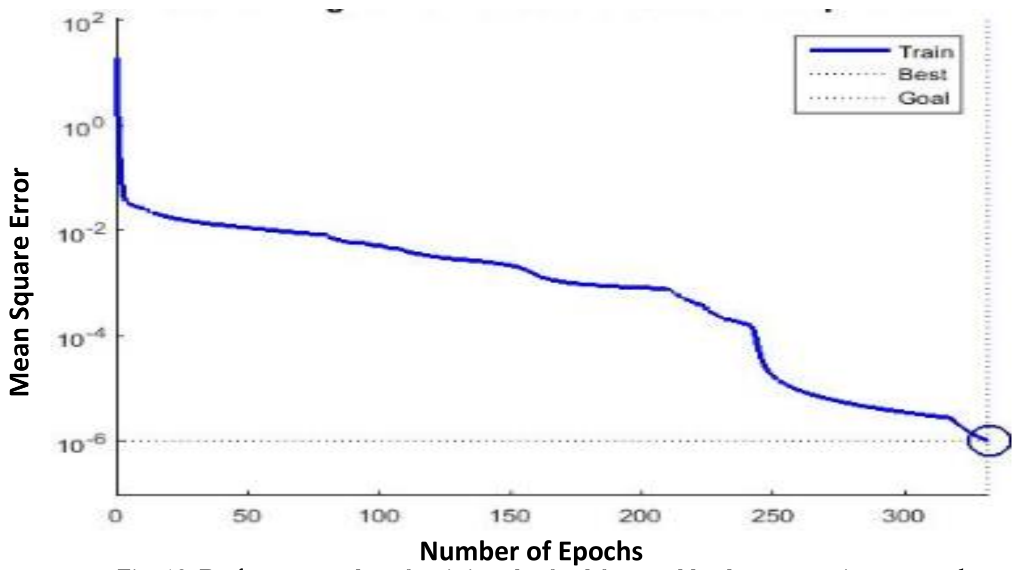


Fig. 10. Performance plot of training the feed forward back propagation network.

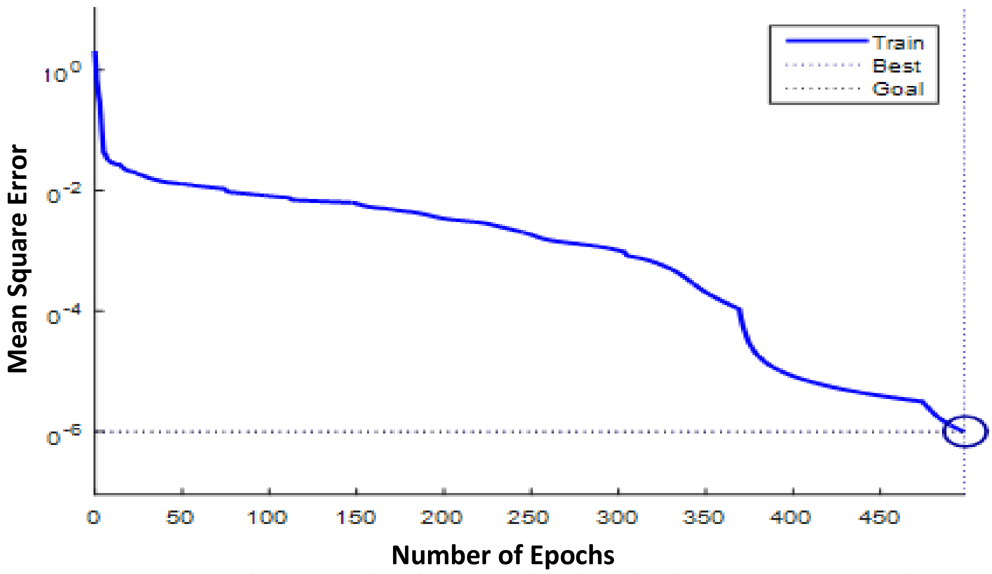
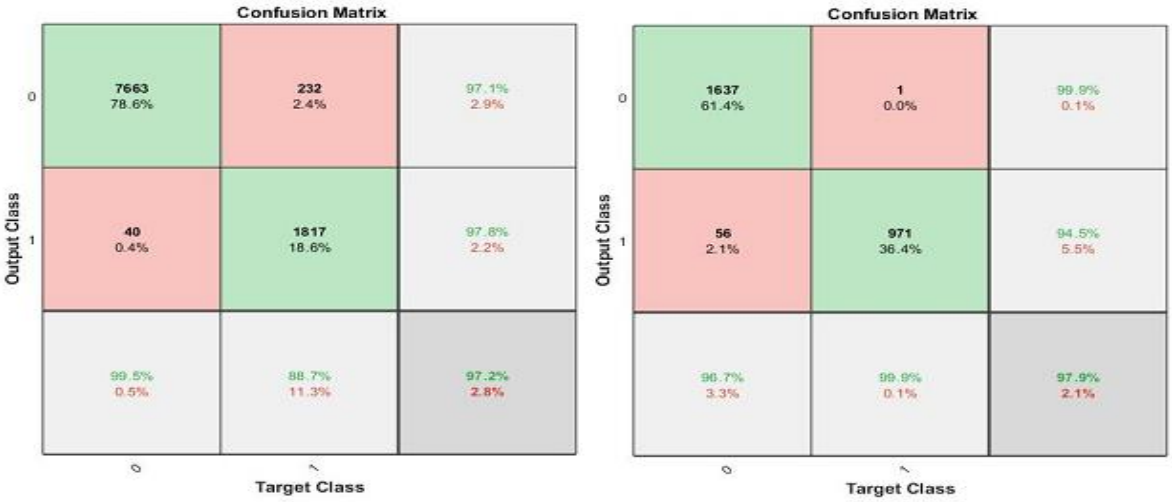


Fig. 11. Performance plot of training the Elman back propagation network.

3.2. Analysis of the Obtained Results

In hybridization, many neural networks are joined together to improve the expected output of a sequential data set. In this case, the best hybrid configuration is determined using the confusion matrix. Two fault cases are taken here: one with the door open (OD) and one with the door closed (CD). The hybrid intelligent model is designed by combining any two of the most accurate proposed models. Cascaded forward back propagation (CFBP) NN, feed forward back propagation (FF) NN, and Elman back propagation (EBP) NN with three layered 10-10-10-1 neurons with the Levenberg-Marquardt training algorithm, $10e^{-5}$ performance goal, and 0.01 learning rate have been taken here with the Tansig-tansig-tansig-purelin transfer function. For each combination, two of these mentioned algorithms are considered. Then, six different combinations, such as CFBP+EBP (OD, CD), CFBP+FF (OD, CD), EBP+CFBP (OD, CD), EBP+FF (OD, CD), FF+CFBP (OD, CD), and FF+EBP (OD, CD), are considered for hybridization. The respective confusion matrices are shown in Figs. 12 to 17. The performance analysis of these hybrid models is given in Table 1.

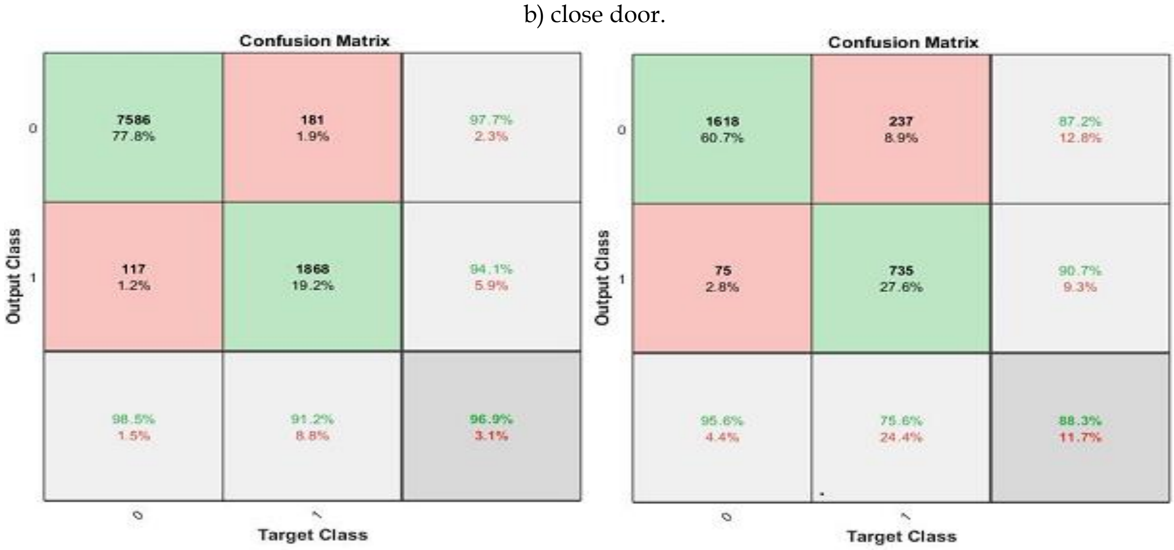


(a) **(b)**

Fig. 12. Confusion matrix of cascaded forward back propagation and Elman back propagation: a) open door; b) close door.

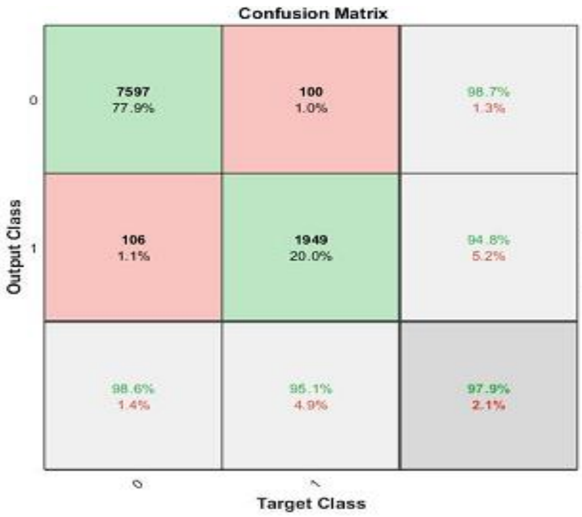
(a) **(b)**

Fig. 13. Confusion matrix of the cascaded forward back propagation and feed forward: a) open door; b) close door.

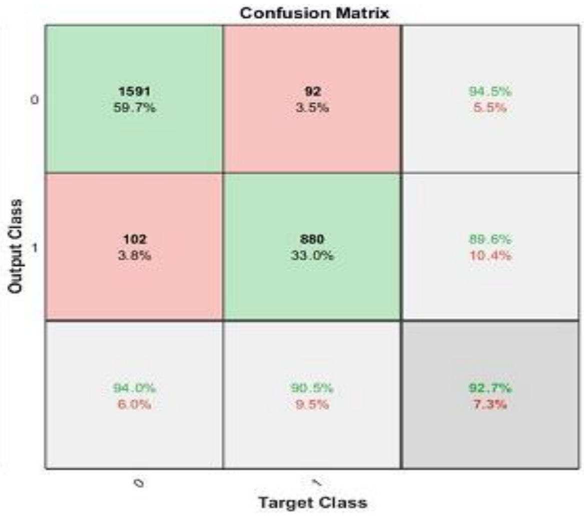


(a) **(b)**

Fig. 14. Confusion matrix of Elman back propagation and cascaded forward back propagation: a) open door; b) close door.

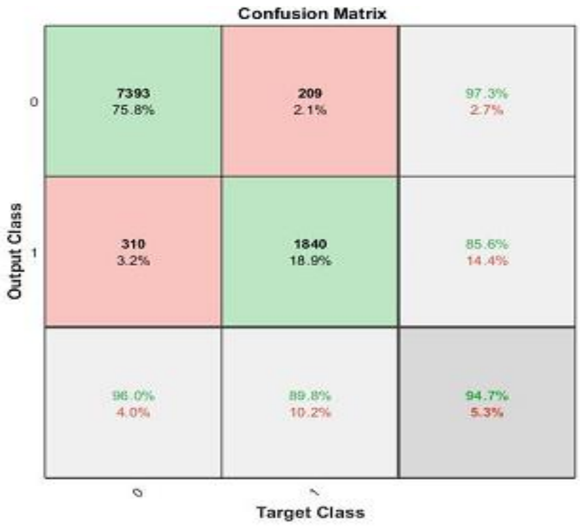


(a)



(b)

Fig. 15. Confusion matrix of Elman back propagation and feed forward: a) open door; b) close door.



(a)

(b)

Fig. 16. Confusion matrix of feed forward and cascaded forward back propagation: a) open door; b) close door.

(a)

(b)

Fig. 17. Confusion matrix of feed forward and Elman back propagation: a) open door; b) close door.

Table 1. Performance analysis of the hybrid models.

Figure number	Methodology	Open Door (OD)	Close Door (CD)
Fig. 12	CFBP+EBP	97.2%	97.9%
Fig. 13	CFBP+FF	96.7%	96.7%
Fig. 14	EBP+CFBP	96.9%	88.3%
Fig. 15	EBP+FF	97.9%	92.7%
Fig. 16	FF+CFBP	94.7%	94.2%
Fig. 17	FF+EBP	96.5%	93.5%

In Table 1, the performance analysis of all the six combinations is given for both door open and door closed cases. We can observe that CFBP+EBP combination is giving accuracy of 97.2% when door is open and 97.9% when door is closed, CFBP+FF combination is giving accuracy of 96.7% when door is open and 96.7% when door is closed, EBP+CFBP combination is giving accuracy of 96.9% when door is open and 88.3% when door is closed, EBP+FF combination is giving accuracy of 97.9% when door is open and 92.7% when door is closed, similarly FF+CFBP is giving accuracy of 94.7% and 94.2% in open door and closed door cases, respectively, and FF+EBP is giving accuracy of 96.5% and 93.5% in open door and closed door cases, respectively. From this result analysis, we can conclude that hybrid combination of cascaded forward back propagation fed to elman back propagation model (CFBP+EBP) has given the best result, which can be described as:

a) Performance with door opened

The proposed ANN based model is tested against the trained network with the data when door is opened. Fig. 12(a) represents the accuracy in terms of confusion matrix. The Figure shows that out of 9762 fault cases, 7663 were correctly classified as door open whereas 1817 cases were door closed. The overall accuracy of the proposed method is 97.2% for all the tested fault cases.

b) Performance with door closed

The proposed ANN based model is tested against the trained network with the data when the door is closed. Fig. 12(b) represents the accuracy in terms of confusion matrix. The Figure shows that out of 2665 fault cases, 1637 were correctly classified as door open whereas 971 cases were door closed. The overall accuracy of the proposed method is 97.9% for all the tested fault cases.

3.3. Comparative Analysis of the Obtained Results with Other Works

Various methods have been proposed for occupancy detection by different researchers. In [19], support vector machine (SVM) is compared with different kernels for occupancy detection. Here, the data has been collected in a research centre and the accuracy has been found in between 55.37% and 79.12%. Genetic algorithm (GA) is used in classification as well as for regression [29]. With the application of genetic algorithm, high predictive power of high-frequency smart meter is also highlighted here. Here, the accuracy is found to be 90%. SVM uses the kernel trick technique. Kernels change a non-separable problem to a separable problem and are widely used in nonlinear separation problem. K-Nearest-Neighbours (k-NN) is another technique used for classification [30]. K-NN has less computation time where the accuracy is found to be 80% to 88%. The proposed method has an accuracy between 97.2% - 97.9%.

Table 2 shows the clear idea about this comparative analysis of different models.

Table 2. Comparison with other models.

Model`	Accuracy
SVM [19]	55.37% to 79.12%
GA [29]	90%
SVM, k-NN [30]	80% to 88%
Proposed method (CFBP+EBP)	97.2% to 97.9%

4. CONCLUSIONS

This paper represented a hybrid technique that has been exploited to approximate buildings' occupancy data. In the literature review section, we have presented a discussion about different occupancy detection methodologies and their advantages and limitations. Then, three different networks are considered for a more efficient and accurate investigation. After doing the proper result analysis, the number of neurons, hidden layer, activation function, transfer function, and performance goal are set at a particular value at which the models are giving their most efficient result. The result analysis was then performed by taking six pairs of combinations, separately for open door and closed door fault cases. In this analysis, we have observed that the cascaded back propagation + Elman back propagation (CFBP+EBP) method gives the best accuracy, i.e., 97.2%–97.9%, among the other proposed methods. Information about occupants is the key to improving building energy performance and reducing building energy consumption, along with occupants' comfort. A significant amount of energy can also be saved by adopting occupancy-based control strategies. Although considerable progress has been made in occupancy detection systems, further work is required to achieve viable applications.

REFERENCES

[1] I. Yüksek, T. Karadayi, "Energy efficient building design in the context of building life cycle," *Energy Efficient Buildings*, pp. 93-123, 2017.

[2] T. Nguyen, M. Aiello, "Energy intelligent buildings based on user activity: a survey," *Energy and Buildings*, vol. 56, pp. 244–257, 2013.

[3] Y. Zhang, X. Bai, F. Mills, J. Pezzey, "Rethinking the role of occupant behavior in building energy performance: a review," *Energy and Buildings*, vol. 172, pp. 279-294, 2018.

[4] Y. Agarwal, B. Balaji, R. Gupta, J. Lyles, M. Wei, T. Weng, "Occupancy-driven energy management for smart building automation," *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, pp. 1–6, 2010.

[5] M. Jin, N. Bekiaris-Liberis, K. Weekly, C. Spanos, A. Bayen, "Occupancy detection via environmental sensing," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 2, pp. 443–455, 2018.

[6] M. Jin, N. Bekiaris-Liberis, K. Weekly, C. Spanos, A. Bayen, "Sensing by proxy: occupancy detection based on indoor co2 concentration," *UBICOMM 2015 : The Ninth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, vol. 14, pp. 1-10, 2015.

[7] J. Conti, P. Holtberg, J. Diefenderfer, A. LaRose, J. Turnure, L. Westfall, "International energy outlook 2016 with projections to 2040," *USDOE Energy Information Administration, Office of Energy Analysis*, Washington, United States, 2016.

[8] H. Yoshino, T. Hong, N. Nord, "Iea ebc annex 53: total energy use in buildings—analysis and evaluation methods," *Energy and Buildings*, vol. 152, pp. 124–136, 2017.

[9] M. Jia, R. Srinivasan, A. Raheem, "From occupancy to occupant behavior: an analytical survey of data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency," *Renewable and Sustainable Energy Reviews*, vol. 68, pp. 525–540, 2017.

[10] R. Melfi, B. Rosenblum, B. Nordman, K. Christensen, "Measuring building occupancy using existing network infrastructure," *Green Computing Conference and Workshops*, pp. 1–8, 2011.

[11] D. Calì, P. Matthes, K. Huchtemann, R. Streblow, D. Müller, "Co2 based occupancy detection algorithm: experimental analysis and validation for office and residential buildings," *Building and Environment*, vol. 86, pp. 39–49, 2015.

[12] S. Wang, X. Jin, "Co2-based occupancy detection for on-line outdoor air flow control," *Indoor and Built Environment*, vol. 7, no. 3, pp. 165–181, 1998.

[13] H. Xie, Y. Wang, Z. Gao, B. Ganthia, C. Truong, "Research on frequency parameter detection of frequency shifted track circuit based on nonlinear algorithm," *Nonlinear Engineering*, vol. 10, no. 1, pp. 592-599, 2021.

[14] D. Chicco, G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, pp. 1-13, 2020.

[15] M. Mahmoud, "Accurate sustainable indoor-lighting design using simultaneous dual control to adjust LEDs lighting and roller blinds' opening to control the daylight for economic power consumption," *Preprints*, 2022.

[16] M. Zuraimi, A. Pantazaras, K. Chaturvedi, J. Yang, K. Tham, S. Lee, "Predicting occupancy counts using physical and statistical co2-based modeling methodologies," *Building and Environment*, vol. 123, pp. 517–528, 2017.

[17] J. Gu, W. Wang, R. Yin, C. Truong, B. Ganthia, "Complex circuit simulation and nonlinear characteristics analysis of GaN power switching device," *Nonlinear Engineering*, vol. 10, no. 1, pp. 555-562, 2021.

[18] K. Lam, M. Höynck, B. Dong, B. Andrews, Y. Chiou, R. Zhang, D. Benitez, J. Choi, "Occupancy detection through an extensive environmental sensor network in an openplan office building," *IBPSA Building Simulation*, vol. 145, pp. 1452–1459, 2009.

[19] B. Ganthia, S. Barik, B. Nayak, "Shunt connected FACTS devices for LVRT capability enhancement in WECS," *Engineering, Technology and Applied Science Research*, vol. 10, no. 3, pp. 5819-5823, 2020.

[20] A. Szczurek, M. Maciejewska, A. Wyłomanska, R. Zimroz, G. Zak, A. Dolega, "Detection of occupancy profile based on carbon dioxide concentration pattern matching," *Measurement*, vol. 93, pp. 265–271, 2016.

[21] B. Ganthia, S. Barik, "Steady-state and dynamic comparative analysis of PI and fuzzy logic controller in stator voltage oriented controlled DFIG fed wind energy conversion system," *Journal of The Institution of Engineers (India): Series B*, vol. 101, no. 3, pp. 273-286, 2020.

[22] A. Ebadat, G. Bottegal, D. Varagnolo, B. Wahlberg, H. Hjalmarsson, K. Johansson, "Blind identification strategies for room occupancy estimation," *2015 European Control Conference*, pp. 1315–1320, 2015.

[23] B. Ganthia, R. Pradhan, R. Sahu, A. Pati, "Artificial ant colony optimized direct torque control of mathematically modeled induction motor drive using pi and sliding mode controller," *Recent Advances in Power Electronics and Drives*, pp. 389-408, 2021.

[24] A. Cardenas, S. Piche, D. Meunier, L. Rueda, K. Agbossou, "On the occupancy measurement and analysis for residential applications," *2019 IEEE 28th International Symposium on Industrial Electronics*, pp. 2189-2194, 2019.
[25] G. Tang, K. Wu, J. Lei, W. Xiao, "The meter tells you are at home! non-intrusive occupancy detection via load curve data," *2015 IEEE International Conference on Smart Grid Communications*, pp. 897–902, 2015.
[26] M. Jin, R. Jia, Z. Kang, I. Konstantakopoulos, C. Spanos, "Presencesense: zero-training algorithm for individual presence detection based on power monitoring," *in Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, pp. 1–10, 2014.
[27] D. Chicco, N. Toetsch, G. Jurman, "The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation," *BioData Mining,* vol. 14, no. 1, pp. 1-22, 2021.
[28] B. Ganthia, S. Barik, B. Nayak, "Wind turbines in energy conversion system: types and techniques," *In Renewable Energy and Future Power Systems,* pp. 199-217, 2021.
[29] R. Razavi, A. Gharipour, M. Fleury, I. Akpan, "Occupancy detection of residential buildings using smart meter data: a large-scale study," *Energy and Buildings*, vol. 183, pp. 195-208, 2019.
[30] W. Kleiminger, T. Staake, S. SantinI, "Occupancy detection from electricity consumption data," *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, pp. 1–8, 2013.