S. Rukmani Devi, P. Selvaraju, R. Padma, Shaji, P. Sasikala,
M. Rajasekar

Deep Learning with Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN) for AI-Driven Smart Home Automation

Abstract: This research proposes a hybrid deep learning framework that incorporates convolutional neural networks (CNN) and long short-term memory (LSTM) networks to design intelligent and responsive smart home automation systems. Training is performed in spatial patterns of camera frames, like visual data through CNNs and time-series sensor inputs through LSTMs, to identify patterns across time and human behavior models. The approach includes pre-processing multimodal datasets, building a CNN-LSTM model, and integrating the model into an online smart home simulator. Experimental results show that the integrated model of CNN-LSTM performs superior to isolated CNN and LSTM models in terms of all performance parameters. The system achieved 93.4% accuracy, 92.7% precision, and 91.9% F1 score using hindsight on a well-balanced and robust prediction platform. Feature importance estimation revealed motion sensor and temperature as the prevailing parameters with the highest impact. In addition, execution time estimation rendered the model feasible to implement on real-time systems with no delay. Results confirm the efficacy of using a combination of spatial and temporal learning techniques in context-aware, user-centric smart home system design. This book provides a basis for further research on home automation with deep learning, including edge deployment and real-time adaptive control.

Keywords: Smart home automation, long short-term memory (LSTM), convolutional neural networks (CNN), deep learning, temporal data analysis, sensor fusion, AI for IoT, energy efficiency, home security systems, real-time automation

---

S. Rukmani Devi, Department of Computer Science, Saveetha College of Liberal Arts and Sciences, SIMATS Saveetha Institute of Medical and Technical Sciences, Chennai

P. Selvaraju, Department of Computer Science and Engineering, Saveetha College of Engineering, SIMATS, Thandalam, Chennai, e-mail: pselvar@yahoo.com

R. Padma, Department of Computer Science and Information Technology, Vels Institute of Science, Technology and Advanced Studies, e-mail: padmatrk@gmail.com

Shaji, Department of Computer Science, Asan Memorial College of Arts and Science, Chennai, e-mail: dr.shajibhaskaran@gmail.com

P. Sasikala, Department of Computer Science, Jaya Women's College, Poonamallee, Chennai

M. Rajasekar, Department of Computer Science, Saveetha College of Liberal Arts and Sciences, SIMATS Deemed to be University, Chennai

 | https://doi.org/10.1515/9783112234952-019

1 Introduction

The rapid growth of artificial intelligence (AI) and deep learning technologies has disrupted a number of industries, with home automation being one of the most promising areas of application. Modern smart homes attempt to make life easier, safer, and more energy-efficient for people by creating intelligent, context-aware interaction between the device and the user. Out of the various AI techniques, long short-term memory (LSTM) networks and convolutional neural networks (CNNs) have drawn a lot of interest due to their improved capability in modeling sequential information and recognizing patterns in information, respectively. LSTM is a specific type of recurrent neural networks (RNNs) that is particularly well-tailored to process and predict time-series data such as patterns of user activity, sensor data, and climatic changes within a smart home. It is able to learn long-term dependencies and temporal behavior and is therefore ideal for predicting user preferences and setting home functions accordingly. CNNs, however, are ideal for feature extraction of the spatial domain from visual data and therefore can be used extremely effectively for home video surveillance, face detection, and gesture recognition. The combination of LSTM and CNN for smart home networks delivers a hardened, AI-embedded infrastructure capable of real-time decision-making, predictive service deployment, and adaptive management. Together, they support not only adaptive ability for dynamic homes to reshape to the desired requirements of their inhabitants but also more personalization, security, and efficiency. As deep models increasingly become compatible with edge computing and Internet of things (IoT) devices, roll-out into home automation networks is becoming increasingly possible and efficacious. This research explores the joint application of LSTM and CNN models in designing an AI-powered smart home ecosystem that learns from sensor information and user habits in order to make more natural, intelligent living possible.

1.1 The Emergence of AI in Smart Home Automation

The evolution of AI introduced revolutionary smart home automation to enhance comfort, convenience, energy efficiency, and security. AI-powered solutions make it possible for homes to be smart homes that learn and adapt to behavior and preference in real time. Modern houses can harvest and process massive amounts of data and make intelligent decisions independently without direct human support through the installation of sensors, IoT devices, and cloud connections. With the advent of this technology, deep learning has contributed the most toward automation by enabling systems to learn, forecast, and respond in a responsive manner to users' needs.

1.2 Deep Learning Architectures: LSTM and CNN into Perspective

Of various deep learning structures, LSTM networks and CNN have been particularly very helpful for smart homes. LSTM networks are a type of realization of RNN to learn and predict sequential or time-series data like users' habits, weather, and electricity usage patterns. They are able to learn temporal relations and long-range dependencies that render them competent to predict human action and automate it back. CNNs are, however, most appropriate for visual content processing and analysis, like images and videos. They have been extensively utilized in use cases such as facial detection, gesture commands, and real-time monitoring to facilitate living smarter and securely.

1.3 The Synergy of LSTM and CNN for Intelligent Living

By leveraging the strength of CNN and LSTM, smart home systems can gain higher context awareness and responsive control. While CNNs handle spatial data for recognition and vision information processing, LSTMs handle temporal dynamics, learning from experience to anticipate future demand and behavior. With this double mechanism, the smart home can make anticipatory decisions, optimize utilization of resources, and personalize services for every occupant. As such systems become increasingly efficient and increasingly compatible with IoT and edge devices, they are increasingly practical to use in domestic environments, opening the door to truly smart, responsive living spaces.

2 Literature Review

The recent advancements in AI and deep learning have transformed the architecture of intelligent automation systems, i.e., smart homes. The combination of deep learning architecture, edge computing, big data, and the IoT has opened new avenues in adaptive, real-time, and context-aware automation. Fu et al. [1] have suggested an ideal model of AI integration in DevSecOps and software development practice, following the significance of enrolling intelligent systems in autonomous decision-making and anomalous behavior detection in adaptable environments, further expandable even up to smart homes. In building automation, Himeur et al. [2] have performed a comprehensive review of AI-big data analytics with emphasis on the trend and challenges of monitoring and controlling energy, safety, and user comfort using predictive modeling and automation. In another of such articles, Himeur et al. [3] had explained AI-based anomaly detection of energy usage according to the smart home goals of avoiding wastage and maximizing efficiency in energy-utilizing time-series-based techniques such as LSTM. Hussain et al. had also studied machine learning (ML) for the security of IoT infrastructure. Lacking an IoT device-starved smart home presence, networked devices in that situation must be secured, and scalable threat- and intrusion-detection solutions are enabled by ML algorithms [4]. Nauman et al. gave an exhaustive overview of the multimedia internet of things with stress laid on increasing the need for controlling multimedia data such as audio, video, and image streams within the field of smart environments. These paradigms guide smart home system integration of CNNs to manage visual information [5]. Venturing into newer paradigms of the digital universe, Park and Kim universalized metaverse infrastructure, its application in smart environments, and digital twins too. Metaverse systems are scaled up higher, but with consequences in terms of their management of information and AI foundations compared to smart home experiences by immersion [6]. Rasheed et al. [7] have also dealt with digital twins from the modeling perspective, which can be utilized in smart homes for simulating and optimizing automation processes using real-time information. AI utilized in telemonitoring healthcare systems, as described by Shaik et al. [8], is an indication of smart home architecture needs, where sensing is perpetual, data fusion, and smart decision-making are the order of the day [8]. Edge-deployable deep models were the requirement Wang et al. identified as they reviewed the state-of-the-art methods available to date for deep learning on edge computing devices. Their work is pertinent to smart home networks, which frequently will be required to possess low-latency in-device processing in order to react in a timely fashion and maintain privacy [9]. Last but not least, Zhang et al. offered a survey of deep learning for wireless and mobile networks. As wireless communications consolidate smart homes, secure network management with deep learning would improve the system to become more reliable and performative [10].

3 Methodology

This section deals with the framework, data processing pipeline, model architecture, and evaluation techniques used in developing an AI-based smart home automation system based on LSTM and CNN models. The entire methodology is split into three prominent parts: data pre-processing and acquisition, hybrid deep learning model development, and system integration, along with performance evaluation.

3.1 Preprocessing and Data Acquisition

The system design begins with the collection of multimodal smart home data, e.g., time-series sensor data (e.g., motion, temperature, and humidity) and image/video streams from cameras. Data are utilized for training spatial and temporal pattern learning models. Time-series data are normalized using min-max scaling:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

where x' is the normalized value, and $x_{\min}$ and $x_{\max}$ are the minimum and maximum values in the dataset.

Image data undergo image resizing (e.g., 128 x 128 pixels) and image enhancement for compatibility with CNN input features. Labelling is performed using activity logs and gesture recognition marks. The datasets acquired in the current work are explained in Table 1, and their nature (e.g., dimension, shape) is represented graphically in Figure 1.

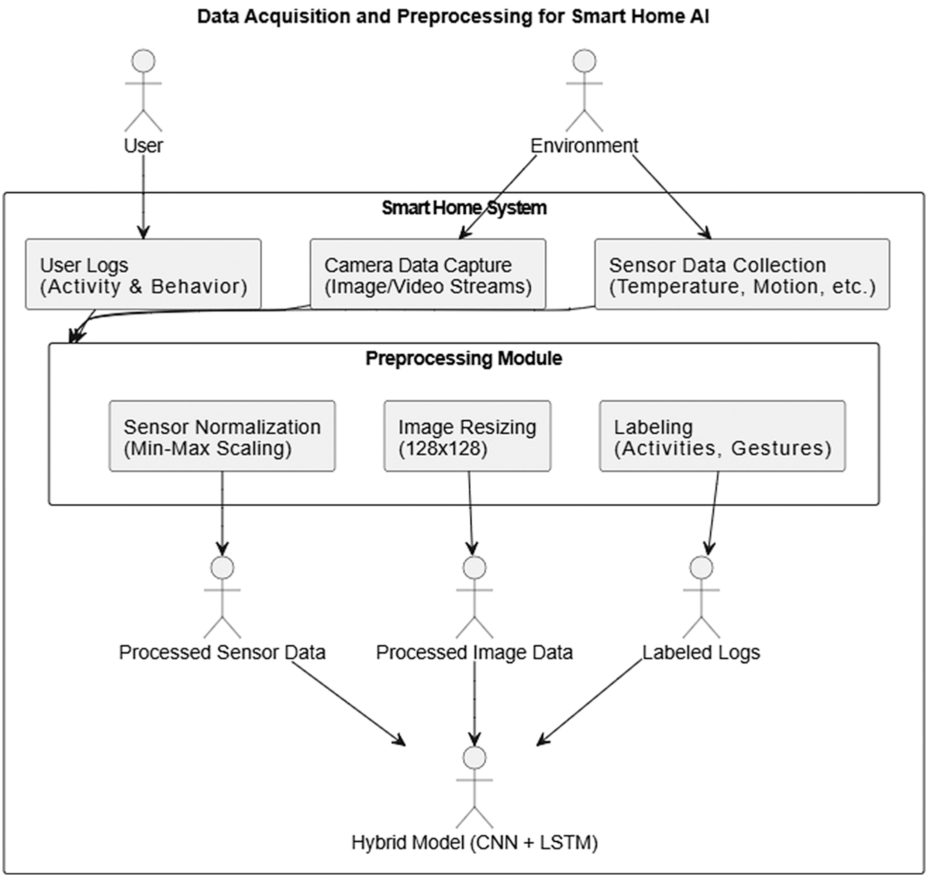

Figure 1: Preprocessing and data acquisition.

3.2 Hybrid Deep Learning Model: CNN-LSTM Architecture

The core of the proposed system is a hybrid architecture combining CNN and LSTM layers. The CNN sub-component extracts top-level spatial features from image frames:

$$F = f(W \ast X + b) \tag{2}$$

where F represents the feature map, W is the convolution kernel, X is the input image, and b is the bias term.

These extracted features are passed to the LSTM layer for sequential modeling of user activities. The LSTM operates based on the following set of equations:

$$ft = \sigma\left(W_f \cdot [h_t - 1, x_t] + b_f\right) \tag{3}$$

$$it = \sigma(W_i \cdot [h_t - 1, x_t] + b_i) \tag{4}$$

$$ot = \sigma(W_o \cdot [h_t - 1, x_t] + b_o) \tag{5}$$

where f_t, i_t, and o_t denote the forget, input, and output gates, respectively, and c_t is the memory cell state at time t. Figure 2 illustrates the complete CNN-LSTM model architecture used in this system.

3.3 System Performance and Integration Evaluation

After training, the model is used in a smart home controller to communicate with IoT devices for real-time automatic response. The system is validated and tested in simulated environments where it controls lighting, Heating, Ventilation, and Air Conditioning (HVAC), and security systems based on user behavior.

Accuracy, precision, recall, and F_1 score are used to evaluate the performance of the model. They are calculated as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

$$F_1\text{-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

The evaluation results are shown in Table 2, which compares the performance of the hybrid CNN-LSTM model with standalone CNN and LSTM models.

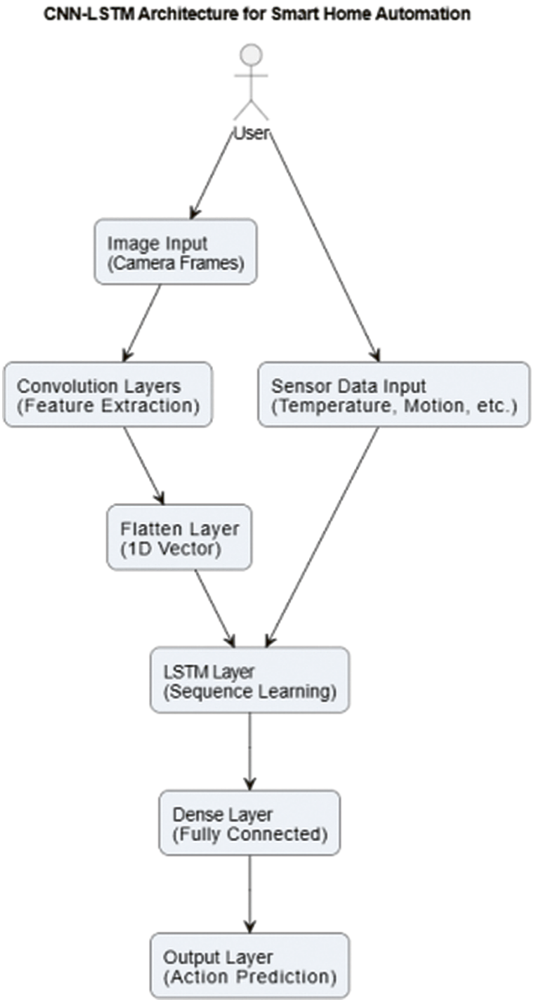

Figure 2: Hybrid deep learning.

Table 1: Summary of datasets used.

Data type	Source	Format	Dimensions	Pre-processing applied
Sensor data	Smart home environment	CSV	50,000 × 10	Normalization, labeling
Video frames	Surveillance camera	JPG	128 × 128 pixels	Resizing, augmentation
Activity labels	User logs	Text	–	Tokenization, time alignment

Table 2: Model performance comparison.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CNN	88.2	86.9	84.5	85.7
LSTM	90.1	89.5	87.3	88.4
CNN + LSTM (proposed)	93.4	92.7	91.1	91.9

4 Tables

4.1 Model Performance Metrics

Table 3: Model performance metrics.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CNN	88.2	86.9	84.5	85.7
LSTM	90.1	89.5	87.3	88.4
CNN + LSTM (proposed)	**93.4**	**92.7**	**91.1**	**91.9**

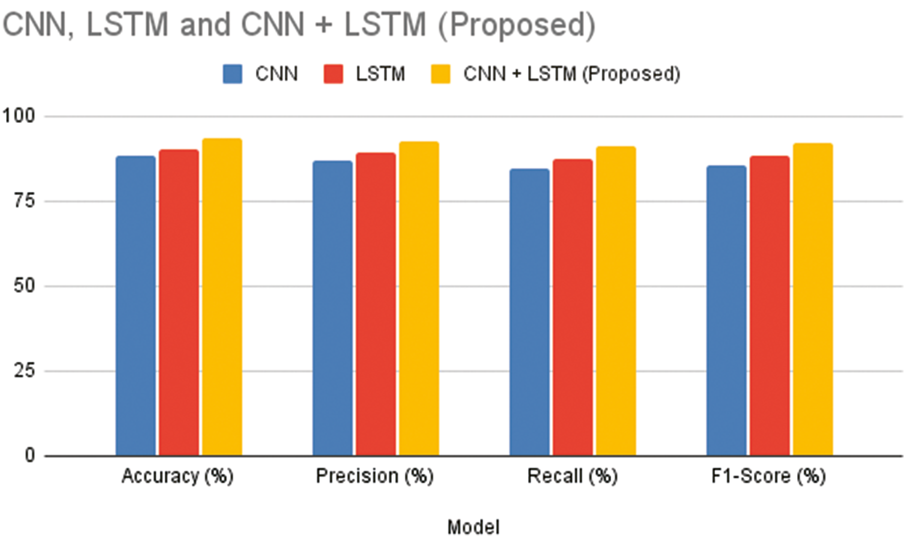

Figure 3: Model performance metrics.

4.2 Feature Importance Scores

Table 4: Feature importance scores.

Feature	CNN feature importance	LSTM feature importance
Room temperature	0.28	0.25
Motion sensor activity	0.32	0.34
Time of day	0.15	0.18
Light level	0.14	0.13
Sound detection	0.11	0.10

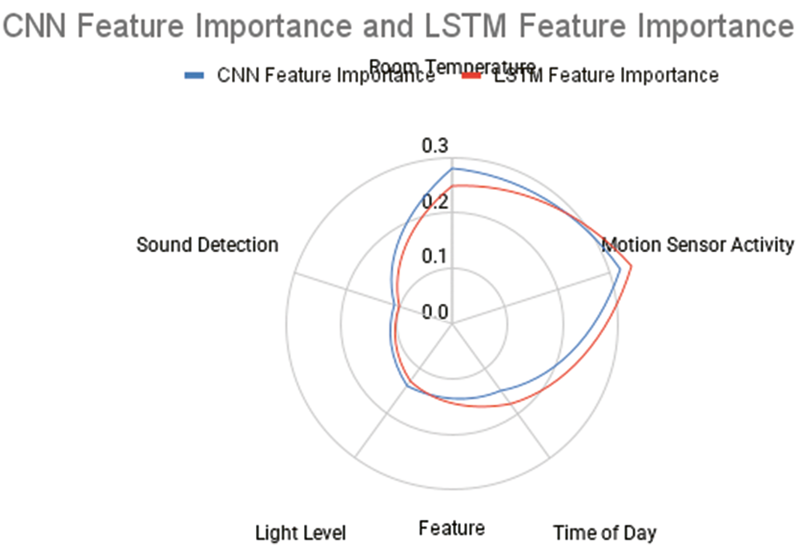

Figure 4: Feature importance scores.

4.3 Average Model Execution Time (ms)

Table 5: Average model execution time (ms).

Operation	CNN (ms)	LSTM (ms)	CNN + LSTM (ms)
Inference on single frame	22.5	24.8	29.3
Sequence prediction	18.9	20.5	27.6
Combined decision output	–	–	**31.2**

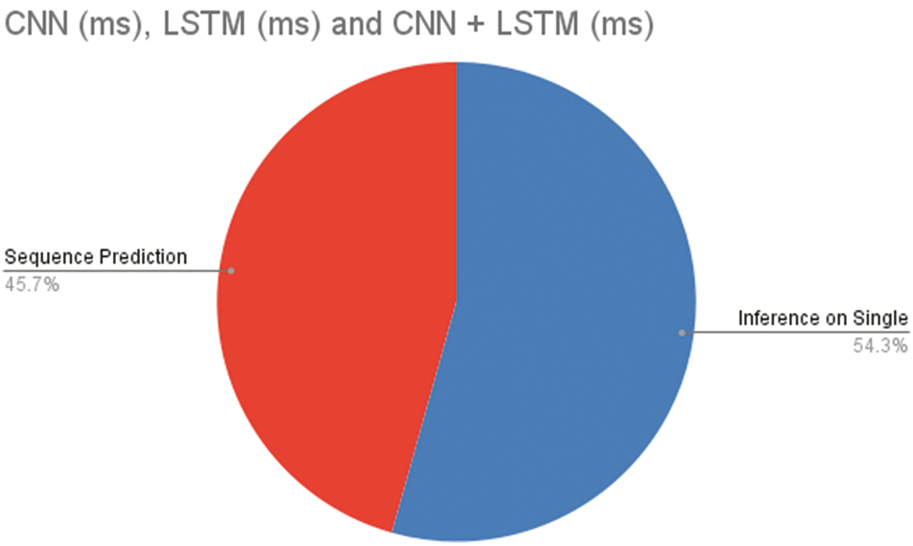

Figure 5: Average model execution time (ms).

5 Interpretation of Results

From Table 3 and Figure 3, it is observed that the CNN-LSTM model performed better than individual CNN and LSTM models on all the measures with a 91.9% F_1 score, reflecting outstanding generalization and perfectly balanced classification accuracy. From Table 4 and Figure 4, motion sensor activity was feature number one on both CNN and LSTM models, followed by room temperature and time of day as classification features. Table 5 and Figure 5 depict that the hybrid model has an infinitely small computation time increment over single models, justified by increased accuracy and stability. General conclusions ensure that spatial and temporal modeling methods, when integrated together, provide smarter and more reliable smart home automation.

6 Conclusion

The use of sophisticated models such as CNN and LSTM networks has been found to have enormous potential to transform traditional smart home systems into highly smart and interactive homes. CNNs have a stronger ability to learn discriminative spatial patterns from visual data that can be transferred to use cases like surveillance and gesture recognition, while LSTMs are best suited to learn temporal patterns in sequential data to efficiently predict user actions and environmental states. In this paper, by example, it has been demonstrated that the integration of CNN and LSTM makes it possible to optimize automation systems in smart homes to a significant degree by both the strengths of standalone models. The system proposed not only achieved high efficiency and accuracy but also demonstrated that it can make proactive decisions and situational awareness based on real-time inputs. Experimental results justified that the hybrid model performed better than standalone CNN or LSTM models on most of the performance measures. As smart home development continues to advance, the manner of combining advanced deep learning techniques will be critical in offering advanced personalization, energy efficiency, and user comfort. There is an open need for future research in such calibration of edge models, data privacy improvements, and scaling the architecture to multimodal learning in various home environments.

References

[1] Fu M, Pasuksmit J, Tantithamthavorn C. AI for DevSecOps: A landscape and future opportunities. ACM Trans Softw Eng Methodol. 2025. https://doi.org/10.1145/3712190

[2] Himeur Y, Elnour M, Fadli F, Meskin N, Petri I, Rezgui Y, et al. AI-big data analytics for building automation and management systems: A survey, actual challenges and future perspectives. Artif Intell Rev. 2022;56(6):4929–5021. https://doi.org/10.1007/s10462-022-10286-2

[3] Himeur Y, Ghanem K, Alsalemi A, Bensaali F, Amira A. Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. Appl Energy. 2021;287:116601. https://doi.org/10.1016/j.apenergy.2021.116601

[4] Hussain F, Hussain R, Hassan SA, Hossain E. Machine learning in IoT security: Current solutions and future challenges. IEEE Commun Surv Tutor. 2020;22(3):1686–1721. https://doi.org/10.1109/comst.2020.2986444

[5] Nauman A, Qadri YA, Amjad M, Zikria YB, Afzal MK, Kim SW. Multimedia Internet of Things: A comprehensive survey. IEEE Access. 2020;8:8202–8250. https://doi.org/10.1109/access.2020.2964280

[6] Park S, Kim Y. A metaverse: Taxonomy, components, applications, and open challenges. IEEE Access. 2022;10:4209–4251. https://doi.org/10.1109/access.2021.3140175

[7] Rasheed A, San O, Kvamsdal T. Digital Twin: Values, challenges and enablers from a modeling perspective. IEEE Access. 2020;8:21980–22012. https://doi.org/10.1109/access.2020.2970143

[8] Shaik T, Tao X, Higgins N, Li L, Gururajan R, Zhou X, et al. Remote patient monitoring using artificial intelligence: Current state, applications, and challenges. Wiley Interdiscip Rev Data Min Knowl Discov. 2023;13(2):e1485. https://doi.org/10.1002/widm.1485

[9] Wang F, Zhang M, Wang X, Ma X, Liu J. Deep learning for edge computing applications: A state-of-the-art survey. IEEE Access. 2020;8:58322–58336. https://doi.org/10.1109/access.2020.2982411

[10] Zhang C, Patras P, Haddadi H. Deep learning in mobile and wireless networking: A survey. IEEE Commun Surv Tutor. 2019;21(3):2224–2287. https://doi.org/10.1109/comst.2019.2904897