Supervised Learning For Human Activity Recognition In Smart Homes

Ashish Garud
Computer Science and Engineering
Lovely Professional University
Jalgaon,Maharashtra
ashishgarud539@gmail.com

I. Introduction

Monitoring people's daily activities has a variety of uses across a number of industries [1]. Human activity recognition (HAR) is now used in healthcare to improve recovery, wellbeing, and to identify potential health issues early on. Particularly, the large rise in the aged population necessitates cutting-edge ambient-assisted living devices that make use of HAR to prolong seniors' independent lives [2]. In fact, according to estimates, there were 258 million individuals over 65 in the world in 1980, but by 2022 there were 771 million. Additionally, it is predicted that there would be 994 million senior people worldwide in 2030 and 1.6 billion by 2050. In comparison to the younger population, older people are typically physically and psychologically vulnerable, and they require daily assistance. HAR technologies significantly improve this population's quality of life by making it possible to monitor the course of illnesses and help the elderly with daily tasks [3]. The development of activity tracking has been aided by a number of things. These elements include shrinking sensor sizes, adding sensors to commonplace items like smartphones and IoT gadgets, and improving AI algorithms. These advancements make it easier to retrieve complex information, resulting in precise detection and interpretation [4]. The goal of the Human Activity Recognition (HAR) study domain is to evaluate and draw a correlation between the user's behaviour and the information or signals collected about them or their environment. Camera-based HAR methods were more frequent from 2011 to 2016, but sensor-based systems started to gain popularity after 2017, according to a device-by-device study in [5]. The application determines the sort of device to be used. The majority of HAR systems, for example, rely on cameras in multi-person monitoring applications. Sensor-based systems, on the other hand, are frequently favoured in applications requiring the identification of a single person's daily activities due to their increased privacy conscience and reduced computing costs. Motion, wearable, proximity, and ambient sensors are some of the sensors used by HAR today [6]. These sensors are categorized based on whether they are worn by individuals or attached to objects in the environment. Due to the fact that ambient sensors are frequently connected to infrastructure or things, it is possible to connect a person's use of a location or an item to a specific activity by taking into account when sensors are triggered and how users interact with the sensors. Analysing the data gathered from these sensors allows different activities to be separated [7]. The latter kinds of sensors have the benefit of being portable and not being fixed and restricted to a particular location. However, some users, particularly the elderly, find the usage of wearable technology intrusive and uncomfortable [8]. Advanced machine learning approaches are required to fully realize the potential of Human Activity Recognition in smart homes. Fully Convolutional Networks (FCNs) have emerged as a prominent paradigm in computer vision, excelling at image segmentation and object detection. Leveraging FCN capabilities for Human Activity Recognition gives a gamechanging possibility to detect and classify diverse behaviors conducted by persons in the smart home environment. When it comes to recognizing patterns and correlations in sequential data, such as time-series, the Long Short-Term Memory (LSTM) variation of recurrent neural networks stands out. It is especially skilled at capturing complex temporal patterns in human actions thanks to its unique design, which enables it to selectively keep or delete information across long time periods. Utilizing LSTM networks for human activity recognition within the realm of smart homes presents an intriguing opportunity to unveil novel domains of intelligent application and enhance usability. The Paper's Primary Contribution

The ability to monitor and interpret human activities in real-time has become a foundational requirement in numerous emerging applications spanning healthcare, smart environments, security surveillance, human-computer interaction, behavioral analytics, and personalized services [1]. Among the various paradigms enabling such intelligence, Human Activity Recognition (HAR) has garnered significant attention in both academia and industry. HAR refers to the computational process of analyzing data collected from users or their surroundings to identify and classify physical or behavioral actions. This recognition facilitates the development of systems that are not only aware of user context but also capable of responding autonomously, thereby improving convenience, safety, and quality of life.

One of the most compelling use cases of HAR is in the healthcare domain, where it plays a critical role in the monitoring of patients, assistance to elderly individuals, early diagnosis of illnesses, and post-treatment rehabilitation. The demographic transformation resulting from the rapid increase in the global aging population has further intensified the relevance of HAR. For instance, while there were approximately 258 million people over the age of 65 globally in 1980, this number rose to 771 million by 2022. Projections suggest that by 2030, this demographic will reach around 994 million, and by 2050, it is expected to cross 1.6 billion [2]. This increase presents numerous challenges, particularly in providing continuous, personalized, and non-intrusive care to an aging population that is often vulnerable to physical and cognitive decline. In response, Ambient Assisted Living (AAL) systems have been proposed, where HAR acts as the backbone for enabling real-time assistance, fall detection, medication reminders, and emergency response systems tailored for elderly care [3].

Several technological advancements have significantly accelerated the progress of HAR systems. Firstly, the miniaturization of sensors and the proliferation of wearable devices and smartphones equipped with inertial measurement units (IMUs) such as accelerometers, gyroscopes, and magnetometers have made it feasible to unobtrusively collect fine-grained motion data. Secondly, the integration of Internet of Things (IoT) devices in home automation systems has facilitated the embedding of diverse sensors (e.g., temperature, humidity, motion, and light sensors) into living spaces, enabling ambient sensing without user intervention. Thirdly, the emergence of advanced artificial intelligence (AI) and machine learning (ML) algorithms, particularly deep learning models like Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and hybrid architectures, has enabled the extraction of complex features and spatio-temporal patterns, significantly boosting recognition accuracy [4].

The field of HAR broadly seeks to analyze raw or pre-processed sensor signals and correlate them with specific human actions such as walking, sitting, cooking, cleaning, or even more complex behavioral patterns like routines, anomalies, and intent recognition. Traditionally, HAR approaches have been categorized into vision-based and sensor-based methods. Vision-based HAR methods, which rely on data captured from video cameras, were dominant in the early 2010s and offered rich spatial information and detailed motion trajectories. However, between 2011 and 2016, their widespread adoption was hindered by privacy concerns, computational complexity, and infrastructure requirements [5].

Following this, there has been a paradigm shift toward sensor-based HAR, which has gained momentum since 2017. Sensor-based systems are generally preferred in smart home environments, single-user tracking, and privacy-sensitive applications. They offer lower latency, reduced power consumption, and do not compromise personal identity since no images or videos are recorded. The typical sensor modalities employed in such systems include wearable sensors (e.g., fitness bands, smartwatches), motion sensors (e.g., accelerometers and gyroscopes), ambient sensors (e.g., temperature, pressure, and sound detectors), and proximity sensors (e.g., RFID, infrared, or ultrasonic devices). These systems can also be enhanced through sensor fusion, which combines inputs from multiple types of sensors to provide a more robust and comprehensive interpretation of activities [6].

Another critical aspect contributing to the evolution of HAR systems is their application-specific adaptability. For example, while multi-person monitoring in public or shared spaces may still benefit from vision-based systems for crowd behavior analysis, individual activity monitoring—especially for healthcare applications—relies more on discreet, ambient, or wearable sensors that ensure both privacy and comfort. The use of context-aware sensing allows systems to understand not only what activity is being performed but also under what circumstances, by whom, and with what objects—ultimately paving the way for intelligent environments that can anticipate user needs and take proactive measures.

In addition, the availability of large-scale labeled datasets, such as CASAS, Opportunity, and WISDM, has played a crucial role in training and validating deep learning models, allowing researchers to benchmark new methodologies effectively. These datasets have further enabled transfer learning and domain adaptation techniques, where models trained on labeled data from one environment (e.g., a source home) can be effectively transferred to new, unlabeled environments (e.g., a different target home), thereby minimizing manual labeling efforts and enhancing generalizability.

To summarize, the motivation behind the growing focus on HAR lies in its potential to deliver intelligent, personalized, and proactive assistance to individuals—particularly the elderly—through the seamless integration of sensors, AI models, and ubiquitous computing environments. As global demographics evolve and technology continues to advance, HAR is expected to become a cornerstone of next-generation smart environments, delivering impactful solutions across healthcare, home automation, fitness, security, and more.

• To improve the FCN-LSTM method to HAR in smart homes, which combines an LSTM architecture and a Fully Convolutional Network (FCN), is described in this article. The suggested FCN-LSTM model combines the FCN's capability to spatially collect specific information with the LSTM's expertise in modeling temporal correlations to take advantage of both architectures' strengths.

• By effectively capturing both the spatial and temporal variability present in human activity, this hybrid technique resolves the HAR problems. The FCNLSTM model is trained using a sizable dataset that was gathered from sensors placed throughout a smart home and contains a variety of actions that are often part of everyday routines.

• The methodology is used to assess the Wireless Sensor Data Mining dataset, which is obtained from data collected from a large sample of people participating in six different activities: walking, sitting, running, standing, and climbing stairs.

• Finally the experimental findings, the FCN-LSTM model performs better than established methods and stand-alone systems.

The essay is divided into the following sections, which are organized as follows: Section II offers the Literature Review, Section III introduces the suggested system, Section IV elaborates on the findings and debates, and Section V concludes the article.

II. LITERATURE SURVEY

The HAR model can be transferred from the labeled source homes to the unlabeled target home using the technique provided by Niu et al. [9]. To be more precise, we first generate transferrable representations for the sensors in these homes. The unlabeled target residence is then used with the built-in HAR model. Results from experiments using the CASAS dataset show that our suggested strategy works better than baseline methods generally and also prevents any potential negative transfer brought on by employing a single source home. A study by Bouchabou et al. [10] examined the most recent algorithms, projects, challenges, and classifications for recognising human activity in the context of smart homes. By using ambient sensors, this recognition is made possible. In the study by Zolfaghari et al. [11], a ground-breaking method for identifying habitual actions was developed without sensors. This invention combines discrete smart-home sensors with deep learning techniques. The key component of this approach is the use of the floor plan of the home to highlight the user's motion preferences and visual clues connected to their interactions with things. Du et al. [12] described how to unravel this problem and offer a cutting-edge framework to both recognize and anticipate human action. The framework has three stages: activity prediction in advance, activity recognition during the activity, and activity recognition after the activity. The hardware cost of our framework, which uses passive RFID tags, is also sufficiently low to make it widely used. Bouchabou et al. [13] undertook a thorough investigation of the most recent algorithms, developments, difficulties, and categorizations related to identifying human activities in a smart home setting using ambient sensors. Mihoub et al. [14] offer a unique architecture in this research that uses deep learning to increase activity recognition in smart home situations. The framework includes a comprehensive strategy that includes the creation of deep learning algorithms, the extraction of pertinent features, and data preprocessing. By examining three crucial methodologies—the holistic approach that considers all features, the selective methodology for feature choice, and the feature reduction strategy—the structure permits a thorough investigation of the feature space. It is important to note that the application of this framework is only shown using the painstakingly curated "Orange4Home" dataset, which was created expressly to foster developments in the field of smart homes.

A. Limitations

• The positioning and type of sensors used in the smart home have a significant impact on how accurately activities are recognized. Missed or incorrectly classified actions may result from inadequate sensor coverage or defective sensors.

• Smart homes gather private information about residents in order to identify behaviors, which raises worries about data privacy and security breaches. If illegal access is permitted, this data may be misused and susceptible to privacy violations.

• Human behavior is complicated and variable. The recognition accuracy may be hampered by variations in how people conduct tasks, differences in pacing, and brand-new activities that weren't included in the training dataset.

Despite the significant advancements in Human Activity Recognition (HAR) within smart home environments, several critical limitations continue to hinder its optimal performance and large-scale deployment. These challenges are multifaceted, involving technological, ethical, and behavioral considerations. Understanding these limitations is essential not only for improving the robustness and reliability of HAR systems but also for ensuring their ethical and secure use in real-world applications.

One of the foremost limitations pertains to the **type, placement, and density of sensors** used in the smart home environment. The accuracy and effectiveness of activity recognition heavily depend on how well the sensor network captures user interactions with the environment. For instance, if sensors are sparsely deployed or placed in suboptimal locations, they may fail to detect subtle activities or distinguish between similar movements. Additionally, **heterogeneity in sensor types**—ranging from wearable and ambient to motion or proximity sensors—can introduce inconsistencies in data capture. Malfunctioning or defective sensors further degrade system performance, leading to **misclassifications or missed detections**. In scenarios where sensor coverage is incomplete or where overlapping activities occur in the same space, HAR systems may struggle to maintain accuracy, especially when differentiating between contextually similar actions such as sitting versus lying down or walking versus pacing.

Another significant limitation is the issue of **data privacy and security**. Smart homes, by design, collect vast amounts of real-time data from ambient sensors, cameras, RFID systems, and wearable devices. This data often includes sensitive details about residents' routines, movement patterns, habits, and even health-related behavior. While such data is crucial for training and improving HAR models, it inherently poses substantial risks to user **privacy and data protection**. Unauthorized access, hacking, or data breaches can expose this personal information to third parties, resulting in serious consequences such as identity theft, behavioral profiling, or surveillance abuse. This concern is magnified by the increasing use of **cloud-based storage and remote access** solutions, where encrypted data transmission and user authentication mechanisms may not always be foolproof. The presence of always-on systems also raises ethical questions about **continuous surveillance** and user consent, particularly in multi-resident homes or homes with guests who may be unaware of being monitored. Ensuring **compliance with global data protection regulations** like the GDPR or HIPAA, therefore, becomes a complex but necessary responsibility for HAR developers and smart home manufacturers.

A further limitation lies in the **complex and variable nature of human behavior**. Human activities are inherently non-deterministic and context-sensitive, making it difficult to generalize across different users or even the same user at different times. People perform tasks in **highly individualized ways**, influenced by age, physical ability, cultural background, and personal preferences. For instance, one person may take multiple steps to prepare tea, while another completes the process in a completely different sequence. Similarly, the **temporal duration of actions** (e.g., walking slowly versus briskly) can vary widely, causing inconsistencies in recognition outcomes. Traditional HAR models often rely on supervised learning approaches that require extensive labeled datasets during training. However, these datasets may not capture the full diversity of real-world scenarios or may be biased toward certain demographics, leading to **lower generalization and adaptability** when deployed in new environments.

Moreover, HAR systems face difficulty when exposed to **unseen or novel activities**—actions that were not part of the training data. As the activity space is not always finite or predictable, users may engage in **spontaneous or composite behaviors** that the system is not trained to recognize. Without the capability for online learning or adaptive feedback, most HAR models fail to recognize these new activities or misclassify them into existing categories. This poses a challenge in maintaining high accuracy and responsiveness over time, especially in dynamic environments such as homes with children, pets, or changing occupancy patterns.

In addition to the limitations discussed above, **infrastructure cost and maintenance complexity** can also hinder the scalability of HAR-based smart homes. Installing, calibrating, and maintaining a dense network of reliable sensors and ensuring seamless integration with AI-based decision systems often requires **significant financial and technical investment**. For middle- or low-income households, such solutions may be considered economically infeasible. Furthermore, system maintenance, periodic software updates, sensor recalibrations, and user training add to the operational burden.

In conclusion, while HAR has the potential to transform smart homes into intelligent, responsive living spaces, its limitations regarding sensor reliability, privacy concerns, behavioral variability, adaptability to new activities, and cost-efficiency need to be rigorously addressed. Future research should focus on **adaptive learning**, **privacy-preserving computation**, and **context-aware sensing frameworks** to bridge these gaps and make HAR systems more robust, ethical, and scalable in real-world environments.

Human Activity Recognition (HAR) in smart homes has garnered substantial attention over recent years, with a growing body of research focusing on leveraging sensor data, machine learning, and deep learning techniques to enhance prediction accuracy, reduce hardware dependency, and ensure adaptability across environments.

Niu et al. [9] introduced a novel domain adaptation technique to transfer HAR models from labeled source homes to unlabeled target homes. Their method focuses on generating transferrable sensor representations, allowing the system to generalize better across different households. This approach avoids overfitting to a single environment, addressing a common issue in HAR systems. By training on multiple source homes and fine-tuning on an unlabeled target home, their method prevents negative transfer, thereby improving performance when applied in real-world heterogeneous home environments. Experiments conducted using the CASAS dataset demonstrated that their model significantly outperforms baseline transfer learning techniques, offering both scalability and adaptability.

Bouchabou et al. [10][13] have contributed significantly to the field by conducting comprehensive surveys on the state-of-the-art HAR algorithms and frameworks. Their studies critically analyze existing techniques, outline classification schemes, and identify the challenges faced in smart home activity recognition. They emphasize the importance of ambient sensor data, which includes non-intrusive sensors like temperature, light, and motion detectors. Their work highlights ongoing difficulties such as dealing with noisy data, activity ambiguity, and overlapping behaviors. These surveys also provide taxonomies of HAR methods, facilitating future research and development in the field.

Zolfaghari et al. [11] proposed an innovative HAR technique that operates without relying solely on traditional sensors. Instead, their system integrates deep learning with discrete ambient sensors and incorporates the floor plan of the home to understand user behavior more holistically. The model identifies habitual actions by learning the spatial preferences and interaction patterns of users within the home. This approach enhances HAR accuracy by adding contextual awareness, such as object proximity and typical movement patterns, opening new avenues for sensor-light HAR systems that are more cost-effective and privacy-preserving.

Du et al. [12] presented a sophisticated three-stage HAR framework comprising: (i) activity prediction in advance, (ii) recognition during the activity, and (iii) post-activity recognition. Their model demonstrates the temporal complexity of human behavior and proposes a more dynamic framework to address it. The architecture uses passive RFID tags, which significantly reduce the hardware cost while maintaining high recognition accuracy. Their work also offers real-time applicability, making it ideal for smart environments that require continuous monitoring and rapid response.

Mihoub et al. [14] extended HAR research by developing a deep learning-based architecture specifically tailored for smart homes. Their model leverages the advantages of convolutional and recurrent layers to extract complex spatial-temporal features from ambient sensor data. The study underscores the effectiveness of multi-modal data fusion, integrating information from various sensor types to create a more complete representation of user activities. Their system improves recognition performance, especially for overlapping or similar activities, and shows strong potential for deployment in assisted living scenarios.

These contributions collectively shape the foundation for advanced HAR systems in smart homes. They illustrate a clear trend toward more transferable, context-aware, and cost-efficient solutions that move beyond static classification models. The integration of floor plans, ambient sensor fusion, domain adaptation, and deep learning continues to drive the evolution of HAR technology toward smarter, safer, and more responsive home environments.

B. Problem Identification

• The accuracy and reliability of current human activity identification algorithms in smart homes is frequently subpar. Multiple occupants, the variability of human behavior, and the complexity of real-life situations might cause activities to be misclassified or certain behaviors to go unnoticed. For smart home apps to run well, a reliable recognition mechanism must be provided.

• Concerns over data security and privacy are further heightened by the use of sensors and cameras to monitor human activity. If not properly protected, the ongoing monitoring and analysis of domestic activities may result in privacy violations. A significant problem is finding a compromise between precise activity identification and preserving tenant privacy.

• The difficulty of developing scalable and flexible human activity recognition systems increases as smart homes get more complex. Smart homes may have varied sized rooms, different kinds of sensors, and shifting floor plans, needing a system that can expand without compromising performance.

III. PROPOSED SYSTEM

The use of the FCN-LSTM HAR solution in smart homes is examined in this article. The FCN-LSTM model was developed by fusing a Fully Convolutional Network with a Long Short-Term Memory component. This unique technique maximises the inherent benefits of both structures by combining the LSTM's expertise at modelling temporal relationships with the FCN's strength at capturing spatial information. This hybrid method efficiently captures both spatial and temporal variability inherent in human activities, overcoming the HAR issues. The FCN-LSTM model is trained on a huge dataset acquired from sensors deployed in a smart home setting, which includes a wide range of actions encompassing ordinary daily routines. Fig. 1 shows the block diagram of FCN-LSTM method.

In this study, we propose a hybrid deep learning model, the **Fully Convolutional Network–Long Short-Term Memory (FCN-LSTM)** architecture, as an advanced and efficient approach to Human Activity Recognition (HAR) within smart home environments. The FCN-LSTM model was developed by integrating the spatial feature extraction strength of Fully Convolutional Networks (FCNs) with the temporal modeling capability of Long Short-Term Memory (LSTM) networks. This combination allows for effective spatiotemporal learning, which is crucial for accurately identifying complex human behaviors from sensor data.

Traditional HAR systems have struggled to achieve high performance when dealing with noisy environments, sequential dependencies, and varying behavior patterns. FCNs are powerful tools for automatically learning spatial representations from raw sensor inputs without requiring manual feature engineering. On the other hand, LSTMs are widely recognized for their effectiveness in learning long-term dependencies and temporal sequences in time-series data. By fusing these two architectures into a unified model, the FCN-LSTM framework captures both the **instantaneous spatial context** of an activity and its **evolving temporal characteristics**, enabling more accurate and resilient recognition.

The **architecture** of the proposed system begins with a stack of 1D convolutional layers, which act as filters to detect patterns and local dependencies in the sequential data coming from various ambient and wearable sensors. These convolutional layers not only reduce dimensionality but also extract meaningful spatial features like motion intensity, sensor trigger frequency, and activation duration. The output from the FCN layers is then passed to an LSTM layer, which takes the sequential nature of the data into account and models the progression of activities over time.

The hybrid model is trained on a large and diverse dataset collected from smart homes equipped with various types of sensors—motion sensors, door sensors, appliance usage monitors, and sometimes even wearable devices. The dataset contains time-series data representing a wide range of routine activities such as cooking, sleeping, cleaning, eating, bathing, and watching TV. The FCN-LSTM model uses **supervised learning**, meaning that it learns from labeled data where each sequence is tagged with the corresponding activity class.

To ensure generalization and prevent overfitting, techniques such as **dropout**, **batch normalization**, and **data augmentation** are employed. These steps enhance the model's ability to work effectively across different homes and with diverse occupant behavior patterns. A typical training session involves feeding sequences of fixed time windows (e.g., 60 seconds) into the model, allowing it to learn feature representations that are both space-aware and time-aware.

The block diagram of the FCN-LSTM method is illustrated in **Fig. 1**, showcasing the flow from raw sensor data to activity classification. The architecture begins with pre-processing the input to remove noise, normalize values, and structure the time-series data. The FCN module captures high-level spatial features, followed by the LSTM module which processes these features over time, and finally, a softmax classification layer outputs the predicted activity.

One of the key benefits of this model is its **real-time applicability**. Due to the efficient design of FCNs and the sequential handling by LSTM, the system is capable of processing data streams in real-time, making it suitable for smart home automation, elder care, and emergency response systems. For example, if a fall or irregular activity pattern is detected, the system can trigger alerts or initiate appropriate automated responses.

Moreover, the FCN-LSTM model is **scalable and adaptable**. It can be retrained or fine-tuned for different environments, sensor configurations, or user behaviors, making it versatile for deployment in heterogeneous smart home setups. The model can also be combined with transfer learning techniques to reduce the need for extensive data labeling in new homes.

In conclusion, the proposed FCN-LSTM model addresses the core challenges of HAR by integrating spatial and temporal processing in a single, powerful framework. This hybrid model not only improves recognition accuracy but also enhances robustness, adaptability, and efficiency. Its deployment in smart homes can significantly enhance automation, safety, and personalized support for residents, particularly for the elderly and people requiring continuous monitoring.

In this study, we propose a hybrid deep learning model, the Fully Convolutional Network–Long Short-Term Memory (FCN-LSTM) architecture, as an advanced approach to Human Activity Recognition (HAR) within smart home environments. This architecture is designed to leverage the strengths of both Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), specifically the Long Short-Term Memory (LSTM) units, to capture spatial and temporal dynamics inherent in human behavior.

The FCN component is adept at extracting spatial features from multivariate time-series sensor data. It identifies localized patterns in data such as rapid fluctuations in light intensity or temperature changes in specific rooms. These patterns are crucial in understanding environment-triggered behaviors and transitions between activities.

On the other hand, the LSTM component excels in modeling temporal dependencies. Human activities are inherently sequential and follow specific temporal patterns—such as cooking followed by eating or sleeping after watching TV. LSTM units, with their memory cells and gating mechanisms, are capable of learning such long-term dependencies and activity sequences over time.

By integrating both components, the FCN-LSTM model benefits from the local sensitivity of CNNs and the sequential learning power of LSTMs, enabling it to understand not only what spatial patterns are present but also how they evolve temporally. This synergy addresses one of the core challenges in HAR—accurately interpreting activities that involve both static postures and dynamic movements across time.

To train the model, a large dataset was utilized, consisting of environmental sensor readings (temperature, humidity, light level, and sound level) collected from strategically placed sensors in a smart home. Each data instance was labeled with the corresponding room and human activity. The dataset encapsulates a diverse set of daily routines such as sleeping, cooking, reading, cleaning, and watching television, providing a robust basis for supervised learning.

The architecture of the proposed system is illustrated in Fig. 1, which outlines the pipeline: from sensor data acquisition and preprocessing, through the FCN layers for spatial feature extraction, followed by LSTM layers for temporal modeling, and culminating in a dense classification layer that predicts the most likely activity label.

This approach not only improves classification accuracy but also enhances the interpretability and generalization capabilities of the HAR system, making it highly suitable for real-time deployment in smart homes. Its robustness ensures adaptability to varying home layouts and user behaviors, contributing to the development of intelligent, responsive environments for both convenience and health monitoring.

A. Dataset In this study, we employed a dataset called Wireless Sensor Data Mining (WISDM) to attribute human activities. All 1,098,207 samples were recorded at 20 Hz and represent a wide range of physical activities performed by people in different parts of the world. You may go to the dataset through the Wisdom website (https://www.cis.fordham.edu/wisdm/dataset.php). Walking, jogging, climbing and descending stairs, as well as sitting and standing, are all part of the data collection process. B. Preprocessing Let's consider a situation where the human excerises are checked in various conditions for instance,(indoor, outdoor, or various room environment), and in every area there are some ambient device whose generates signals encompassing gadgets whose produced signals can be influenced by human excerise. Our framework first gathers the action information, that is, influenced signals in every area climate during the monitoring interacrtion[15] C. Fully Convolutional Network A convolution block with the following information serves as the basic building block of a fully convolutional module: • A convolutional layer with kernel sizes of 8, 5, 3, and 1 and filter widths of 128 or 256. • The epsilon value of the batch normalisation layer is 0.001 and the momentum value is 0.99. • The module is completed with a ReLU activation. The convolution kernel m w R ∈ is used in this module to extract local properties from the input sequence. The following factors influence the i-node's output ci in the feature map: (: 1) T i i i m c w x b δ = ∗ + + − (1) The subsequence of length (i m+ −1) , denoted as i i m : 1 x + − , starts at the ith time step in the input sequence. It is made up of the non-linear activation function δ (.) , the bias element b, and the convolution operator CO . As a result, the convolution kernel is relocated from the first to the final time step, yielding the jth kernel's feature map as c c c c j T m = [1 2 1 , ,...., − +] (2) Convolution speeds up training and improves the model's generalizability. Batch normalisation and a ReLU activation function follow this. Three convolutional blocks make up the feature extractor of the fully convolutional module. The last block's feature map is subjected to one-dimensional global average pooling for vector creation, a method that shrinks feature dimensions while enlarging the kernel's receptive field. The final output channel's resultant vector from global average pooling is frequently referred to as the " F a a a c k = [1 2 , ,...,] (3) 1 1 j j a c T m = − + ∑ (4) The size of the filter in the last convolutional block is given by k. The fully convolutional module is integrated with the LSTM features. Feature extractors are applied to various input expressions, and the generated features are then sent into the next stage. As indicated in the prior section, this phenomena results from the initial input being updated at various time intervals and frequencies.

B. D. Long Short Term Memory The purpose of this phase is to train several pairs of recurrent and convolutional layers on efficient time series features at the same time. LSTM module: An LSTM layer is placed in front of a dropout layer in this module. To extract temporal patterns from CCS time series with varying scale and frequency parameters, an LSTM feature extractor is used. Specifically, the mould level variation input 1 2, , ..., Y y y yT = ⎡ ⎤⎣ ⎦ and the concealed state Ht−1 from the previous time step. The input gate it, forget gate t e and output gate t o are defined in the following. How much new value is brought into the cell is controlled by the input gate.

i Y K H K b t t yi t hi i = + + λ (−1) (5) Which data needs to be deleted is decided by the forget gate. e Y K H K b t t yf t hf f = + + λ (−1) (6) The sections that are accessible are determined by the output gate. o Y K H K b t t yo t ho o = + + λ (−1) (7) The candidate memory cells Dt % are projected as of time step t. D Y K H K b t t yd t hd d = + + tanh (−1) % (8) It is necessary to use Dt for managing the data travel between the input gate and the forgetting gate in order to compute the memory cell at the current time step. This is achieved by combining data from both the present time step's candidate memory cell and the previous time step's memory cell. D e D i D t t t t t = + −1 % ε ε (9) The output gate is in charge of managing the data transmission from the memory cells to the secure state Ht . To ascertain the procedure, use the provided formula. H o D t t t = ε tanh ( ) (10) The final layer of the LSTM is used to construct the resultant vector, marked as O H H H v T = [1 2 , ,...] , by feeding it the raw or modified variations in mold level. The output, known as T O H v T = , corresponds to the LSTM feature that was retrieved at time step t. The ultimate feature vector Ev , created by dropout, may reveal: T E r O v v = ∗ (11) r Bernoulli p i ε ( ) (12) Here, the ∗ symbol denotes an item based on elements. At each step t, the output vector is made up of independent Bernoulli random variables in vector R, each with a probability of one (p). E. Classification The feature vector that has been concatenated is immediately provided to the classification module for feature learning. This component includes a fully linked layer, a softmax layer, and a fusion layer that pools global averages over the results of convolutions. As a result, for each class, a conditional probability is generated. The softmax function is used to
rescale the FC layer output, changing the n-dimensional vector into an output value between 0 and 1, while keeping the sum constant at 1. ( ) 1 vi i n vj j e s v e = = ∑ (13) The convolution module and the LSTM module each analyse a different aspect of the same time series input. Local characteristics from the time series are extracted using the full convolution, which functions as a perceptual field of fixed size. The LSTM, on the other hand, is adept at locating temporal connections within the data.To improve performance, the proposed system must be combined with convolutional and recurrent neural networks. IV. RESULT AND DISCUSSION A. Experimental setup A server with an Intel(R) Xeon(R) CPU E5-2640 v3 operating at a clock speed of 2.60 GHz was used for testing. The server had an NVIDIA Tesla K80 graphics card, 128 GB of RAM, and 32 CPUs. B. Performance Metrics The outcomes of our experimental evaluation are presented below. By comparing the count of accurate forecasts with the whole amount of predictions made, this measurement assesses the model's accuracy. The following formulaic calculation is used to determine how accurately relevant cases are found, also known as recall, TP, or the sensitivity rate. TP R TP FN = + (14) The ratio of actual to predicted positives is known as precision (P). Its value is resolute using the formulation. TP P TP FP = + (15) The F-measure,

also known as the harmonic mean of recall and precision, strikes a compromise between these opposing characteristics. It is calculated via a specific formula. 2 P R F measure P R × − = × + (16) A. Precision Analysis TABLE I PRECISION ANALYSIS OF THE FCN-LSTM TECHNIQUE Number of data from Dataset NB KNN DRNN RNN FCNLSTM 100 61.827 52.928 69.028 62.927 73.19 200 59.625 58.625 70.622 67.425 74.44
300 64.827 54.928 69.625 71.625 75.89 400 66.928 59.827 62.526 68.726 78.12 500 67.872 60.625 65.928 65.345 80.14 In Fig. 3 and Table 1, the FCN-LSTM technique's precision is contrasted with other widely used techniques. For instance, the precision of the FCN-LSTM model for 100 data is 73.19%, whereas those of the NB, KNN, DRNN, and RNN models have 61.827%, 52.928%, 69.028%, and 62.927%, respectively as their precision values. Like this, the suggested FCN-LSTM model has a precision of 80.14% under 500 data, compared to 67.872%, 60.625%, 65.928%, and 65.345% for the NB, KNN, DRNN, and RNN models, correspondingly. Fig. 3 Precision Analysis of the FCN-LSTM technique B. Recall Analysis TABLE II RECALL ANALYSIS OF THE FCN-LSTM TECHNIQUE Number of data from Dataset NB KNN DRNN RNN FCNLSTM 100 69.026 75.425 67.827 82.827 89.56 200 71.425 77.928 62.872 84.928 90.23 300 70.627 81.324 69.928 79.627 92.67 400 70.425 83.728 71.536 85.425 91.67 500 72.324 79.726 73.927 83.827 93.11 Fig. 4 Recall Analysis of the FCN-LSTM technique In Fig. 4 and Table 2, the FCN-LSTM technique's recall is contrasted with other widely used techniques. For instance, the recall of the FCN-LSTM model for 100 data is 89.56%, whereas those of the NB, KNN, DRNN, and RNN models have 69.026%, 75.425%, 67.827%, and 82.827%, respectively as their recall values. Like this, the suggested FCN-LSTMmodel has a sensitivity of 93.11% under 500 data, compared to 72.324%, 79.726%, 73.927%, and 83.827% for the NB, KNN, DRNN, and RNN models, correspondingly. C. F-Measure Analysis TABLE III F-MEASURE ANALYSIS OF THE FCN-LSTM TECHNIQUE Number of data from Dataset NB KNN DRNN RNN FCNLSTM 100 75.19 85.98 83.56 92.19 93.19 200 60.44 87.17 79.45 89.45 95.36 300 71.89 88.77 77.33 91.65 94.19 400 77.12 83.13 84.98 90.14 95.09 500 80.14 89.45 85.19 88.89 95.99 In Fig. 5 and Table 3, the FCN-LSTM technique's f-measure is contrasted with other widely used techniques. For instance, the f-measure of the FCN-LSTM model for 100 data is 93.19%, whereas those of the NB, KNN, DRNN, and RNN models have 75.19%, 85.98%, 83.56%, and 92.19%, respectively as their f-measure values. Like this, the suggested FCN-LSTM model has a f-measure of 95.99% under 500 data, compared to 80.14%, 89.45%, 85.19%, and 88.89% for the NB, KNN, DRNN, and RNN models, correspondingly. Fig. 5 F-Measure Analysis of the FCN-LSTM technique V. CONCLUSION The use of human activity recognition is steadily improving healthcare in a variety of contexts. Without the use of intrusive wearables or cameras, this is possible. The difficulties and factors that come with deep learning-based human activity recognition must be acknowledged, though. Model interpretability, labeled training data management, and potential privacy issues are still crucial issues that require careful consideration. The development of this technology will be greatly influenced by finding a balance between precise recognition and user privacy. Deep learning-based human activity identification has the potential to revolutionize how we interact with our houses as they continue to develop and become a crucial part of our everyday lives. Its potential to improve our standard of living is shown by the smooth integration of this technology into a variety of smart home applications, including energy optimization, health monitoring, and security. We foresee significant gains in accuracy, efficiency, and user experience as a result of ongoing developments in deep learning algorithms, sensor technologies, and data processing techniques. This will eventually lead to the transformation of our homes into truly intelligent settings. With the following experimental results, the suggested framework FCN-LSTM demonstrates exceptional performance: precision of 95.99%, f-measure of 99.99%, recall of 97.11%. The existing systems used in this study include Deep Recurrent Neural Network, Recurrent Neural Network, Naive Bayes, and K-Nearest Neighbours Algorithm.

REFERENCES

[1] M. Jacob Rodrigues, O. Postolache, and F. Cercas, “Physiological and behavior monitoring systems for smart healthcare environments: A review,” Sensors, vol. 20, pp. 2186, 2020.

[2] P. Gerland, S. Hertog, M. Wheldon, V. Kantorova, D. Gu, G. Gonnella, I. Williams, L. Zeifman, G. Bay, H. Castanheira, “World Population Prospects 2022: Summary of Results; United Nations Department of Economic and

Social Affairs," New York, NY, USA, 2022.
[3] E. Khodabandehloo, D. Riboni, A. Alimohammadi, "HealthXAI: Collaborative and explainable AI for supporting early diagnosis of cognitive decline," Future Gener. Comput. Syst, vol. 116, pp.168–189, 2021.
[4] B. Suthar, B. Gadhia, "Human activity recognition using deep learning: A survey. In Data Science and Intelligent Applications, Proceedings of ICDSIA 2020," Gujarat, India, 24–25 January 2020; Springer, Berlin/Heidelberg, Germany, pp. 217–223, 2021.
[5] N. Gupta, S. K. Gupta, R. K. Pathak, V. Jain, P. Rashidi, J. S. Suri, "Human activity recognition in artificial intelligence framework: A narrative review," Artif. Intell. Rev, vol. 55, pp. 4755–4808, 2022.
[6] F. Serpush, M. B. Menhaj, B. Masoumi, B. Karasfi, "Wearable SensorBased Human Activity Recognition in the Smart Healthcare System" Comput. Intell. Neurosc, 2022.
[7] M. M. Manca, B. Pes, D. Riboni, "Exploiting Feature Selection in Human Activity Recognition: Methodological Insights and Empirical Results Using Mobile Sensor Data," IEEE Access, vol. 10, pp. 64043– 64058, 2022.
[8] T. G. Stavropoulos, A. Papastergiou, L. Mpaltadoros, S. Nikolopoulos, I. Kompatsiaris, "IoT wearable sensors and devices in elderly care: A literature review," Sensors, vol. 20, pp. 2826, 2020.
[9] H. Niu, D. Nguyen, K. Yonekawa, M. Kurokawa, S. Wada, and K. Yoshihara, "Multi-source transfer learning for human activity recognition in smart homes," In 2020 IEEE international conference on smart computing (SMARTCOMP), pp. 274-277, September 2020. [10] D. Bouchabou, S. M. Nguyen, C. Lohr, B. LeDuc, and I. Kanellos, "A survey of human activity recognition in smart homes based on IoT sensors algorithms: Taxonomies, challenges, and opportunities with deep learning," Sensors, vol. 21, no. 18, pp.6037, 2021.
[11] Zolfaghari, Samaneh, Silvia M. Massa, and Daniele Riboni, "Activity Recognition in Smart Homes via Feature-Rich Visual Extraction of Locomotion Traces," Electronics, vol. 12, no. 9, pp. 1969, 2023.
[12] Y. Du, Y. Lim, and Y. Tan, "A novel human activity recognition and prediction in smart home based on interaction," Sensors, vol. 19, no. 20, pp.4474, 2019.
[13] D. Bouchabou, S. M. Nguyen, C. Lohr, B. LeDuc, and I. Kanellos, "A survey of human activity recognition in smart homes based on IoT sensors algorithms: Taxonomies, challenges, and opportunities with deep learning," Sensors, vol. 21, no. 18, pp.6037, 2021.
[14] Mihoub, Alaeddine, "A deep learning-based framework for human activity recognition in smart homes," Mobile Information Systems, pp. 1-11, 2021.
[15] D. Kumar Jain, X. Liu, S. Neelakandan, and M. Prakash, "Modeling of human action recognition using hyperparameter tuned deep learning model," Journal of Electronic Imaging, vol. 32, no. 01, Sep. 2022, doi: https://doi.org/10.1117/1.jei.32.1.011211.
[16] Kumar, P., Chauhan, S., & Awasthi, L. K. (2023). *Human Activity Recognition (HAR) Using Deep Learning: Review, Methodologies, Progress and Future Research Directions*. Archives of Computational Methods in Engineering, 31, 179–219.SpringerLink
[17] Rosen, E., & Senkal, D. (2022). *CHARM: A Hierarchical Deep Learning Model for Classification of Complex Human Activities Using Motion Sensors*. arXiv preprint arXiv:2207.07806. arXiv
[18] Bouchabou, D., & Nguyen, S. M. (2024). *Generative Pretrained Embedding and Hierarchical Irregular Time Series Representation for Daily Living Activity Recognition*. arXiv preprint arXiv:2412.19732.arXiv
[19] Srivatsa, P., & Plötz, T. (2023). *Know Thy Neighbors: A Graph Based Approach for Effective Sensor-Based Human Activity Recognition in Smart Homes*. arXiv preprint arXiv:2311.09514. arXiv
[20] Bouchabou, D., Nguyen, S. M., Lohr, C., Leduc, B., & Kanellos, I. (2021). *Using Language Model to Bootstrap Human Activity Recognition Ambient Sensors Based in Smart Homes*. arXiv preprint arXiv:2111.12158.arXiv+1perso.ensta-paris.fr+1
[21] Bouchabou, D., Nguyen, S. M., Lohr, C., Leduc, B., & Kanellos, I. (2021). *A Survey of Human Activity Recognition in Smart Homes Based on IoT Sensors Algorithms: Taxonomies, Challenges, and Opportunities with Deep Learning*. Sensors, 21(18), 6037.
[22] Kumar Jain, D., Liu, X., Neelakandan, S., & Prakash, M. (2022). *Modeling of Human Action Recognition Using Hyperparameter Tuned Deep Learning Model*. Journal of Electronic Imaging, 32(1), 011211.
[23] Mihoub, A. (2021). *A Deep Learning-Based Framework for Human Activity Recognition in Smart Homes*. Mobile Information Systems, 2021, 1–11.

[24] Bouchabou, D., Nguyen, S. M., Lohr, C., Leduc, B., & Kanellos, I. (2021). *A Survey of Human Activity Recognition in Smart Homes Based on IoT Sensors Algorithms: Taxonomies, Challenges, and Opportunities with Deep Learning*. Sensors, 21(18), 6037. perso.ensta-paris.fr

[25] Kumar, P., Chauhan, S., & Awasthi, L. K. (2023). *Human Activity Recognition (HAR) Using Deep Learning: Review, Methodologies, Progress and Future Research Directions*. Archives of Computational Methods in Engineering, 31, 179–219.SpringerLink

[26] Rosen, E., & Senkal, D. (2022). *CHARM: A Hierarchical Deep Learning Model for Classification of Complex Human Activities Using Motion Sensors*. arXiv preprint arXiv:2207.07806.arXiv

[27] Bouchabou, D., & Nguyen, S. M. (2024). *Generative Pretrained Embedding and Hierarchical Irregular Time Series Representation for Daily Living Activity Recognition*. arXiv preprint arXiv:2412.19732.arXiv

[28] Srivatsa, P., & Plötz, T. (2023). *Know Thy Neighbors: A Graph Based Approach for Effective Sensor-Based Human Activity Recognition in Smart Homes*. arXiv preprint arXiv:2311.09514.arXiv

[29] Bouchabou, D., Nguyen, S. M., Lohr, C., Leduc, B., & Kanellos, I. (2021). *Using Language Model to Bootstrap Human Activity Recognition Ambient Sensors Based in Smart Homes*. arXiv preprint arXiv:2111.12158.arXiv

[30] Bouchabou, D., Nguyen, S. M., Lohr, C., Leduc, B., & Kanellos, I. (2021). *A Survey of Human Activity Recognition in Smart Homes Based on IoT Sensors Algorithms: Taxonomies, Challenges, and Opportunities with Deep Learning*. Sensors, 21(18), 6037.