People’s Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
Mohamed El Bachir El Ibrahimi University of Borj Bou Arréridj
Faculty of Mathematics and Computer Science
Department of Computer Science

THESIS
In order to obtain the Doctorate degree in LMD (3rd cycle)
Branch: Computer Science
Option : Ingénierie de l’informatique décisionnel

THEME

Cognitive Psychology in Service of the Machine: Towards the Study of the Human Cognition

By: AGTI Nadia

Publicly defended on: 24/09/2025

In front of the jury composed of:

Pr. Behloul Ali	University of Batna2	**President**
Dr. Sabri Lyazid	University of BBA	**Supervisor**
Pr. Kazar Okba	University of Kalba, UAE.	**Co-Supervisor**
Pr.Bouziane Abderraouf	University of BBA	**Examiner**
Pr.Seghir Rachid	University of Batna2	**Examiner**
Pr. Nouioua Farid	University of BBA	**Examiner**
Pr.Akhrouf Samir	University of M’sila	**Examiner**

2024/2025

Abstract

The cognitive space plays an important role in structuring an action and the possible actions to be executed; therefore, actions are not only governed by a dynamic perception of knowledge but also by previously acquired knowledge. The neurophysiologist Alain Berthoz sees the brain as “a simulator of action and an emulator of reality.” He believes that it is through action—and not language—that we construct our perception of the world. The goal here is to integrate knowledge about actions/events from memory or past experiences and combine it with information perceived from the environment. The complexity of the problem increases when we take into account the unpredictable nature of human behavior.

Today, technological breakthroughs are attempting to enable direct communication between machines and the human brain, with the aim of performing actions through thought. The question now is whether it is possible to understand the physical or mental state that led a person to: 1) undertake actions they should have avoided, or conversely, 2) be encouraged to carry out actions they had previously abandoned.

We are interested in these theories in order to reproduce a person’s psychological state, with the aim of understanding and explaining the reasons behind the two points above. This is done by relying on semantic interdependencies and the properties of actions carried out in the past or present.

Keywords: Artificial Intelligence, Semantics, Spatio-temporal Reasoning, Behavior Recognition, Cognitive Psychology Theory.

Résumé

L'espace cognitif joue un rôle important pour structurer une action et les possibilités d'action à exécuter ; de ce fait, les actions n'obéissent pas uniquement à une perception dynamique des connaissances mais plutôt aussi à des connaissances acquises aux préalables. Le neurophysiologiste Alain Berthoz, voit le cerveau comme un simulateur d'action et un émulateur de la réalité. Il considère qu'à partir de l'action, et non du langage, que nous construisons notre perception du monde. Il s'agit ici d'intégrer des connaissances sur des actions/évènements à partir de la mémoire ou du vécu et les intégrées avec les informations perçues de l'environnement. La complexité du problème s'accroit lors de la prise en compte le caractère imprévisible de l'humain. Aujourd'hui des prouesses technologiques tentent de faire communiquer une machine directement avec le cerveau humain en vue de réaliser des actions via la pensée. La question maintenant est de savoir si on pourrait comprendre l'état physique ou mental ayant conduit l'humain à 1) entreprendre des actions qu'il devrait éviter, ou au contraire 2) l'inciter à exécuter des actions qu'il a abandonnées. Nous nous intéressons à ces théories pour tenter de reproduire l'état psychologique d'un humain en vue de comprendre et d'expliquer la(les) raison(s) expliciter par les deux points ci-dessus en s'appuyant sur des interdépendances sémantiques et des propriétés des actions produites dans le passé/présent.

Mots-clés : Intelligence artificielle, sémantique, raisonnement spatio-temporel, reconnaissance de comportement, théorie de la psychologie cognitive.

الملخص

يلعب الفضاء المعرفي دورًا مهمًا في هيكلة الأفعال وإمكانيات الفعل الممكن تنفيذها؛ فالأفعال لا تخضع فقط للإدراك الديناميكي للمعرفة، بل ترتكز أيضًا على معارف مكتسبة مسبقًا. يرى عالم الأعصاب ألان بيرتوز الين ـرتهز أن الدماغ هو محاكي للفعل ومُحاكي للواقع، ويعتبر أن الإدراك الذي نبنيه عن العالم ينبع من الفعل وليس من اللغة. في هذا السياق، يتم دمج المعارف المتعلقة بالأفعال أو الأحداث المستمدة من الذاكرة أو التجربة الحياتية مع المعلومات التي يتم إدراكها من البيئة المحيطة. وتزداد تعقيدات هذه المسألة عند أخذ الطابع غير المتوقع للسلوك البشري في الحسبان. في الوقت الراهن، تحاول الابتكارات التكنولوجية الربط بين الآلة والدماغ البشري مباشرة بهدف تنفيذ الأفعال عبر التفكير فقط. والسؤال المطروح الآن هو: هل يمكننا فهم الحالة الجسدية أو النفسية التي دفعت الإنسان إلى: ١) القيام بأفعال كان من المفترض تجنبها، أو على العكس، ٢) الإحجام عن تنفيذ أفعال كان يجب القيام بها؟ تهتم هذه الدراسة بهذه النظريات في محاولة لمحاكاة الحالة النفسية للإنسان من أجل فهم وتفسير الأسباب الكامنة وراء السيناريوهين المذكورين، بالاعتماد على الترابطات الدلالية وخصائص الأفعال المنجزة في الماضي أو الحاضر.

الكلمات المفتاحية

الذكاء الاصطناعي، الدلالات، الاستدلال الزماني المكاني، التعرف على السلوك، نظرية علم النفس المعرفي.

Acknowledgements

This journey would not have been possible without the grace and mercy of God. In moments of doubt and exhaustion, His presence gave me strength, clarity, and the courage to continue. For that, I am deeply grateful.

I would like to express my heartfelt thanks to my supervisor, Dr. Lyazid SABRI. His support has been a constant light throughout this work. With patience, generosity, and a deep belief in the value of this research, he guided me with care and wisdom. I am truly thankful for the trust he placed in me, and for the many discussions that pushed me to think deeper and work harder.

My sincere gratitude also goes to my co-supervisor, Professor Okba KAZAR, for his insightful advice and unwavering support. His thoughtful input helped shape this thesis in meaningful and lasting ways. Working under his guidance has been an enriching experience.

I am especially thankful to Dr. Soraya HAMIDA, whose encouragement, kindness, and support have meant more than words can say. She was always there when I needed reassurance or simply someone to believe in me.

To the distinguished members of my jury — Prof. Ali BEHLOUL, Prof. Samir AKROUF, Prof. Farid NOUIOUA, Prof. Rachid SEGHIR, and Prof. Abderraouf BOUZIANE — thank you for taking the time to read and evaluate my work. Your feedback, expertise, and valuable insights are deeply appreciated.

Above all, I owe everything to my mother, to my family, whose unconditional love, prayers, and encouragement have been my anchor through every step. Their faith in me has given me the strength to keep going. To all my loved ones who stood by me with patience and care. Thank you from the bottom of my heart.

Contents

List of Figures

List of Tables

This work is dedicated to:
My mother, whose unwavering support and love have guided me to this day.
My late father, who taught me to honor knowledge and cherish the pursuit of science.
My husband Amar,
My brothers: Rachid, Mourad, and Yahia,
My dear children: Ahmed Yassine and Maria,
My entire family, my in-laws, and all my cherished friends.

Chapter 1

General Introduction

1.1 Research Motivation and Problem Statement

Over the past decade, intelligent technologies have become increasingly embedded in our daily lives. From smart homes to wearable health monitors, these systems have evolved from passively collecting data to actively interacting with, assisting, and adapting to users. A key element driving this transformation is Human Activity Recognition (HAR), which focuses on detecting patterns in human behavior through sensor data.

Current HAR systems are generally effective at recognizing what a person is doing—such as walking, cooking, or sleeping. However, they often stop there. They rarely address the more complex and crucial question of why a person engages in a certain activity. Yet in many real-world applications, particularly in healthcare, assisted living, and mental health, the underlying motivation behind a behavior is just as important as the behavior itself. For instance, a person repeatedly entering the kitchen at night may be driven by habit, emotional distress, or cognitive confusion. While the action appears identical on the surface, the appropriate system response—whether reassurance, intervention, or monitoring—depends entirely on the context and motivation. Without this layer of interpretation, HAR systems risk misclassifying behaviors, missing critical cues, or losing user trust.

To advance from simple recognition toward genuine understanding, this research draws from cognitive psychology, particularly theories of intrinsic motivation, extrinsic motivation, and amotivation. These concepts explain how internal and external factors influence human behavior. For example, intrinsic motivation reflects the desire to engage in an activity for its own sake, while extrinsic motivation is driven by external rewards or pressures. Amotivation, by contrast, refers to a lack of intent, often observed in older adults or individuals with cognitive disorders.

Incorporating such theories into HAR provides new insight into the intentions, decision-making processes, and contextual factors that underlie human actions. By combining this with semantic reasoning and temporal modeling of behaviors, we aim to infer not only what someone is doing, but also why they are doing it—or failing to act. This capability is vital in adaptive human-computer interaction, assistive technologies, and mental health support, where recognizing the user's psychological state is as important as identifying their physical actions.

This research is guided by the vision of building intelligent systems that simulate human-like understanding. Inspired by neurophysiologist Alain Berthoz—who describes the brain as "a simulator of action and an emulator of reality"—this thesis seeks not only to recognize human actions but also to model the motivational and cognitive processes that drive them. By integrating artificial intelligence, semantic ontologies, and cognitive theory, we aim to develop a more human-centered and interpretable HAR framework.

The core limitation here is that most systems are not equipped to interpret the cognitive or emotional drivers behind behavior. They can detect physical movement, but often miss the deeper story. For example, two individuals sitting still may be in vastly different states—one is relaxing, while the other may be experiencing confusion, sadness, or amotivation. Treating both cases equally can result in delayed care or safety risks.

Cognitive psychology teaches us that actions are shaped by internal factors such as memory, goals, emotions, and motivation. The field differentiates between intrinsic, extrinsic, and amotivated behaviors—distinctions that are critical in domains like mental health or aging, where changes in motivation often precede observable behavioral changes.

Yet, most existing HAR systems lack mechanisms to model these distinctions. They typically classify actions based on sensor inputs, without considering temporal patterns, user context (e.g., age, cognitive status), or environmental variables. These factors are essential for interpreting human behavior in a nuanced and meaningful way.

This thesis responds to this gap by proposing a hybrid, psychologically informed framework that blends data-driven recognition with semantic and cognitive reasoning. By grounding HAR in motivational theory, temporal logic, and user-aware ontologies, we aim to build systems that not only detect behavior but understand it to deliver responses that are adaptive, informed, and aligned with human needs.

1.2 Research Objectives and Methodology Overview

To address the issues outlined above, we propose an interdisciplinary framework that combines semantic knowledge, fine-grained action recognition, and cognitive modeling to reproduce and interpret psychological states through human actions.

Our methodology proceeds in three main phases:

- **Ontological Modeling:** We examine existing ontologies in cognitive psychology, such as OntoCogMémo [2], as well as multilayer structures proposed in [3, 4], to formalize key concepts related to memory, self-regulation, and motivation. This results in a robust semantic foundation for reasoning about human behavior.
- **Spatiotemporal Action Analysis:** We apply fine-grained activity recognition techniques to analyze the temporal and contextual relationships between actions, enabling us to detect inconsistencies, disruptions, or abnormal transitions that may indicate underlying cognitive or motivational states.
- **Hybrid Reasoning and Inference:** In the final phase, we combine data-driven deep learning models with logical reasoning methods, to infer human behavior. This hybrid approach allows us to consider not only observed data but also the broader context in which actions take place.

Together, these phases contribute to a holistic understanding of human activity, supporting applications in personalized coaching, neurological rehabilitation, cognitive monitoring, and support for vulnerable populations.

1.3 Thesis Contributions

This thesis makes several original contributions that advance the state of the art in Human Activity Recognition (HAR) by integrating deep learning, semantic reasoning, and cognitive modeling:

- **A novel hybrid deep learning model:** We propose an effective architecture that combines Convolutional Neural Networks (CNN) with Multi-Layer Perceptrons (MLP) to enhance the recognition of complex human activities from sensor data.
- **A stream reasoning approach for medication risk detection:** Using the LARS (Logic-based framework for Analytic Reasoning over Streams) framework, we develop a method for continuous reasoning over human behavior to detect patterns associated with medication adherence risks in real time.
- **An ontological querying framework for cognitive activity recognition:** We design a system that leverages cognitive psychology theories and ontological modeling to query and interpret human activities from a cognitive perspective.

- **Modeling Human Cognition in Smart Environments:** We introduced a hybrid deep learning and ontology-based framework to infer the motivations behind human activities. By combining sensor data with cognitive concepts like goals and intentions, the system moves beyond activity recognition to explain *why* actions occur.

Together, these contributions aim to bridge the gap between low-level data-driven approaches and high-level cognitive understanding, enabling a more holistic and interpretable model of human activity in intelligent systems.

1.4 Thesis Outline

The structure of this thesis is organized to progressively build an understanding of the theoretical background, methodological developments, and experimental evaluations related to human activity recognition from a cognitive perspective. The content of each chapter is outlined as follows:

- **Chapter 2** presents the theoretical foundations of Human Activity Recognition (HAR) and Cognitive Psychology. It includes an in-depth review of activity classification techniques, cognitive processes, and how these domains intersect in understanding human behavior. In addition, this chapter investigates the state of the art in deep learning models and ontological systems relevant to HAR. It critically examines their respective strengths and limitations and highlights opportunities for hybrid integration.
- **Chapter 3** details the core contributions of this thesis, the development of the proposed hybrid framework, and the reasoning mechanisms employed for cognitive-aware activity recognition.
- **Chapter 4** discusses the practical implementation of the proposed approaches. It describes the datasets used, the experimental setup, evaluation metrics, and the results obtained from comprehensive performance assessments.
- **Chapter 5** concludes the thesis by summarizing the key findings, discussing the implications of the results, and proposing directions for future research and applications.

Chapter 2

Foundations and State of the Art in Human Activity Recognition and Cognitive Psychology

2.1 Introduction

Human activity recognition (HAR) has become one of the most critical research issues. Its importance has been highlighted in various intelligent and distributed surveillance and monitoring systems, Ambient Assisted Living (AAL), well-being management, elderly care, medical diagnosis, and pervasive computing. These fields have witnessed significant development such as the availability of sensors and accelerometers, live data streaming, and particularly the recent advancement in computer vision, machine learning, Artificial Intelligence (AI) and the Internet Of Things (IoT) [5].

HAR is the process of discovering human activity through analysing raw time series data from various categories of sensors embedded in wearable devices or other sensing platforms. Human activity recognition provides a detailed understanding of how humans perform different physical activities in their daily lives, such as walking, running, sitting, standing, and even more complex activities like cooking or cleaning.

Based on how they utilize and combine prior knowledge and data for modelling and recognition purposes, the methods of human activity recognition are classified into i) Data-driven approaches that learn an activity model from the existing data of the user's actions using machine learning and deep learning techniques [6], ii) Knowledge-driven approaches that perform activity recognition through the reliance on logical reasoning and taking advantage of knowledge and sensor data modelling [7] and iii) hybrid approaches that combine both of the previous methods to perform a task.

Knowledge-driven methodologies rely on the incorporation of predefined domain knowledge to establish a semantic representation of human activity, often through the application of structured knowledge modeling techniques. Subsequently, this semantic framework is used to perform inference over the data acquired from various sensor inputs [8]. In contrast, data-driven strategies employ computational techniques such as data mining, machine learning, and deep learning to construct activity models directly from collected datasets [9]. These models frequently use probabilistic or statistical mechanisms for reasoning [10]. Typically, machine learning algorithms are trained using extensive and diverse datasets, where performance is optimized by aligning new sensor observations with previously learned activity templates. During testing, the system attempts to recognize actions by comparing incoming sensor data against these trained models. Nonetheless, such methods often encounter the ”Cold-Start” issue, as they depend heavily on the availability of substantial training data [11]. Moreover, these models frequently lack robustness when deployed in unfamiliar environments, since they tend to be highly specialized to the training context and user [12].

On the other hand, knowledge-driven strategies offer mechanisms for encoding domain knowledge and leveraging logical reasoning to infer human activities [13]. This reasoning process may utilize various artificial intelligence approaches such as case-based reasoning, ontological reasoning, or rule-based inference systems [14]. Despite their strengths in interpretability, these approaches often struggle with issues such as limited generalizability, difficulties in segmenting continuous real-time sensor streams, scalability challenges, and high computational demands [15]. Figure 2.1 depicts the overall pipeline for sensor-based human activity recognition, highlighting the three principal methodological paradigms used to perform the recognition task.

This chapter explores the essential concepts behind Human Activity Recognition and cognitive psychology, with a particular emphasis on where these two fields converge. It begins by outlining the core aspects of HAR, its scope, types of activities, sensor technologies, and key challenges. The discussion then moves through various recognition approaches, including data-driven, knowledge-based, and hybrid methods, supported by commonly used datasets. The chapter concludes by examining how cognitive principles and ontological modeling can deepen our understanding of human activities, setting the stage for the hybrid framework introduced in this thesis.

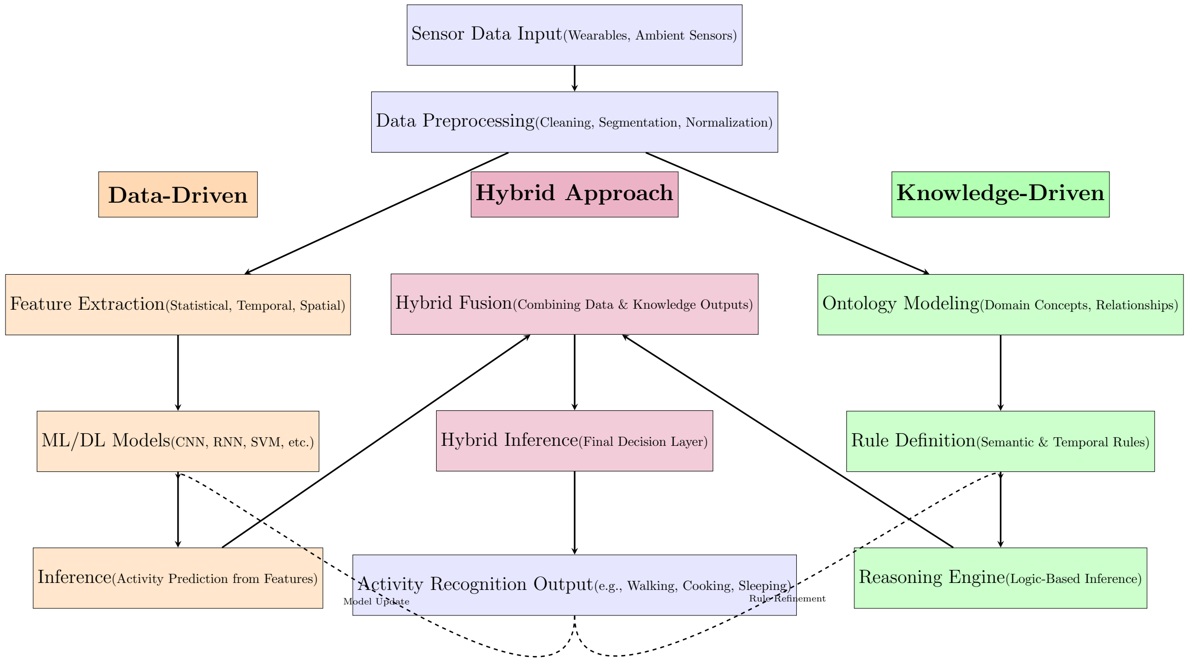


Figure 2.1: Illustration of the human activity recognition framework combining three approaches: data-driven, knowledge-driven, and a hybrid approach that fuses both outputs for improved decision-making.

2.2 Human Activity Recognition: An Overview

This section provides a comprehensive overview of Human Activity Recognition, covering its definition, types of activities, commonly used sensor modalities, and key challenges along with future research directions.

2.2.1 Definition and Scope

HAR refers to the interdisciplinary field that focuses on the automatic identification and classification of human behaviors or actions through data collected from various sensing modalities. These data sources include wearable devices (e.g., smartphones, smartwatches), ambient or environmental sensors (e.g., motion detectors, temperature sensors), and vision-based systems (e.g., video cameras, depth sensors). The activities detected can range from simple motor actions such as walking, sitting, or turning—to more complex and abstract routines like cooking, cleaning, or managing medication intake.

The scope of HAR extends across a wide array of applications including health monitoring, elderly care, ambient assisted living (AAL), sports analytics, smart homes, and human-computer interaction. The primary objective is to develop systems capable of not only detecting what activity is occurring but also understanding the context, purpose, and possible implications of that activity. This involves challenges such as recognizing subtle behavioral cues, handling noisy or incomplete data, and adapting to diverse environments or individual user patterns.

2.2.2 Types of Activities

Human activities can generally be classified based on their timing and structure:

- **Atomic (or simple) activities:** These are short, clearly defined actions that typically occur on their own—such as standing, walking, sitting, or opening a door. They represent the basic building blocks of more complex behaviors [16, 17].
- **Composite (or complex) activities:** These consist of sequences or combinations of atomic actions that together form a more involved task. For example, the activity of "cooking dinner" may include chopping vegetables, boiling water, stirring ingredients, and serving the meal. Recognizing such activities requires understanding the order and relationship between individual steps, often through hierarchical modeling [18, 19].
- **Concurrent activities:** These refer to actions that take place simultaneously or overlap in time, such as walking while talking on the phone or checking the oven while setting the table. Identifying concurrent activities is particularly challenging, especially when dealing with sensor fusion and real-time data processing [16, 18].

In the rest of this thesis, we use the term HAR to refer to sensor-based human activity recognition.

2.2.3 Sensor Modalities in HAR

The accuracy and reliability of HAR systems depend heavily on the types of sensors used for collecting data. These sensors differ in where they are placed, the level of detail they provide, and the kind of information they capture:

- **Wearable sensors:** Commonly embedded in devices like smartphones, smartwatches, or fitness trackers, these include accelerometers, gyroscopes, magnetometers, and heart rate monitors. They are well-suited for capturing body movements and physiological signals, but their effectiveness can vary depending on how and where the user wears the device.
- **Ambient sensors:** Positioned within the environment, ambient sensors detect changes in surroundings and interactions with physical spaces or objects. Examples include passive infrared (PIR) motion detectors, pressure mats, temperature and humidity sensors, light

switches, and door contacts. These sensors are especially valuable in smart home and assisted living environments, offering unobtrusive and continuous monitoring.

- **Vision-based sensors:** Devices such as cameras and depth sensors (e.g., Microsoft Kinect) allow for the visual analysis of human posture, gestures, and activities using images or video. While they provide rich contextual information, their use raises privacy concerns and often demands significant computational resources for real-time processing.

Table 2.2 summarizes some of the most commonly used sensors in the literature for recognizing human activities.

2.2.4 Open Challenges and Future Research Directions

Despite significant advances in sensor-based HAR, several open challenges remain. These challenges not only hinder the deployment of robust HAR systems in real-world settings but also define key areas for future research:

1. **Scarcity of Labeled Data:** High-quality labeled datasets are crucial for training supervised learning models. However, the collection and annotation of sensor data are labor-intensive, time-consuming, and often application-specific. This issue—widely recognized as annotation scarcity—has motivated research into self-supervised learning, unsupervised domain adaptation, and synthetic data generation [20–22].
2. **Class Imbalance:** In real-world datasets, some activities—especially rare or emergency-related ones—are severely underrepresented. This imbalance negatively impacts model generalization, especially for minority classes. Techniques such as re-sampling, cost-sensitive learning, and ensemble models are commonly used to address this issue [23–26].
3. **Complex and Hierarchical Activity Modeling:** Many daily activities are composite in nature, consisting of temporally dependent sequences of atomic actions. Recognizing such structured patterns requires sophisticated data segmentation techniques, temporal modeling, and often hierarchical representations [15, 27–29].
4. **Robust Feature Extraction:** Differentiating between activities that exhibit similar sensor patterns—such as walking vs. running—remains challenging. Feature extraction methods must be sensitive to fine-grained variations while being invariant to irrelevant noise. Recent efforts have explored temporal [27], multimodal [30], and statistical features [29] to improve discriminative performance.
5. **Recognition of Novel and Unexpected Activities:** HAR systems often fail to recognize activities that were not part of the training data, such as unexpected falls or

anomalous behavior. Addressing the distribution shift between training and testing scenarios requires few-shot learning, continual learning, and open-set recognition techniques [31–33].

6. **Concurrent and Multi-User Activities:** In multi-person environments, individuals often engage in concurrent or interleaved activities. Recognizing overlapping actions (e.g., walking while talking on the phone) or distinguishing among multiple users in shared spaces (e.g., smart homes) increases the complexity of recognition models [34, 35]. Personalization and context-awareness are essential in such settings.

7. **Real-Time and Online Recognition:** Many HAR applications, such as fall detection or activity-based reminders, require real-time response and low-latency inference on resource-constrained devices. Online learning approaches that process streaming data incrementally, while managing concept drift, are increasingly important for practical deployments [36, 37].

8. **Security and Privacy:** HAR systems often handle sensitive user information. Ensuring data confidentiality, secure sensor communication, and user anonymity is critical, particularly in ambient intelligence and health-related applications. Privacy-preserving techniques such as federated learning and differential privacy are gaining traction [38, 39].

9. **Noisy and Incomplete Data:** Sensor readings are often noisy, missing, or affected by environmental factors. Models must be robust to such imperfections and capable of distinguishing signal from noise. Attention-based mechanisms and robust statistics have shown promise in mitigating these issues [40].

10. **Adaptive Data Segmentation:** Traditional segmentation approaches—such as fixed or sliding windows—may be suboptimal, particularly for activities with irregular durations. Knowledge-driven or semantic segmentation approaches offer more adaptive solutions that align better with natural activity boundaries [41–44].

11. **Intention and Early Activity Recognition:** Recognizing user intention or anticipating upcoming activities based on initial cues and context is an emerging research direction. Such anticipatory systems can proactively assist users and enhance safety. Integrating commonsense reasoning and knowledge graphs can improve early prediction performance [45].

12. **Sensor Fusion and Management:** Combining data from heterogeneous sensors—wearables, ambient devices, cameras—offers richer context but introduces challenges in synchronization, calibration, and data alignment. Effective sensor fusion requires context-aware modeling and ontology-driven frameworks [30, 46].

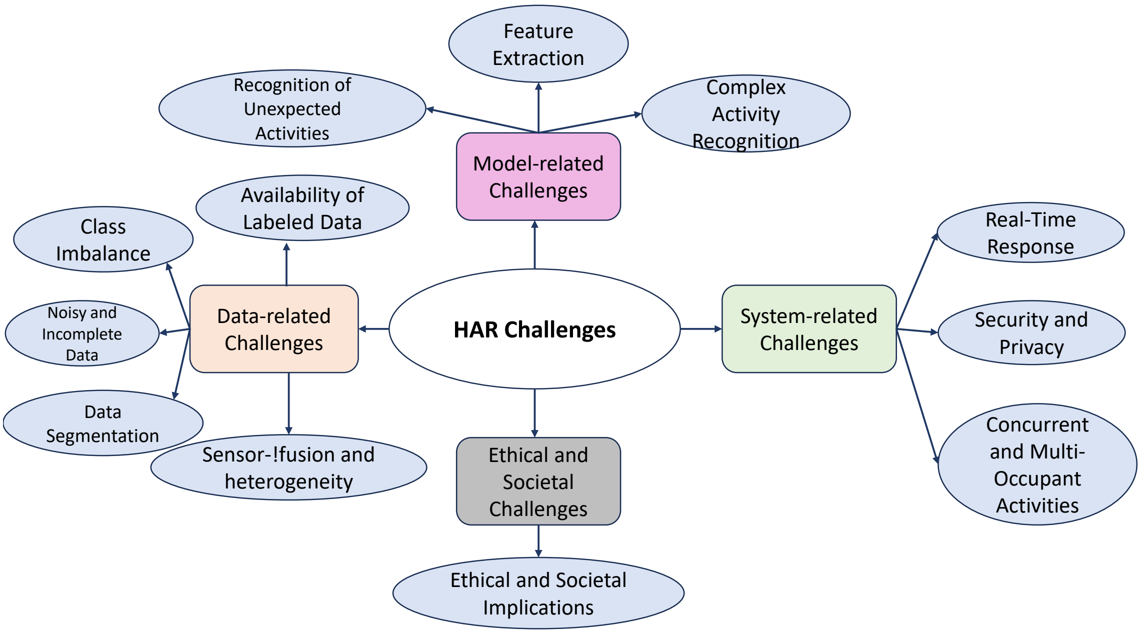


Figure 2.2: An overview of the key challenges in human activity recognition, categorized into model-related, data-related, system-related, and ethical/societal domains.

13. **Generalization, Overfitting, and Efficiency:** Striking a balance between accuracy, computational efficiency, and model generalizability remains a core concern. Overfitting on user-specific patterns or specific environments can hinder real-world applicability. Lightweight architectures and cross-domain adaptation techniques are active research areas [47, 48].

Future research in HAR should emphasize adaptive, context-aware, and personalized models that can learn continuously from streaming data, generalize across diverse populations, and respect user privacy. The integration of symbolic reasoning with deep learning, zero-shot activity recognition, and human-in-the-loop systems represent promising directions to further advance the state of the art. Figure 2.2 summarizes the challenges of HAR ctegorized to: model-related, data-related, system-related, and ethical/societal domains.

2.3 State of the Art in Sensor-Based Human Activity Recognition

In this section, we present a comprehensive survey of the current methods and approaches used in sensor-based HAR, including commonly used datasets, data-driven approaches, semantic-based techniques, and hybrid frameworks.

2.3.1 Datasets

Human Activity Recognition (HAR) research depends heavily on the availability and quality of datasets collected through various sensing modalities. These datasets differ across several dimensions, including sensor types, sampling rates, types of activities, diversity of participants, and the environments in which the data were collected. Most studies focus on sensor-based datasets due to concerns around privacy and the high computational demands of video processing.

Among the most widely used sources are datasets hosted by the UCI Machine Learning Repository [49], valued for their accessibility, standardization, and suitability for benchmarking. Smartphone-based datasets, such as *UCI HAR*[50] and *WISDM*[51], are particularly popular due to the ubiquity of mobile phones equipped with built-in sensors. These datasets typically capture simple activities like walking, standing, and sitting, and often include a relatively large number of participants. However, they generally assume a fixed device position (e.g., in a pocket) and are limited to short-duration, atomic actions, which can restrict their applicability in real-world, dynamic settings.

To overcome these limitations, wearable sensor datasets such as *PAMAP2*[52], *MHEALTH*[53], *Skoda*[54], and *HHAR*[55] incorporate multiple sensors placed on various parts of the body. For instance, PAMAP2 includes heart rate monitoring; MHEALTH integrates ECG and magnetometer data; Skoda captures data from 30 accelerometers worn by a factory worker; and HHAR combines readings from smartphones and smartwatches across nine participants. These datasets offer richer, multimodal representations of activity, though their generalizability may vary depending on participant diversity and recording conditions.

Hybrid datasets, such as *Opportunity*[56] combine wearable, ambient, and location-aware sensors to enable context-aware activity recognition. Opportunity captures activities within a smart environment using inertial sensors and RFID tags, whereas RealWorld HAR collects data in more naturalistic, unconstrained settings. Although these datasets better reflect real-world complexity, they also introduce new challenges, such as positional noise and increased variability.

Ambient sensor datasets—including *CASAS Aruba*[57] and *MIT PlaceLab*[58] shift the focus from motion to environmental context. These datasets rely on event-based sensors like motion detectors, pressure mats, and light switches to infer activities through sequences of environmental events. Such datasets are especially relevant for applications in smart homes and elder care, and typically require models capable of capturing temporal dependencies and handling sparse, asynchronous data.

Another important consideration is *sampling frequency.* High-frequency datasets such as PAMAP2 and Skoda provide fine-grained motion data, which is useful for capturing subtle physical patterns. Others, like WISDM, operate at lower frequencies, while event-driven datasets like CASAS Aruba rely on discrete sensor activations rather than continuous data streams—requiring models to handle irregular temporal patterns.

Finally, *subject diversity* plays a crucial role in model generalizability. Datasets like UCI HAR and RealWorld HAR include data from many users, supporting the development of robust, generalizable models. In contrast, datasets such as Skoda and CASAS Aruba involve data from a single individual, which may benefit personalized modeling but also increase the risk of overfitting to a specific user or setting.

Several specialized datasets have been developed to support research in niche application domains. For example, the *Daphnet FoG* dataset focuses on freezing of gait episodes in patients with Parkinson's disease, while the *Nursing Activity* dataset captures daily tasks performed by medical staff in hospital environments. In the field of sports analytics, *Swimmaster* provides sensor data for analyzing swimming movements. Additionally, behavioral datasets such as *HASC2010* and *HASC2012*[59] offer detailed recordings of daily activities collected via wearable devices, making them valuable for general-purpose HAR tasks.

HAR datasets span a wide spectrum—from readily accessible, smartphone-based collections to complex, multimodal datasets designed for context-rich environments. The selection of an appropriate dataset should align with the specific objectives of the research, whether the focus is on general activity recognition, health monitoring, assistive living, or context-aware automation. Table 2.1 presents a summary of the major datasets discussed in this study, highlighting the diversity in sensor configurations, activity complexity, and target application domains available to researchers in the field.

Table 2.1: Comprehensive Overview of Publicly Available Datasets Commonly Used in Sensor-Based HAR

Dataset	**Source**	**Sensors**	**Instances**	**Activities**	**Subjects**	**Application Context**	**Data Format**	**Notable Features**
HAR	[50]	Accelerometer, Gyroscope	10,299	6	30	Daily living activities	Time series	One of the most widely used datasets.
WISDM	[51]	Accelerometer	1,098,209	6	51	Mobile activity recognition	Time series	Collected using smartphones.
PAMAP2	[52]	Accelerometer, Gyroscope, Heart Rate Monitor	3,850,505	18	9	Physical activity monitoring	Time series	Multi-sensor dataset with comprehensive activities.
OPPORTUNITY	[60]	Accelerometer, Gyroscope, Magnetometer, ambient, and Location tracking	1,164,867	18	4	Real-world activity recognition	Time series	Rich contextual information; multiple sensor modalities.
Aruba	[61]	Motion, Temperature	1,034	11	4	Smart home activity recognition	Time series	Focus on smart environments and daily living.
Milan	[57]	Various sensors (smartphone, wearable)	2 million	10	20	Smart environment monitoring	Mixed	Utilizes both smartphones and wearables for data collection.
MHEALTH	[53]	Accelerometer, Gyroscope, Heart Rate Monitor	120	12	10	Health-related activity recognition	Time series	Focus on health monitoring with real-time data.
ADLs	[62]	Wearable sensors	2,747	2	2	Daily Living Activities	Mixed	Focuses on common activities performed by individuals.
HAPT	[63]	Accelerometer, Gyroscope	10,000	12	30	Daily living activities	Time series	Data collected in real-world conditions.

Table 2.2: Overview of sensor modalities commonly employed in HAR systems, including examples, strengths, limitations, and typical application areas. This comparison aids in evaluating suitable sensors for various recognition tasks.

Sensor Type	**Examples**	**Advantages**	**Limitations**	**Common Applications**	**References**
Accelerometers	Smartphone IMUs, Wearables	Low cost, widely available, energy-efficient	Limited to motion data, prone to drift	Activity tracking, Fall detection	[64]
Gyroscopes	Smartphones, Wearables	Captures angular velocity, complementary to accelerometers	Drift over time, requires calibration	Gesture recognition, Sports analysis	[65]
Magnetometers	Smartphones, Smartwatches	Provides orientation data, used in navigation	Affected by magnetic interference	Indoor navigation, Sports tracking	[66]
Pressure Sensors	Barometers, Footwear	Measures force or altitude changes	Limited data context, requires integration	Posture detection, Climbing activities	[67]
Depth Sensors	Kinect, LiDAR	Captures 3D spatial data, robust to lighting conditions	Expensive, Limited range	Elderly monitoring, Rehabilitation	[68]
Ambient Sensors	Temperature, Light	Non-intrusive, context-aware data collection	Limited activity details, prone to noise	Smart homes, Energy usage monitoring	[69]

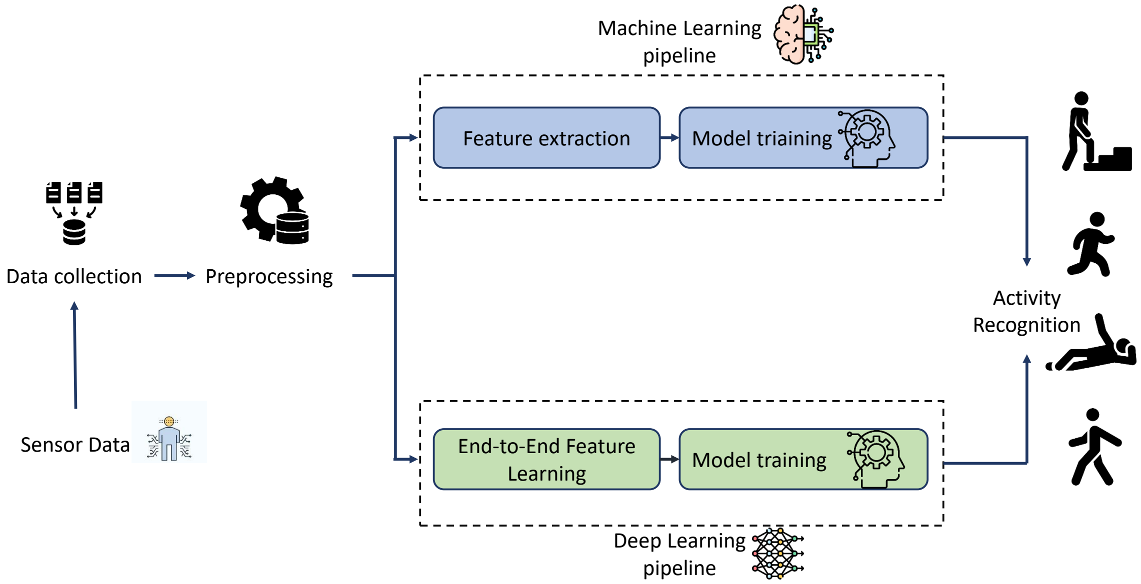


Figure 2.3: General workflow of machine learning and deep learning approaches for Human Activity Recognition (HAR). The figure illustrates the typical stages, including data acquisition, preprocessing, feature extraction or representation learning, model training, and activity classification.

2.3.2 Data-based Approaches for Human Activity Recognition

Learning-based approaches, including machine learning and deep learning, play a key role in sensor-based HAR. These methods enable automatic recognition of activities by identifying patterns in sensor data. This section provides an overview of commonly used learning techniques, highlighting both classical and deep learning models. Figure 2.3 shows the learning approaches workflow for human activity recognition.

2.3.2.1 Machine learning approaches

Machine learning (ML) algorithms have been extensively applied to solve the HAR problem. Classical techniques, including Random Forest (RF) [70], Markov models [71], and Support Vector Machines (SVM) [72, 73], have shown considerable success, particularly in controlled environments where datasets are limited. Despite their effectiveness, traditional ML approaches often depend heavily on hand-crafted features and require significant preprocessing. This reliance can lead to inefficiencies, particularly in real-time applications or when scaling up to larger datasets. Additionally, these methods tend to struggle in incremental or unsupervised learning contexts, where shallow feature-based techniques often fail to capture the complexity of human activities. The authors of [74] introduce a new method known as Class Incremental Random Forests (CIRF). During the incremental learning phase, it employs the Gini index or information gain for leaf node splitting, along with a splitting method based on the separating axis theorem for adding interior nodes. Building upon the importance of capturing temporal relationships, a Hidden Markov Model (HMM) and static machine learning classifiers are utilized in [75] to enhance activity recognition in patients with motor disabilities. This approach captures the temporal relationships among successive samples by using the outputs of static classifiers as observations in the HMM. Furthermore, the researchers in [76] propose a two-stage continuous HMM framework that utilizes the hierarchical structure of fundamental activities for human activity recognition. This architecture supports a variable number of states and iterations, allowing for the application of different feature subsets to various subclasses, thereby reducing computational overhead. In addition to these approaches, the work in [77] describes a methodology that combines Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA), and modified Weighted Support Vector Machines (WSVM). Initially, the features obtained from LDA are supplemented with significant principal components to improve recognition accuracy.The authors of [78] present an infrared–ultrasonic sensor fusion method for fall detection using Support Vector Machines (SVMs), focusing on senior healthcare. In this study, the user's position, dimensions, and temperature profile are estimated through an innovative sensory fusion method. The research examines several feature sets within the SVM framework and evaluates their impact on the accuracy of fall detection.

While traditional machine learning methods like Random Forest, Markov models, and Support Vector Machines have demonstrated success in HAR, their limitations become evident in more complex settings. These models are less efficient for large-scale or real-time applications due to the need for extensive preprocessing and hand-crafted features. Moreover, they often underperform in incremental or unsupervised learning scenarios because of their reliance on shallow features. However, recent advancements, such as Class Incremental Random Forests and hybrid models that combine static classifiers with Hidden Markov Models, provide more robust solutions by addressing these challenges and improving performance in dynamic environments.

By introducing new techniques for feature selection, temporal relationships, and sensor fusion, these methods enhance the scalability and resilience of HAR systems across various applications, such as incremental learning and healthcare. Despite these advancements, there is still room to further improve the flexibility and generalization of these models in real-world settings.

2.3.2.2 Deep learning approaches

In the domain of Human Activity Recognition (HAR), various deep neural network architectures have been designed and utilized to enhance performance. Among the most prominent and widely adopted are Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), along with their respective variants. CNNs are characterized as deep learning models composed of several hierarchical layers, including convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers, typically culminating in a softmax activation function for classification.

Moreover, Convolutional Neural Networks (CNNs) are well-suited for capturing localized patterns in multimodal sensory input, effectively leveraging spatial locality to achieve robust recognition performance through translational invariance. In a study by Avilés-Cruz et al. [79], the authors proposed an approach for human activity analysis based on three parallel CNN streams. Each network independently performs local feature extraction, and their outputs are subsequently integrated during the classification phase to enhance overall prediction accuracy. The whole CNN architecture is based on a feature fusion of three CNN networks that are: a fine CNN, a medium CNN and a coarse CNN. Chen et al. [80] proposed an acceleration-based human activity recognition algorithm using a convolution neural network (CNN). They modify the network convolution kernel to adapt the characteristics of tri-axial acceleration signals. On the other hand, Khan et al. [81] proposed a transfer learning motivated Heterogeneous Deep Convolutional Neural Network (HDCNN) thet is specifically adapted to the properties of CNNs. The model supposes that the weights' relative distribution in the various layers is not variant since the set of the observed activities set is constant and does not show any change. Moreover, Almaslukh et al. [82] have suggested a deep learning architecture where they employ time-domain features and learned ones for an independent HAR that is marked with an efficient position.

A HAR framework has been suggested in [83], based on CNNs. The authors propose an efficient approach to human activity recognition using deep ensemble learning. They address the challenge of recognizing confusing activities by combining feature extraction from multiple sensors, deep ensemble learning for model training, and feature fusion to improve recognition accuracy. An attention-based convolutional neural network for HAR has been presented by Wang et al. [40] from weakly labelled data. The model can concentrate on labelled activity amid a long data sequence and, at the same time, filter many signals of background noise In another study, Wan et al. [84] proposed a convolutional neural network-based real-time human activity recognition method using an ensemble classification algorithm with CNN for local feature extraction. They relied on CNN, LSTM (Long Short Term Memory), BLSTM (Bidirectional LSTM), MLP (Multilayer Perceptron), and SVM (Support Vector Machine) models to address the problem of the complex structure of the networks that can deal with large datasets, and the problem of ascertainment in case of high similar human activities. In [85], the authors proposed a light CNN using Lego filters of lower dimensions which are used as stackable Lego bricks for conventional filters to efficiently reduce computational and memory costs.

A new recurrent setting of dual attention method has been suggested by Gao et al. [86]. It combines temporal attention and channel attention on a CNN. Besides, [87] introduced the

Self HAR method that combines self-training, multi-task self-supervision, and the techniques of semi-supervised learning to make the networks of CNN transferable to new situations.

Rachid et al.[88] proposed an Adaptive CNN for energy-efficient HAR (AHAR). They constructed a novel adaptive CNN architecture that uses an output block predictor to select a part from the initial architecture to be used in the inference step. Besides, Cheng et al. [89] proposed a computationally efficient CNN using conditionally parameterized convolution for real-time HAR on mobile and wearable devices by replacing conventional convolutions with a conditionally parameterized convolution (CondConv).

Another architecture of deep learning neural networks is RNN (Recurrent Neural Network) where the connections between the nodes produce a directed graph along a temporal sequence. RNNs extract the temporal dependencies and can progressively learn information over time intervals that suit streaming sensory data in HAR [90].

The LSTM is one of the famous RNN-based architectures which can process entire sequences of data and single data points. By setting logic gates in a calculation cell, the LSTM network gathers efficient time series information. In [91], authors proposed a personalized human activity recognition framework by building a RNN model for each individual. Ordonez et al. [92] proposed a generic deep framework for activity recognition based on convolutional and LSTM recurrent units which are suitable for multimodal wearable sensors; it can naturally perform sensor fusion; it does not require specialized knowledge in feature design; and it explicitly models the temporal dynamics of feature activations.

In [93], the authors constructed a binarized BLSTM-RNN model where the input and output of all hidden layers and weight parameters are binary values in this model. A framework that uses a Deep Recurrent Neural Networks (DRNNs) has been suggested by Murad et al. [94] for building models of recognition which can capture long-range dependencies in variable length input sequences. Cascaded, bidirectional, and unidirectional architectures are presented based on LSTM-DRNNs. The authors in [95] gave a new model of HAR that is weakly supervised based on recurrent attention learning. An agent had been trained for information extraction from weakly labelled sensor data by adaptively selecting a location sequence. Relying on new rewarding strategies, the authors used reinforcement learning to train the model.

Machoe et al. [96] proposed a RNN-based subject-independent human activity recognition method using a windowing approach. The weight initialization process aims at avoiding the hidden neurons' behaviours that are not wanted. Residual bidirectional long short-term memory (Res-BidirLSTM) has been introduced by Zhao et al. [97]. To speed up the algorithm's training, the memory employs modified deep RNNs using residual connections among stacked cells to avoid the problems of gradient vanishing.

Wang et al. [98] proposed a new structure called hierarchical deep LSTM (H-LSTM) to recognize human activities after preparing sensor data and selecting and extracting features by the time-frequency domain method. Zhou et al. [99] designed a semi-supervised deep learning framework to get a more accurate HAR. The authors used reinforcement learning by developing a Deep Q-Network (DQN) using a distance-based reward function to enhance the learning efficiency in IoT environments. Then they developed a multi-sensor-based data fusion method to integrate body sensor data and used a LSTM-based classification network to extract contextual features and identify perfect patterns.

In the literature, given the advantages of both CNNs and RNNs, some researchers proposed deep hybrid approaches. Some deep learning architectures are combined to construct a deep HAR model. A hybrid deep framework called CNN-LSTM-ELM has been presented by Sun et al. [100]. The framework combines an LSTM network, a deep CNN, and an Extreme Learning Machine (ELM) classifier. It aims at decreasing the time of executing the HAR task. A framework that blends a CNN model and a Deep Bidirectional Long Short-Term Memory (DBLSTM) has been proposed by Su et al. [101] to get an automated HAR system. DBLSTM preprocesses sequential raw data, and a bidirectional vector is generated. CNN uses this vector to extract local features, and in the end, human activities are recognized via the fully connected layer using a SoftMax function.

A framework that combines LSTM with convolutional layers has been suggested by Xia et al. [102]. The raw data are given to a two-layer LSTM followed by convolutional layers. Then, an average pooling layer is applied to substitute the fully connected layer after convolution to reduce the model parameters, followed by a batch normalization layer to accelerate the convergence. Another architecture of activity recognition that is based on deep learning has been suggested by Mutegekiet al. [103].A long short-term memory network combined with a convolutional neural network is created. This CNN-LSTM network is spatially and temporally deep. Gumaei et al. [104] proposed a multi-sensors-based and hybrid framework that consists of different types of recurrent units (SRUs and GRUs) for human activity recognition. Zhou et al. [105] presented a WiFi-based activity recognition framework using an Autoencoder to feature extraction. Besides, they relied on a CNN-LSTM architecture to extract local and global features. Yuki et al. [106] utilized a dual-stream Conv-LSTM network where the first stream deals with a smaller time length and the second stream deals with a longer time length for an efficient human activity recognition system. The work of Agti et al. [107, 108] presents a novel hybrid deep learning network that combines Convolutional Neural Network (CNN) and Multi-Layer Perceptron (MLP) layers to autonomously learn and classify activities from sensor data. The CNN-MLP approach facilitates the multi-class categorization of human activities by effectively capturing temporal data.

Table 2.3: An overview of the most important hybrid approaches features.

Reference	Year	Method Features
Yamada et al. [109]	2007	Ontologies + Naive Bayes Model
Riboni et al. [10]	2009	Ontologies + Statistical classifiers
Ye et al. [110]	2014	Ontologies + Machine learning algorithms
Chen et al. [111]	2014	Ontologies + Incremental Learning Algorithm
Gayathri et al. [112]	2015	Ontologies + Hierarchical clustering
Riboni et al. [113]	2015	Supervised Learning + Markov Logic Network method
Ahmadi et al. [114]	2017	Clustering model + Ontologies
Gayathri et al. [115]	2017	Ontologies + Markov Logic Network
Civitares et al. [116]	2018	Ontologies + Active learning(Random Forests algorithm) + Markov Logic Network
Mojarad et al. [117]	2018	Ontologies + Random Forests algorithm + Bayesian network
Civitares et al. [118]	2019	Machine learning + ontologies
Rueda et al. [119]	2019	Deep Learning + computational state space models (CSSM)
Sukor et al. [120]	2019	Description Logic Reasoning(ontology) + Machine learning model
Li et al. [15]	2019	Marcov Logic Network + Rule-based modeling and inference
Civitares et al. [121]	2019	Markov Logic Network + ontologies + Active learnig
Civitares et al. [122]	2019	Semi-supervised Active and incremental Machine Learning + Ontologies
Mojarad et al. [123]	2020	Deep Learning(CNN-LSTM model) + Ontologies
Bettini et al. [124]	2020	semi-supervised learning + semantic context-aware reasoning(ontology-based)
Riboni et al. [125]	2020	Unsupervised machine learning(HMM) + ontologies
Bettini et al. [126]	2021	Machine learning model + ontology

Deep learning techniques are characterized by their deep architecture [127], meaning they have multiple layers of neurons. This allows them to learn hierarchical features in a scalable way, making them highly effective for processing large and complex datasets. This scalability is a crucial aspect of deep learning, enabling it to achieve superior performance in various applications.

Multiple deep learning architectures encode features under various perspectives. CNNs are competent in capturing the local connections of multimodal sensory data. The translational invariance introduced by locality leads to accurate activity recognition [128]. Recurrent neural networks (RNNs) incrementally learn information through time intervals and extract temporal dependencies. Therefore, they are suitable for streaming sensory data in human activity recognition. The structure of deep neural networks allows them to be composed easily into united networks with a total optimization function. The latter makes allowance for miscellaneous deep learning techniques including deep active learning [129], and deep attention systems [130]. Furthermore, most existing models of pattern recognition concentrate in first place on learning from static data; while real-life activity data flows require robust incremental online learning

The network allows the automatic learning of features in deep learning rather than manual extraction. Besides, a high-level representation in the deep layer can be extracted by the deep neural network, making it suitable for activity recognition tasks, particularly complex activities. Unlabeled data for model learning can be exploited by deep learning models, with a large quantity of unlabeled data. On the other hand, they can generally be transferred to new tasks where there are few or no labels.

Nevertheless, deep learning methods face different limits which affect deep learning-based human activity recognition models. The need for domain knowledge in feature extraction tasks can affect the time processing and accuracy of constructing a human action recognition system. Furthermore, the features learned are shallow and may only be used for the recognition of basic actions like ”sitting” or ”running”, and not for complex activities like ”drawing” for example, or activities that need context-awareness. Moreover, traditional approaches require large labelled datasets to construct the model, even though most of the datasets of human activities are not labelled. Therefore, in unsupervised learning tasks, these models' efficiency is undermined. However, for model training, existing deep generative networks can take advantage of unlabeled samples. In contrast, existing deep generative networks can exploit unlabeled samples for model training.

As to data-driven approaches, they do not integrate a semantic viewpoint of the analyzed activities causing the absence of semantic reasoning on the observed actions, the situation of the actors, the motivation to act, and the objectives [6]. As a solution, knowledge-based approaches introduced human behaviour reasoning and action context recognition to improve the activity recognition task.

2.3.3 Semantics in human activity recognition systems

HAR holds immense potential for impactful applications across various fields, including smart homes, healthcare, and human-computer interaction. Ongoing research in HAR is poised to reveal transformative solutions that not only enhance human-centric technology but also improve quality of life by addressing both methodological and ethical considerations.

We can classify the approaches of Knowledge-based human activity recognition into three classes according to the reasoning way used to recognize activities. In this line, we get: Logic-based, rule-based, and ontology-based approaches according to the reasoning way used to recognize activities. Some works have combined two kinds of methods to construct a knowledge-based framework for human activity recognition. Today, the technique that is the most used in knowledge-driven human activity recognition problems is Ontology.

Ontologies can clearly describe the knowledge, and can also offer an organized, structured and machine-readable representation of this knowledge. Moreover, they have the expressive power to support the reasoning process [131]. Their main drawback is that knowledge-driven approaches are adaptation problems since representation is usually perceived as generic and in static conditions [132]. In this context, according to [133], it is hard to generate a complete model of the environment.

Such knowledge-based methods have been suggested in many studies. A knowledge-driven approach has been presented by Okeyo et al. [134]. This method uses temporal knowledge and ontological modelling formalisms to recognize complex activities, including concurrent and interleaved ones. Along the same line, it employs rule-based temporal inference for identifying complex activities, while for the simple ones, it uses ontological reasoning. Noor et al. [8] proposed a new reasoning algorithm for activity recognition in an intelligent environment. The method integrates ontological reasoning with Dempster-Shafer's proof theory. Uncertainty is quantified, while contextual information is aggregated. Moreover, a degree of belief is provided. The latter facilitates more robust decision-making in activity recognition.

Recognition of Activities of Daily Living RADL is a system that has been provided by Bae et al. [135] for the analysis and identification of ADLs for the elderly in smart homes. The constructed ontology supports the semantic discovery of devices, location, and the environment in such homes. For the present situation, the activities and the appropriate service are detected by the reasoning procedure.

A continuous sensor sequence can be segmented automatically into significant partitions by KCAR (A Knowledge-driven approach for Concurrent Activity Recognition) exploring the semantics of sensor events in the ontological model [136]. The complex problem of concurrent activity recognition is transformed into a simple single-user sequential activity recognition task by KCAR through the segmentation of the interwoven sensor sequence for each ongoing activity. It uses ontological reasoning and statistical methods to measure similarity and match hierarchical concepts. A continuous sensor sequence can be segmented automatically into significant partitions by KCAR by exploring the semantics of sensor events in the ontological model.

SPARQL-based reasoning has been presented by Meditskos et al. [137]. It supports crucial tasks of inference in the domain of activity interpretation. However, they are not supported by the standard OWL 2 semantics. They improve the framework with a conceptual layer through the ontology models of DOLCE + DnS Ultralite(DUL). Khattak et al. [138] proposed an approach to improve the elderly health and life quality by analyzing the nutritional consumption and health activity information. Authors rely on an ontology-based method to model the daily life activities and the details of the patient's profile, allowing the analysis of fine-grained situations for personalized service recommendations. Ahmadi et al. [139] proposed an ontology-based approach that allows separating the input data into categories of non-overlapping activity events regarding the context information. Then, these datasets are segmented and then mapped into activity classes with a pre-trained classification model. Relying on ontology, Noor et al. [46] introduced a sensor fusion-based approach. They aim to resolve uncertainties resulting from missing sensor data through wearable and ambient sensors. A framework that captures new knowledge from activity models has been given by Safyan et al. [140] for the development of behavioural changes in the activity recognition model through an ontology. This framework learns specialized and extended activities from the available activity models. The objective is to incorporate new characteristics into established activity patterns, recognize previously unseen activities, and enrich the ontology by capturing evolving representations that contribute to the continuous refinement and expansion of the knowledge model.

An approach that exploits an algorithm of unsupervised machine-learning algorithm has been suggested by Gayathri et al. [141]. This approach extracts knowledge from unlabeled data through contextual pattern clustering. Moreover, it represents it as an ontology which provides enhanced activity recognition thanks to clear reasoning and representation. An activity model is represented by the proposed Fuzzy Ontology Activity Recognition (FOAR) in the form of a fuzzy temporal ontology. Activity recognition is accelerated by the Fuzzy SWRL rules modelled in the ontology. An approach for timely daily activity recognition has been presented by Liu et al. [142] from an incomplete stream of sensor events. It allows the process to begin when a daily activity begins.

Activity features are extracted from a subset of the most relevant sensor events rather than from the entire set of sensor activations triggered by daily routines. In this context, Hooda et al. [143] developed an ontology-based HAR model that utilizes the *Activity Recognition Ontology* to formalize the activity recognition process—specifically by capturing the relationships between inferred high-level activities and fundamental actions—and the *Sensor Measurements Ontology* to represent sensor-derived data.

While knowledge-driven approaches may be considered static and limited in scope [6], they offer semantic clarity and are relatively easy to deploy. However, they often lack robustness in managing uncertainty and handling temporal dynamics. Typically, these methods do not incorporate mechanisms to evaluate the comparative effectiveness of different reasoning strategies, as long as the strategies adequately represent human actions. Moreover, they generally do not integrate learning capabilities with logical inference techniques [7]. A commonly acknowledged limitation is the static nature of the inferred models, which are not easily adaptable to individual user preferences or behavioral variations [143].

Recognizing a wide range of activity types presents a significant challenge, particularly in indoor environments where constructing comprehensive activity models is inherently difficult. To address this limitation, the integration of data-driven techniques with knowledge-based approaches has been proposed as a promising strategy. This hybridization facilitates the creation of more complete and flexible activity models capable of capturing diverse human behaviors. Moreover, such models can evolve over time, allowing them to adapt dynamically to changes in users' behavior patterns [120].

2.3.4 Hybrid approaches for human activity recognition

Many researchers have proposed hybrid methods that integrate data-driven and knowledge-based techniques for human activity recognition (HAR), aiming to combine the strengths of learning algorithms and symbolic reasoning.

Yamada et al. [109] introduced an ontology-based approach combined with Naive Bayes to model object relationships in diverse environments. Riboni et al. [10, 144] combined ontological reasoning with statistical inference, allowing inference of activities beyond the reach of statistical methods. Ye et al. [110] developed USMART, which segments continuous sensor streams and leverages sensor semantics using an ontological model.

Gayathri et al. introduced EPAM [115], using unsupervised learning to extract patterns from unlabeled data within an ontology-based framework. In [112], they also proposed a hybrid method using Markov Logic Networks (MLN) to model both simple and composite activities using contextual attributes and expert-defined temporal rules. Riboni et al. [113] presented FABER, combining supervised learning with symbolic reasoning to detect abnormal behaviours in smart homes.

Ahmadi et al. [114] proposed a framework integrating clustering for action detection, ontology-based reasoning for activity recognition, and an estimation module for energy waste. Mojarad et al. [117] improved HAR by combining random forests, ontological reasoning, and Bayesian networks. Civitarese et al. [116, 118] developed a hybrid real-time recognition system integrating contextual reasoning with supervised learning, enhancing recognition through environmental context. In another work, Sukor et al. [120] introduced a system that evolves user activity models dynamically, merging initial ontological inference with data-based adaptation. Rueda et al. [15, 119] combined deep learning with symbolic models using MLNs to model complex daily activities probabilistically.

Civitarese et al. [145] presented POLARIS, which recognizes ADLs by combining ontological reasoning with probabilistic refinement. Contextual and semantic correlations are inferred through ontology, refined by probabilistic reasoning for online recognition. Mojarad et al. [146] developed a context-aware AmI system using LSTM models for location/object detection and HACON ontology for reasoning. PASP [147, 148] was used to identify abnormal behaviour through probabilistic rules. Riboni et al. [125] proposed a system combining unsupervised resident separation with knowledge-based multi-resident activity recognition.

Civitarese et al. proposed CAVIAR [122, 124], a semi-supervised hybrid system using semantic context to refine classifier predictions, which are then used in active learning. In [149], Pro-CAVIAR optimized this pipeline with probabilistic knowledge-based refinement of machine-learned distributions using context data. In addition, Mojarad et al. [123] introduced a dual-component HAR system: a CNN-LSTM-based recognition module and a symbolic reasoning module using ASP over HAT ontology to interpret user behaviour.

Table 2.3 summarizes the combined features of the studied hybrid approaches.

Table 2.4: Comparative study of hybrid approaches for sensor-based HAR.

Reference	Year	Advantages	Limitations
Yamada et al. [109]	2007	Excellent noise tolerance; High accuracy in activity space detection.	Low recognition of independent activities.
Riboni et al. [10]	2009	Good accuracy; Capable of handling complex activities.	Temporal relationships of activities not considered; Incapable of recognizing similar activities.
Ye et al. [110]	2014	Automatic semantics-driven segmentation; Captures user intention; Does not require annotated training data.	Sensitive to sensor noise; Considers at most two consecutive events; Struggles with similar nearby activities.

Reference	Year	Advantages	Limitations
Chen et al. [111]	2013	Flexible, reusable, and scalable.	Sensitive to noise; Limited activity types; Does not address complex activities.
Gayathri et al. [112]	2015	Reduces computation and memory cost.	Does not model uncertainty; Does not handle complex activities.
Riboni et al. [113]	2015	High accuracy.	Does not address multi-inhabitant or interleaved activities.
Ahmadi et al. [114]	2018	Good accuracy; Real-time response; Easy adoption in real-world setups.	Ignores action sequence and activity dependencies; High initialization effort.
Gayathri et al. [115]	2017	Good F-measure; Handles uncertainty and incomplete data; Recognizes composite activities.	Lacks temporal activity characterization; Does not support multiple occupants; Limited inconsistency handling.
Civitares et al. [116]	2018	No initial training set required.	Requires protecting user privacy and data integrity; High correlation mapping effort.
Mojarad et al. [117]	2018	Detects and corrects context inconsistencies.	Limited consideration of complex activities.
Civitares et al. [118]	2019	Supports context reasoning.	Requires large training datasets.
Rueda et al. [119]	2019	Adds contextual information; Recognizes actions, locations, and dependencies.	Low performance.
Sukor et al. [120]	2019	High recognition rate; High scalability.	Cannot distinguish concurrent or nonsensical object usage.
Li et al. [15]	2019	Accurate under noise and missing data; Supports complex activity recognition.	Does not consider inter-activity relations or user profiles in multi-resident scenarios.
Civitares et al. [121]	2019	Reduces number of feedback requests.	Re-initializes recognition model after updates.
Civitares et al. [145]	2019	Supports complex activity recognition; Probabilistic and deterministic rules; Handles noise; Adapts to changes.	Not reusable for all activity types; Ignores health-related concept drift.

Reference	Year	Advantages	Limitations
Civitares et al. [122]	2019	High recognition rate; Fewer queries via active learning.	Ignores semantic uncertainty and temporal activity sequences.
Mojarad et al. [123]	2020	Models human activities and contexts; Infers behavioral knowledge.	Ignores uncertainty in data and rules.
Bettini et al. [124]	2020	Considers context data; Good accuracy.	Cannot handle incomplete knowledge and sensor limitations.
Bettini et al. [149]	2020	High recognition rate; Works in unusual scenarios; Fewer active learning queries.	Ignores personal habits and specific contexts.
Riboni et al. [125]	2020	Multi-resident recognition without labeled data.	Assumes at most two persons; Does not model inter-inhabitant interaction.
Bettini et al. [126]	2021	Provides semantic explanations.	Does not present temporal dependencies between states.

A recent approach for explainable activity recognition (XAR) has been proposed by Bettini et al. relying on machine learning models (Random Forests and SVM) [126]. The researchers extracted a stream of temporally characterized semantic states from each sensor data stream, giving input to a trained classifier. After that, considering the feature's importance and values generates explanations from a set of features over the semantic states. Nonetheless, the drawback of the approach is that temporal dependencies between states are not shown by the semantic explanations. Table 2.4 presents a comparative study of the most important hybrid approaches for human activity recognition and highlights the advantages and limits of each work.

Deep learning techniques have achieved impressive results in Human Activity Recognition (HAR), especially when large labeled datasets are available. These models excel at learning complex patterns but may underperform when dealing with subtle or rare activities. In contrast, knowledge-driven methods—such as rule-based or ontology-based systems—offer semantic reasoning and work well in low-data scenarios. However, they often lack adaptability and struggle in uncertain environments.

Hybrid approaches have emerged to bridge this gap by combining the reasoning strengths of knowledge-based models with the learning power of deep neural networks. These integrated systems show improved robustness and generalization, particularly in complex or data-limited settings. Ultimately, the choice of method depends on factors like activity complexity, data

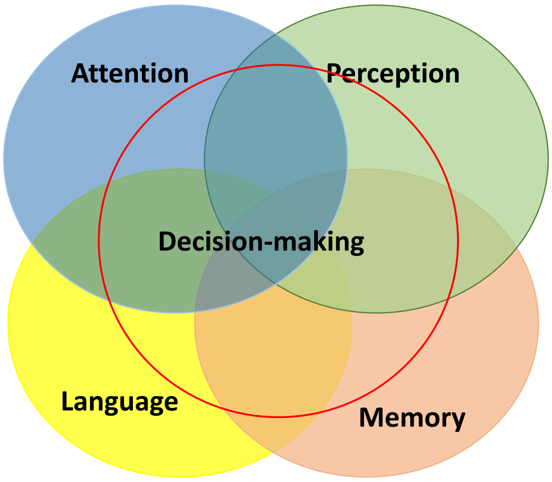


Figure 2.4: Interrelated cognitive processes—including perception, attention, memory, and language—that collectively support and enable human decision-making. These components form the cognitive foundation for interpreting information, reasoning, and acting within the environment.

availability, and the trade-off between interpretability and performance. Hybrid models present a promising path toward more context-aware and reliable HAR systems.

2.4 Cognitive Psychology: Theoretical Foundations

Cognitive psychology focuses on the scientific study of mental functions such as perception, attention, memory, reasoning, language, and problem-solving. It emerged in response to the limitations of behaviorism, framing the mind as an information-processing system and offering key insights into human cognition that influence artificial intelligence and intelligent system design [150, 151].

Core cognitive processes—perception, attention, memory, learning, and language—work in tandem to help individuals interpret, respond to, and navigate their environments. Perception interprets sensory input, attention filters relevant stimuli, memory stores and retrieves information, learning adapts behavior through experience, and language enables communication. Together, they form a dynamic and interconnected system essential for decision-making and problem-solving. Computational modeling of these processes underpins adaptive, human-like AI (see Figure 2.4).

Cognitive architectures formalize these mental functions in software. Symbolic models like ACT-R and SOAR use rule-based reasoning [152], while connectionist models (e.g., neural networks) mimic learning via weighted patterns [153]. Hybrid models combine both to support flexible reasoning and adaptation, advancing applications in HCI, robotics, and cognitive modeling. Mental representations such as schemas, scripts, and mental models structure human knowledge and support reasoning, anticipation, and problem-solving [154]. These ideas inspire computational representations like ontologies and knowledge graphs [155], though capturing their richness remains a challenge.

Cognitive psychology also explains how internal states (e.g., beliefs, emotions, intentions) influence behavior. Processes like attention, memory, and decision-making mediate real-time human actions, providing a foundation for designing AI that can predict or replicate human responses [156, 157].

By translating cognitive processes into algorithms—such as rule-based inference or neural learning—researchers create systems capable of human-like reasoning and interaction. This convergence of psychology and computer science fuels advancements in natural language processing, decision support, and autonomous agents [158].

Despite progress, modeling human cognition is difficult. Individual variation, unconscious processes, emotional influences, and sociocultural context introduce complexity that current AI systems cannot fully replicate [159]. Situated and embodied cognition theories emphasize that thought is context-dependent and distributed across mind, body, and environment [160]. Tools, gestures, and social interactions extend cognitive capacities. This necessitates context-aware and affective computing that enables machines to adapt to real-world dynamics [161].

Integrating cognitive principles into Artificial Intelligence(AI) and Human-Computer Interaction(HCI) supports the creation of human-centered systems. These can interpret behavior, adapt interfaces, recognize emotions, and collaborate effectively leading to intelligent systems that are explainable, responsive, and aligned with human needs [162, 163].

2.5 Cognitive Foundations for Human Activity Recognition

Understanding human behavior and the motivations that drive action is a complex and multidimensional challenge. According to Previc [164], the cognitive space plays a central role in structuring actions and determining the range of possible actions. Human actions are not solely governed by the real-time perception of environmental stimuli, but also by prior knowledge and experiences. In this sense, actions emerge from the integration of perceptual input with stored knowledge and memory.

This thesis is grounded in cognitive psychology, particularly focusing on theories of intrinsic and extrinsic motivation, as well as the concept of amotivation. Intrinsic motivation refers to engaging in an activity for the sake of personal satisfaction, while extrinsic motivation involves performing actions to gain rewards or avoid negative consequences. Deci and Ryan [165] introduce a third dimension—amotivation—where an individual perceives no connection between their actions and the outcomes. This is especially relevant in the context of aging populations or individuals with dementia, where motivation to act may be diminished or absent. In some cases, we observe pathological perseveration, where individuals persist in repeating irrelevant actions, or conversely, exhibit inaction or abandonment of activities. These patterns are particularly evident in neurocognitive disorders or under specific motivational deficiencies. A major challenge, then, is to model such behaviors in a way that reveals the psychological or neurophysiological state behind them. Can one infer an individual's mental or physical state from the analysis of their past and present actions, particularly when these actions appear inconsistent with rational expectations?

The neurophysiologist Alain Berthoz, a professor at the Collège de France, supports a view shared by philosophers such as Maurice Merleau-Ponty and Edmund Husserl: the brain is not merely a processor of information but a simulator of action and an emulator of reality. According to Berthoz, action precedes language in shaping our perception of the world. He illustrates this through the case of patients suffering from perseveration syndrome, where individuals become trapped in a loop of repeated actions due to an inability to suppress ongoing behavior in favor of initiating new ones. This condition highlights a possible internal competition between past and future actions, where the inability to transition from one to another impairs planning and decision-making.

Understanding and modeling human activity requires more than just recognizing observable actions; it necessitates insight into the underlying cognitive mechanisms that drive and structure those actions. For systems to move beyond surface-level classification and achieve human-like interpretation, they must integrate cognitive processes and models into their architecture.

Several key cognitive processes are particularly relevant to HAR systems that seek psychological realism and contextual awareness:

- **Perception:** This is the process through which humans interpret and make sense of sensory inputs. In HAR, perception maps onto how systems interpret sensor data to infer user context. A cognitive-inspired design of sensor systems can enhance their ability to capture meaningful environmental and behavioral cues by aligning data processing with how humans naturally perceive their surroundings.
- **Attention:** Attention governs how cognitive resources are selectively focused on specific stimuli or tasks. In HAR, attentional modeling is especially important for multitasking

or concurrent activities. By simulating human attentional mechanisms, HAR systems can prioritize certain sensor inputs or behavioral cues over others, improving accuracy in complex or noisy environments.

- **Memory:** Both short-term (working) and long-term memory play critical roles in human behavior. Memory influences how actions are initiated, sequenced, and repeated over time. Incorporating memory-inspired architectures—such as recurrent or temporal models—enables HAR systems to recognize activity patterns based on historical context and user-specific routines.
- **Executive Functions:** These include high-level cognitive capabilities such as planning, decision-making, task switching, and inhibitory control. They are essential for modeling activity sequences involving choices, interruptions, and goal-directed behavior. HAR systems informed by executive function theories are better equipped to handle dynamic and adaptive behavior, such as switching from one task to another or responding to unexpected events.
- **Self-Regulation:** This refers to an individual's ability to monitor and adjust behavior in alignment with internal goals and social norms. In HAR, insights into self-regulation can help systems detect deviations from typical behavior patterns that may indicate intentional changes, errors, or even potential health issues. Modeling self-regulatory behavior allows HAR applications to interpret not just *what* a person is doing, but *why*.

Grounding HAR systems in these core cognitive processes enables the development of models that not only detect activities but also interpret behaviors in a way that mirrors human understanding. This enhances the system's intelligence, adaptability, and relevance in real-world contexts.

In parallel, cognitive science has proposed several foundational models to formalize how the mind functions—each offering frameworks that can inspire the design of HAR systems:

- **Information-Processing Models:** These models conceptualize the mind as a sequential computational system, where inputs are processed through stages to yield outputs. They align well with symbolic reasoning systems and rule-based inference mechanisms commonly found in early artificial intelligence research.
- **Connectionist Models:** Inspired by the brain's architecture, these models—often implemented as neural networks—perform parallel distributed processing. They are particularly effective in capturing emergent patterns and behaviors from large data sets and form the basis of many modern deep learning-based HAR systems.

- **Cognitive Architectures:** Frameworks like ACT-R and Soar offer unified theories of cognition by integrating symbolic and sub-symbolic processing. These architectures simulate multiple cognitive tasks, including memory, attention, and problem-solving, and provide a strong foundation for modeling sequential, goal-directed behaviors and decision-making processes [166].

A crucial aspect of human activity is its temporal and goal-oriented nature. Human behavior is typically purposeful and unfolds over time, rather than being random or isolated. Accurately capturing this temporal structure is essential for HAR systems aiming to model or predict behavior. Time-sensitive reasoning mechanisms enable systems to infer causality, sequence dependencies, and even future intentions from partial or ongoing activities. In particular, goal-driven modeling supports anticipatory recognition and enhances system robustness in real-world scenarios, where users may perform multiple or overlapping tasks [167].

By integrating these cognitive principles—both in terms of process and theoretical models—HAR systems can move toward a more human-centered approach. Such systems are capable not only of recognizing actions but also of interpreting them within a meaningful psychological and contextual framework.

2.6 Conclusion

This chapter provided a broad overview of Human Activity Recognition (HAR), combining sensor-based methods with cognitive psychology and semantic modeling. It reviewed data-driven, deep learning, and hybrid approaches, emphasizing how hybrid methods balance interpretability and adaptability. Cognitive theories were introduced to explain the motivations behind human actions. The integration of these disciplines points toward more intelligent, human-centered HAR systems.

In the following chapter, we present the core contributions of this thesis, detailing how our proposed approaches advance the integration of cognitive, semantic, and deep learning techniques in HAR.

Chapter 3

Research Contributions

3.1 Introduction

This chapter presents the main contributions of the thesis, which collectively aim to advance the field of human activity recognition by integrating data-driven techniques, semantic reasoning, and cognitive modeling. First,a stream reasoning approach is introduced to detect medication-related behavior risks in real time using the LARS framework. Second, a novel hybrid deep learning model is proposed, combining convolutional and fully connected layers to enhance recognition accuracy across multiple datasets. Third, a cognitive framework based on ontological querying is presented, providing a high-level perspective on interpreting human behavior. And finally, a hybrid framework that combines deep learning and ontological reasoning is proposed for modeling human cognition in smart environments.

These contributions mark an important step toward bringing cognitive psychology into the design of intelligent systems. By modeling and interpreting human thoughts and behaviors using computational methods, this work helps create environments that are more aware of people's needs and contexts. It bridges the gap between raw sensor data and a deeper understanding of human activity, making ambient intelligence and human-centered systems more responsive, adaptive, and meaningful.

3.2 Stream Reasoning approach on Human Behavior for Medication Risks Detection using LARS Framework

Given the constantly changing nature of human behavior and the continuous flow of sensor data, traditional methods—which often rely on static analysis—struggle to keep up. These approaches are generally not equipped to handle the dynamic characteristics inherent in streaming data.

To address this, stream reasoning has emerged as a powerful solution, allowing systems to process and interpret data on the fly. This capability is especially important in time-sensitive applications such as healthcare, where immediate recognition of activities can support prompt intervention. Stream reasoning techniques rely on temporal queries to monitor and interpret ongoing data, enabling the detection of patterns as they develop.

Recent advancements highlight a growing trend toward knowledge-driven models that incorporate semantic understanding. These approaches aim to create richer representations of human behavior by linking contextual factors such as location, objects, and actions through semantic relationships [168, 169].

Stream reasoning extends traditional stream processing by incorporating logic-based approaches to effectively handle and analyze continuous, high-velocity data streams. One foundational concept in this area is the use of *windows*, first introduced in the Continuous Query Language (CQL) [170], which makes it possible to perform queries on recent subsets of incoming data, mimicking traditional database operations. This concept was later adapted to the Semantic Web domain, allowing for structured, static-like queries over dynamic data streams.

Among the most influential developments in stream reasoning is the LARS framework [171], which builds on Answer Set Programming (ASP) [172] by introducing temporal operators and window functions for logical reasoning over time-stamped data. A LARS rule is typically expressed as $\lambda_0 \leftarrow \lambda_1, \ldots, \lambda_n$, where each λ_i denotes a logical formula. Operators such as $\boxplus^w$ for windowing, $\diamond\lambda$ and $\Box\lambda$ for expressing eventuality and invariance over time, and $@_t\lambda$ for referencing specific time points, enable precise reasoning over temporal contexts.

A stream reasoning model based on LARS is proposed to support the detection of potential medication-related risks among elderly individuals. This application demonstrates how stream-based logic reasoning can be applied to human activity analysis, particularly in healthcare settings. The use of LARS in this way highlights its potential for real-time interpretation of activity patterns, enabling proactive support in scenarios where timely interventions are critical.

3.2.1 LARS: a Logic-based framework for Analyzing Reasoning over Streams

LARS offers a structured approach for reasoning over time-constrained and data-driven streams. In this framework, a data stream is represented as a pair (T, f), where T denotes a finite interval of natural numbers, and f defines a mapping between a subset of atoms $\mathcal{A}$ and their associated time points. A key element in this model is the use of window functions, which serve to isolate and operate on specific portions of the stream. These functions enable temporal slicing of the data, allowing the reasoning process to focus on relevant segments within defined time intervals.

3.2.2 LARS Window Functions

In the LARS framework, given a stream S and a time point $t \in \mathbb{N}$, a window function (w) produces a substream of S, referred to as a window. The syntax of LARS formulas follows a defined grammar, where p is an atom from $\mathcal{A}$:

$$\phi ::= p \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi \mid \Diamond\phi \mid \Box\phi \mid @t\phi \mid \boxplus^{w}\phi$$

Shorthand notations are also used in this context: $\bot := p \wedge \neg p$, $\top := \neg\bot$, and $\phi \rightarrow \beta := \neg\phi \vee \beta$. A pointed interpretation in LARS is expressed as $\mathcal{I} = (S^{\star}, S, t)$, where $S^{\star} = (T^{\star}, v^{\star})$ is the full stream, $S = (T, v)$ is a substream such that $S \subseteq S^{\star}$, and $t \in T^{\star}$ represents a specific time point. Whether a LARS formula ϕ holds at a given interpretation is determined using the inductive semantics defined as follows:

$$\begin{aligned}
\mathcal{I} &\models p && \iff && p \in v(t), \text{ for } p \in \mathcal{A} \\
\mathcal{I} &\models \neg\phi && \iff && \mathcal{I} \not\models \phi \\
\mathcal{I} &\models \phi \wedge \beta && \iff && \mathcal{I} \models \phi \text{ and } \mathcal{I} \models \beta \\
\mathcal{I} &\models \phi \vee \beta && \iff && \mathcal{I} \models \phi \text{ or } \mathcal{I} \models \beta \\
\mathcal{I} &\models \Diamond\phi && \iff && \exists t' \in T : (S^{\star}, S, t') \models \phi \\
\mathcal{I} &\models \Box\phi && \iff && \forall t' \in T : (S^{\star}, S, t') \models \phi \\
\mathcal{I} &\models @t'\phi && \iff && (S^{\star}, S, t') \models \phi \text{ and } t' \in T \\
\mathcal{I} &\models \boxplus^{w}\phi && \iff && (S^{\star}, w(S, t), t) \models \phi
\end{aligned}$$

If a generic window function of type w_i is used, along with a stream function of type ch and a parameter vector of type i, the notation $\boxplus^{\mathbf{x}}_{\iota,ch}$ is applied to denote the corresponding window operator.

Two common window types in LARS are the tuple-based and time-based windows. The *tuple-based window* $w_{\#}(S, t, \mathbf{y})$ extracts a substream from S within the interval $[t_k, t_l]$, returning a defined number of tuples. In contrast, the *time-based window* $w_{\tau}(S, t, \mathbf{y})$ selects a time-relative segment of S with respect to the reference point t', capturing all preceding tuples if $k = \infty$, and all succeeding ones when $l = \infty$.

In addition to these windowing mechanisms, LARS incorporates temporal operators to address time-based reasoning. The operators $\Box$ and $\diamond$ are used to express whether a condition holds throughout a window or at least once, respectively—similar to constructs in modal logic. Specific time references can also be made using the `@` operator, which allows direct evaluation of formulas at designated time points.

3.2.3 LARS Program Structure

A LARS program is defined as a finite set $\mathcal{R}$ of rules, each expressed in the form $r = \phi \leftarrow \beta$, where both ϕ and β are LARS formulas.

A stream S is said to satisfy a rule r at a specific time point t if and only if $(S, t) \models \beta \rightarrow \phi$, which is denoted by $(S, t) \models r$.

A program is interpreted as the conjunction of its individual rules. Accordingly, a stream S satisfies the entire program $\mathcal{R}$ at time t if and only if all rules $r \in \mathcal{R}$ are satisfied at that time point. This is expressed as $(S, t) \models \mathcal{R}$.

Within the set of atoms $\mathcal{A}$, a distinction is made between extensional atoms $\mathcal{A}^{\epsilon}$ and intentional atoms $\mathcal{A}^{I}$. Extensional atoms represent input data and, by definition, do not appear in the heads of rules.

3.2.4 Proposed Method

This section outlines the proposed method for LARS-based stream reasoning in the context of human activity analysis [173], with a specific focus on identifying potential medication-related risks. The overall architecture of our proposed method is depicted in Figure 3.1.

Scenario

To introduce the approach, consider the following illustrative scenario:

Mary is an elderly woman who lives alone in her home. She experiences challenges due to the progressive decline of her short-term memory and motor abilities. To improve her safety and support her cognitive functioning, Mary utilizes a tailored Ambient Assisted Living (AAL) service.

As part of her daily routine, Mary is required to take medication multiple times per day, with a specific time interval between each dose. Given her condition, there is a risk that she may forget whether she has already taken a dose, leading to what is referred to as the *missed dose risk*.

In addition to this, there is a possibility that Mary might take the same medication more frequently than recommended, particularly if she forgets that she has already taken it. This situation presents an *overdose risk*, such as when multiple doses are taken within a short time interval.

The reasoning system proposed here is designed to monitor Mary's medication intake behavior and identify such risks. It generates alerts to minimize the likelihood of missed doses and ensures the safe administration of medication by detecting both missed doses and potential overdoses due to repeated intake within a short period.

- **Missed dose risk**: This risk refers to situations in which Mary unintentionally fails to take her medication at the scheduled times. Such omissions may negatively affect her health, potentially aggravating her condition and increasing the likelihood of hospitalization.
- **Overdose risk**: This risk occurs when Mary inadvertently consumes a dose that exceeds the prescribed amount, which may happen due to memory lapses or misunderstandings regarding the dosage schedule. Overdosing poses significant health threats, including possible organ damage, and in severe cases, it may be life-threatening.

A stream reasoning approach grounded in the LARS framework is proposed, with a focus on detecting potential medication-related risks among elderly individuals. The objective is to analyze data streams and monitor behavioral patterns to infer hazards such as missed doses and overdose situations. This is particularly pertinent for older adults who live independently. To facilitate this, a set of rules is formulated using the LARS syntax and temporal operators, thereby enhancing the reasoning capability of the system and contributing to the overall quality and safety of medication administration for senior patients. Importantly, this contribution constitutes the first known application of the LARS framework for stream-based reasoning in the domain of human behaviour analysis.

3.3 A novel CNN-MLP deep learning model for sensor based human activity recognition

Human Activity Recognition (HAR) is a foundational component within the domains of artificial intelligence and machine learning, with broad applications spanning healthcare, sports performance analysis, smart home systems, and intelligent environments. The development of precise HAR models is essential for improving individuals' quality of life, particularly by supporting personalized health monitoring and treatment strategies [174]. In the context of sports, these models contribute to performance evaluation and the design of tailored training regimens [175]. Additionally, in smart environments such as homes and urban infrastructures, HAR facilitates improved automation, safety, and energy management [176].

Sensor-based HAR relies on data collected from three main types of sensors: body-worn devices (such as accelerometers and gyroscopes [177, 178]), object-based sensors (including RFID tags

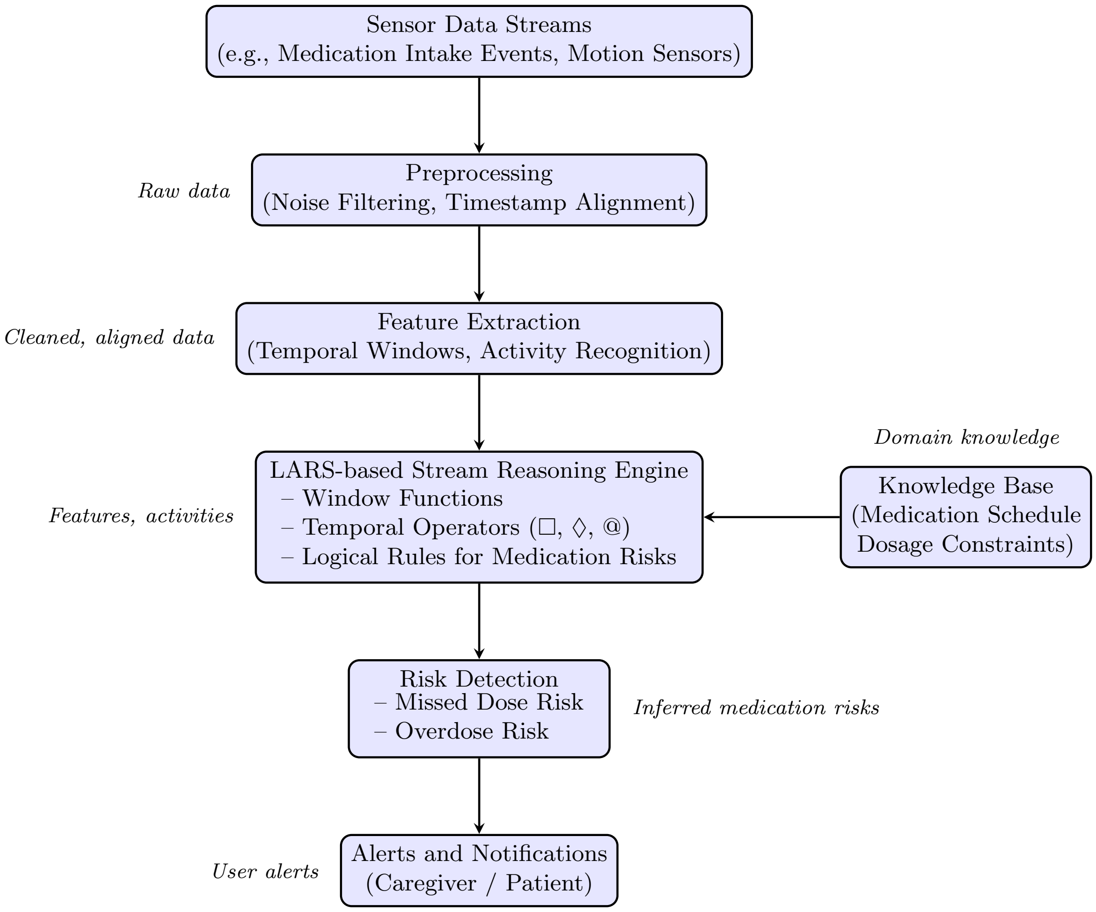


FIGURE 3.1: Architecture of the LARS-based stream reasoning system for detecting medication-related risks from continuous human activity data streams.

and motion sensors [179]), and ambient sensors (like sound, light, and temperature sensors [180]). While each of these sensor types brings valuable capabilities, effectively combining data from multiple sources remains a significant challenge in practical applications.

Recently, deep learning methods have made remarkable strides in HAR by automatically learning hierarchical features directly from raw sensor inputs [180]. Among these, CNNs and LSTMs have become especially popular, with hybrid models such as CNN-LSTM, Bi-LSTM, and attention-based approaches delivering strong improvements in accuracy and robustness on benchmark datasets [84, 102, 103, 181]. In particular, CNN-LSTM architectures effectively capture both spatial patterns and temporal dependencies, enhancing recognition of complex activity sequences.

However, many existing methods face significant challenges, such as achieving real-time performance, adapting to various sensor types, and maintaining robustness in uncontrolled environments. A number of prior studies focus on individual datasets or traditional architectures, resulting in models that often lack scalability and cross-domain generalizability. Additionally,

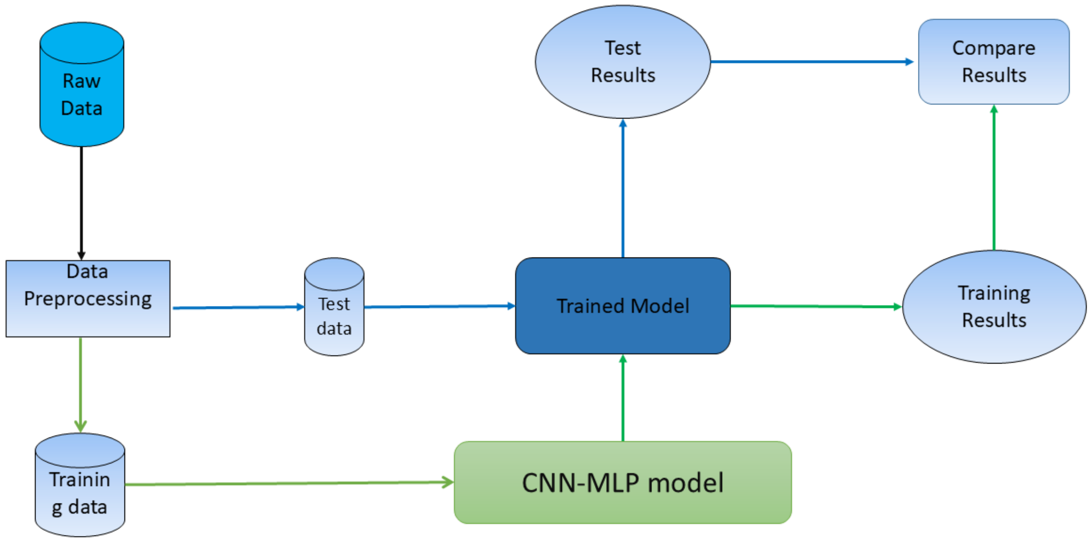


Figure 3.2: Architecture of the proposed model of CNN-MLP architecture for human activity recognition.

deploying these models on edge devices is limited by high computational and memory requirements [182]. To overcome these issues, we introduce a novel hybrid deep learning model that combines Convolutional Neural Networks (CNN) with Multi-Layer Perceptrons (MLP). The CNN layers are designed to extract spatial and temporal features from raw sensor data, while the MLP layers improve classification of these extracted representations. The architecture incorporates Batch Normalization, Global Average Pooling, L2 regularization, and Dropout techniques to enhance generalization, prevent overfitting, and reduce computational complexity—making it well-suited for real-time applications.

3.3.1 Proposed Architecture

Our CNN-MLP model is carefully designed to capture both the spatial and temporal characteristics of sensor data, enabling accurate multi-class classification. The architecture is composed of distinct layers, each contributing to the process of extracting meaningful features and making predictions. This design is especially effective for handling time-series data, where understanding the sequence and structure of sensor inputs is crucial. By combining CNN and MLP components, the model is able to recognize both motion patterns over time and spatial relationships across different sensors.

Figure 3.2 visually depicts the overall architecture of our approach, encompassing all necessary components.

Convolutional and pooling layers serve as the core building blocks of our model architecture, playing a crucial role in capturing intricate spatial patterns from the input data. Through a series of convolutional operations, the model is able to detect meaningful features and patterns that emerge across different segments of the dataset. By stacking these layers, the network builds hierarchical feature representations that reflect the underlying complexity of human activities.

Our design incorporates three dedicated 1D convolutional layers, each progressively refining the spatial characteristics extracted from the raw data. The first convolutional layer performs the following transformation:

$$X_i^{(l)} = \phi\left(\sum_j W_{ij}^{(l)} * X_j^{(l-1)} + b_i^{(l)}\right) \tag{3.1}$$

where $X_i^{(l)}$ is the output of the i-th filter at layer l, $W_{ij}^{(l)}$ represents the convolutional kernel weights, $b_i^{(l)}$ is the bias term, and $\phi(\cdot)$ is the ReLU activation function.

The first convolutional layer includes 64 filters and is followed by a max-pooling layer that reduces the spatial dimensions of the resulting feature maps while retaining key information. The second convolutional layer, equipped with 128 filters, increases the model's capacity to capture more intricate spatial patterns. This is again followed by a max-pooling layer to streamline the learned features. The third and final convolutional layer, using 256 filters, focuses on extracting fine-grained spatial characteristics from the input. Each max-pooling operation is computed as:

$$P_i^{(l)} = \max_{k \in \mathcal{N}(i)} X_k^{(l)} \tag{3.2}$$

where $P_i^{(l)}$ is the pooled output, and $\mathcal{N}(i)$ defines the receptive field for the pooling operation.

The architecture also integrates Multi-Layer Perceptron (MLP) layers, which play a key role in capturing the dynamic temporal patterns present in human activity data. These layers help the model understand the evolving relationships between consecutive time steps by analyzing how sensor values change over time. The MLP component includes two fully connected layers: the first with 512 neurons and the second with 256 neurons. Each dense layer applies the following transformation:

$$h^{(l)} = \sigma(W^{(l)} h^{(l-1)} + b^{(l)}) \tag{3.3}$$

where $h^{(l)}$ represents the activation at layer l, $W^{(l)}$ is the weight matrix, and $b^{(l)}$ is the bias term.

ReLU activation functions are employed to introduce non-linearity, allowing the model to learn and represent more complex patterns in the data. To summarize the extracted spatial and temporal features, a Global Average Pooling layer is applied. This layer computes the mean value of each feature map, helping reduce dimensionality while retaining essential information:

$$GAP_i = \frac{1}{N} \sum_{j=1}^{N} X_{ij} \tag{3.4}$$

where N represents the number of spatial locations.

To enhance training stability and efficiency, batch normalization is applied. This normalization process follows:

$$\hat{x}^{(l)} = \frac{x^{(l)} - \mu^{(l)}}{\sqrt{\sigma^{2(l)} + \epsilon}} \tag{3.5}$$

where $\mu^{(l)}$ and $\sigma^{2(l)}$ are the batch mean and variance, respectively, and ϵ is a small constant for numerical stability.

The output layer acts as the final decision stage, responsible for classifying human activities. It is a fully connected layer where each neuron corresponds to a specific activity class, and it uses the softmax activation function to produce a probability distribution over the possible classes:

$$p_i = \frac{e^{z_i}}{\sum_j e^{z_j}} \tag{3.6}$$

where z_i is the input to the i-th neuron before activation.

To improve convergence during training, a learning rate scheduling strategy is employed. The learning rate starts at 0.001 and stays constant for the first ten epochs. After that, it gradually decreases following an exponential decay rule defined as:

$$\eta_t = \eta_0 e^{-0.1t} \tag{3.7}$$

where η_t is the learning rate at epoch t and η_0 is the initial learning rate.

The model is trained using the Adam optimizer, with an initial learning rate set to 0.001. By integrating convolutional, pooling, dense, and normalization layers, the architecture effectively captures relevant features and supports accurate classification. This design results in a reliable and high-performing HAR system.

The proposed CNN-MLP hybrid architecture represents a significant contribution to advancing Human Activity Recognition (HAR). By combining the spatial feature extraction strength of Convolutional Neural Networks (CNNs) with the classification capabilities of Multi-Layer Perceptrons (MLPs), the model demonstrates improved performance across diverse application domains, including smart homes, healthcare, and personal fitness.

Through rigorous evaluation on both smartphone-based datasets (UCI HAR, WISDM, and PAMAP2) and a smart home dataset (CASAS Aruba), the architecture exhibits strong adaptability and effectiveness across heterogeneous sensor modalities and activity contexts. These results validate the model's robustness and generalizability in practical HAR scenarios.

The general design of the enhanced CNN-MLP model is illustrated in Figure 3.3.

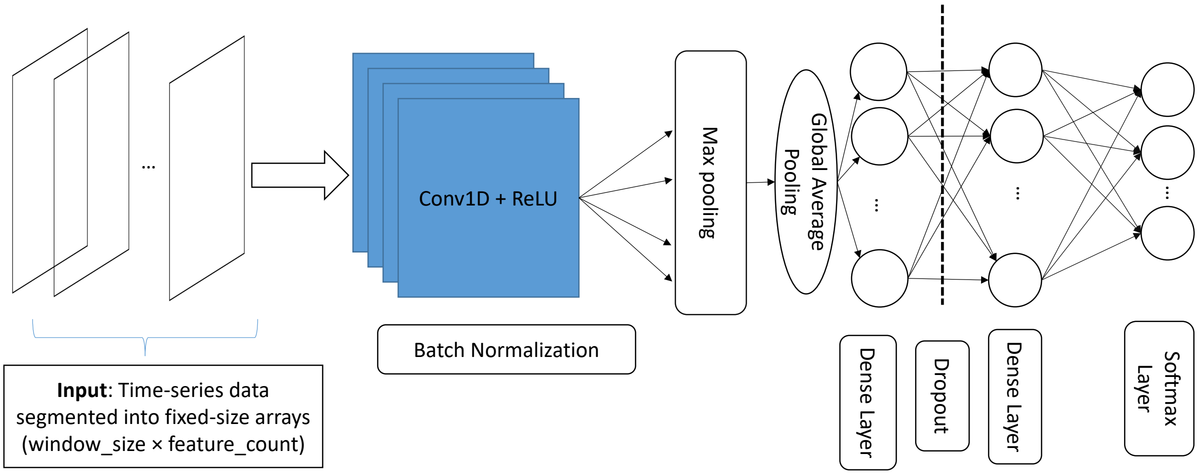


FIGURE 3.3: Proposed Enhanced CNN-MLP Architecture for Human Activity Recognition

The proposed model offers several advantages:

- Effective extraction of spatial and temporal features from time-series data using the CNN layers.
- Robust performance against overfitting through techniques like Batch Normalization and Dropout.
- Lightweight design optimized for deployment on resource-constrained devices, achieved through reduced complexity and computational optimizations.
- Scalability to handle multiple datasets and diverse sensor modalities, ensuring broad applicability in real-world scenarios.

The training process incorporates an adjustable learning rate schedule, with an initial value of 0.001. The learning rate decays exponentially during training:

$$\eta_{t+1} = \eta_t \cdot e^{-0.1} \tag{3.8}$$

where η_t is the learning rate at epoch t, helping the model converge more efficiently.

This CNN-MLP module forms the foundation of the proposed system by accurately identifying human activities from sensor data. It extracts high-quality features that enhance subsequent ontology reasoning, contributing to the overall performance and adaptability of the system.

This contribution underscores the value of hybrid deep learning strategies in enhancing recognition accuracy and system reliability. It lays a solid foundation for the development of intelligent, resource-efficient HAR systems capable of recognizing complex human behaviors across multiple environments.

3.4 Ontological Querying-based Cognitive Perspective for Activity Recognition.

Recent developments in human activity recognition (HAR) within smart home environments have increasingly emphasized leveraging sensor data to interpret and predict user behavior. While conventional machine learning approaches have shown promise for recognizing simple, well-defined activities, they often encounter limitations when faced with complex temporal sequences or rich contextual dependencies. To address these gaps, deep learning techniques have emerged as powerful alternatives. For example, Depthwise Separable Convolutional Neural Networks (DS-CNN) [183] offer a computationally efficient design by decoupling spatial feature extraction from feature fusion, though they may fall short in capturing dynamic environmental context. In parallel, Deep Recurrent Neural Networks (DRNNs) [184] have demonstrated strong capabilities in modeling temporal dependencies in sensor data from mobile devices, yet their adaptability to diverse, real-world settings remains an ongoing challenge.

Explainable Artificial Intelligence (XAI) techniques, including LIME, SHAP, and Anchors, have been widely adopted to improve the interpretability of machine learning models [185]. These methods often generate explanations in natural language or visual formats to support user understanding. Despite their benefits, ensuring a meaningful balance between explanation transparency and user trust remains an open issue. For example, SHAP explanations, while mathematically rigorous, can sometimes undermine user confidence by revealing overly complex or unexpected model behavior. Recent advancements, such as the integration of fine-tuned Word2Vec embeddings [186], have contributed to better recognition performance and have eased the burden of manual labeling. Nevertheless, such techniques may struggle to generalize across diverse household settings and are frequently trained on benchmark datasets that offer limited representational diversity.

In a related direction, Sahoo et al. [187] employed ESP32-based Wi-Fi sensing to gather Channel State Information (CSI) for real-time human activity recognition, leveraging lightweight machine learning models for on-device inference at the edge. This strategy demonstrates promising low-latency performance; however, the simplicity of the features extracted can hinder the system's ability to accurately capture more nuanced or complex activities. Similarly, Fahad et al. [188] proposed a framework that combines a probabilistic neural network with autoencoders to identify anomalous patterns in user behavior. Nevertheless, their reliance on boxplots to define the ground truth for anomalies introduces potential issues in terms of reliability and consistency during model evaluation.

To overcome the limitations of conventional HAR techniques, hybrid deep learning models have been explored. For example, Mekruksavanich et al. [189] developed a CNN-LSTM framework that jointly captures spatial and temporal characteristics of sensor data. While this integration improves recognition performance, its adaptability to diverse and dynamic environments remains limited. In a similar vein, Li et al. [190] proposed a method that incorporates a spatial distance matrix alongside a sensor contribution analysis mechanism to enhance recognition accuracy through the use of wide convolutional networks. Although these models demonstrate strong performance in controlled conditions, they often face difficulties in generalizing across varied contextual settings encountered in real-world applications.

In parallel, semantic-based approaches have gained attention due to their potential to deliver high-level abstractions and interpretations of sensor inputs [191]. Ontology-driven frameworks, often leveraging RDF and SPARQL, provide a formal structure to represent complex interrelations among entities such as users, activities, and their environments [9, 137, 192]. These semantic models enable advanced reasoning over activity sequences and contextual cues. Nonetheless, traditional ontological representations frequently fall short in expressing the dynamic nature of human behavior. To mitigate this limitation, Larhrib et al. [193] advocate for augmenting ontologies with behavior flow concepts, thereby supporting more effective validation and testing of activity patterns.

Ontology-based frameworks for human activity recognition have shown promise in capturing the structural and temporal complexities of everyday behaviors. For instance, Hooda et al. [143] introduce a model that encapsulates human actions, interactions with objects, and contextual information from the environment, thereby enabling the inference of high-level activities. Similarly, Ni et al. [194] develop a foundational ontology grounded in the NeON methodology, which integrates user profiles, smart home contexts, and Activities of Daily Living (ADLs).

Their design is built upon the DOLCE+DnS Ultralite (DUL) ontology, promoting semantic interoperability and reusability across applications.

In parallel, context-aware HAR systems have incorporated dynamic environmental information to refine activity recognition processes [123]. These systems leverage spatial, temporal, and interactional data to derive more accurate interpretations of user behavior. A notable example is the hybrid architecture proposed by Mouloued et al. [195], which combines ontological context modeling with reasoning mechanisms such as event calculus and probabilistic planning, implemented via Answer Set Programming (ASP) and Partially Observable Markov Decision Processes (POMDP). Their follow-up work [196] further integrates deep learning techniques with commonsense reasoning, resulting in a more resilient and adaptive framework for activity prediction in smart environments. Recent advancements reflect a clear shift toward integrating machine learning techniques with semantic models and contextual reasoning. This multidisciplinary approach supports the development of more adaptive, transparent, and semantically enriched HAR systems, aligning well with the complex and dynamic nature of smart home environments.

3.4.1 Proposed framework

In the following, we outline the core components of the proposed methodology, including the design of the model architecture. The figure 3.4 illustrate the different compenents of this proposed framework [197].

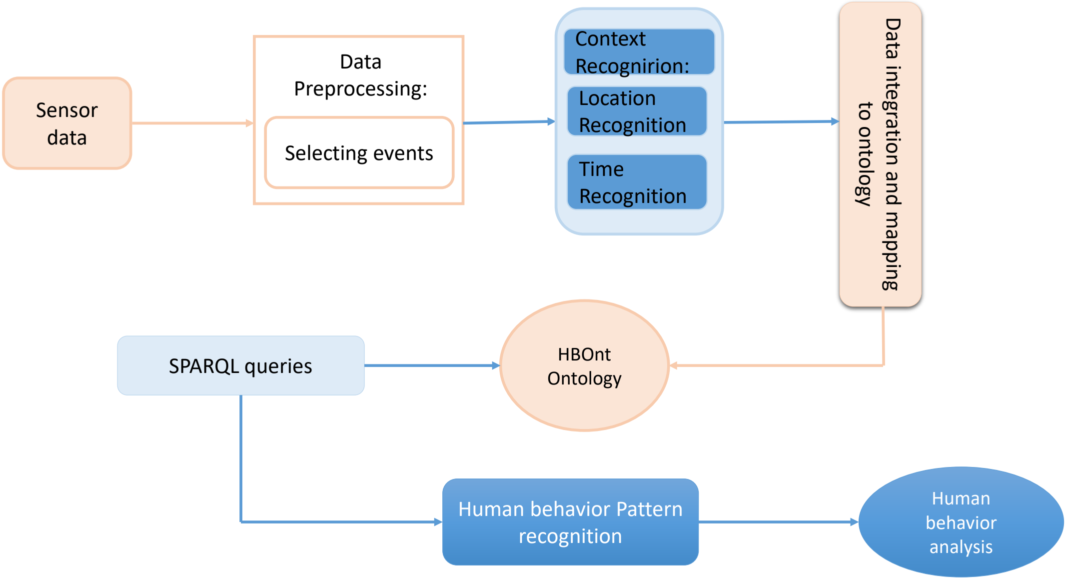


Figure 3.4: General architecture of the proposed ontological quering-based cognitive perspective framework for human behavior recognition

3.4.1.1 Mapping the dataset to an ontology

This section describes the process of translating information from the smart home environment into an ontology that semantically captures the relationships among actions, sensors, and contextual elements. The resulting ontology establishes a formalized framework for representing human activities alongside their corresponding sensor data, thereby enabling sophisticated querying and reasoning capabilities.

The ontology encompasses several key classes, including *Activity*, *ActivityTime*, *Location*, *Sensor*, *Object*, and *Subject*. Instances of the *Activity* subclasses differentiate multiple occurrences of the same activity by day and time, allowing for precise temporal representation of activities across different periods. Sensors are classified into three distinct categories, each associated with specific locations within the smart home environment.

These concepts are interconnected through a variety of relationships such as *occursOn*, *hasStartTime*, *hasEndTime*, *hasActivityLocation*, *hasActor*, *hasTimeOfDay*, and *hasSensorID*, among others. These properties collectively capture the intricate links between activities, sensors, and their contextual attributes. By providing a structured semantic representation, this ontology supports advanced querying and inferencing to uncover meaningful patterns in human behavior.

- *Activity* ($\mathcal{A}$): This class represents the human activities described in the datasets. It contains various activity types as subclasses, with instances corresponding to each occurrence of an activity distinguished by day and time.
- *ActivityTime* ($\mathcal{D}$): Represents the specific day and time of day when an activity takes place.
- *Location* ($\mathcal{L}$): Denotes the different rooms or spaces within the smart home where sensors are installed (e.g., Bedroom, Kitchen).
- *Sensor* ($\mathcal{S}$): Covers the range of sensor types deployed in the smart home environment, subdivided into:
    - *DoorSensor* ($\mathcal{S}_D$)
    - *MotionSensor* ($\mathcal{S}_M$)
    - *TemperatureSensor* ($\mathcal{S}_T$)

  Individuals in these subclasses correspond to the actual sensors recorded in the datasets.
- *Object* ($\mathcal{O}$): Represents various objects involved in the activities, typically categorized by the location where they are found (e.g., desktop, toothbrush).

- *Subject* ($\mathcal{S}_B$): Represents the subjects performing activities; in our datasets, this typically involves a single subject.

These classes and subclasses are linked through both object and data properties, examples of which include:

- *hasActor* ($hasActivity : \mathcal{A} \to \mathcal{S}_B$): Associates each activity occurrence with a specific *Subject.*
- *occursOn* ($occursOn : \mathcal{A} \to \mathcal{D}$): Connects each *Activity* to a particular day and time.
- *isMonitoredBy* ($isMonitoredBy : \mathcal{A} \to \mathcal{S}$): Links each *Activity* to one or more sensors monitoring it within the corresponding location.
- *hasActivityLocation* ($hasActivityLocation : \mathcal{A} \to \mathcal{L}$): Associates each activity occurrence with a specific *Location* inside the smart home.

To represent the temporal characteristics of activity occurrences, we define the following data properties: *hasStartTime* ($hasStartTime : \mathcal{A} \to xsd : dateTime$) indicates the start time of each activity instance, while *hasEndTime* ($hasEndTime : \mathcal{A}_O \to xsd : dateTime$) denotes its end time.

Here, XSD (XML Schema Definition), a recommendation by the World Wide Web Consortium, provides a formal specification for describing elements. The datatype `xsd:dateTime` from the XML Schema language defines time values, which can be conceptualized as objects comprising integer attributes for year, month, day, hour, and minute.

For each record in the datasets, we instantiate objects of the class *Activity* ($\mathcal{A}$) that encapsulate contextual information such as the day, location, start and end times, and related objects. Activity occurrences ($\mathcal{A}_O$) correspond directly to the activity labels provided in the datasets. Each such occurrence is modeled as an instance of a subclass of *Activity*, enriched with temporal attributes through the properties *hasStartTime*, *hasEndTime*, and spatial context via *hasActivityLocation.*

The sensors recorded in the datasets are organized into subclasses: *DoorSensor* ($\mathcal{S}_D$), *MotionSensor* ($\mathcal{S}_M$), and *TemperatureSensor* ($\mathcal{S}_T$). Each sensor is represented as an individual entity, characterized by properties such as *SensorID*, *Location*, and *SensorType.* Locations ($\mathcal{L}$) denote the various locations within the smart home, with each sensor linked to its corresponding location using the property *isInstalledIn.*

Each *Activity* type ($\mathcal{A}$) is connected to an *ActivityTime* instance ($\mathcal{D}$) and a *Location* ($\mathcal{L}$), and is linked to a *Subject* ($\mathcal{S}_B$) through the object properties *occursOn*, *hasActivityLocation*, and *hasActor*, respectively.

These formal relationships are defined as follows:

$$\forall a_0 \in \mathcal{A}, \exists d \in \mathcal{D}, \exists l \in \mathcal{L} \mid occursOn(a_0, d) \wedge hasActivityLocation(a_0, l) \tag{3.9}$$

$$\forall a_0 \in \mathcal{A}, \exists s \in \mathcal{S}_B \mid hasActor(a_0, s) \tag{3.10}$$

$$\forall a_o \in \mathcal{A} \mid hasStartTime(a_o, t_s) \wedge hasEndTime(a_o, t_e) \tag{3.11}$$

The dataset is transformed into RDF triples, which semantically capture the relationships among activities, their occurrences, and contextual elements within the smart home, following the structured ontology mapping. This representation supports the inference of activity patterns and behaviors, allowing for advanced querying through SPARQL. Figures 3.5 and 3.6 provide an overview of the HBOnt-Ontology taxonomy. The detailed inset emphasizes a portion of the Stimulus-Sensor-Observation pattern from the SSN ontology [198]. To accommodate sensor value representation within observations, the *hasValue* property was incorporated into the HBOnt ontology.

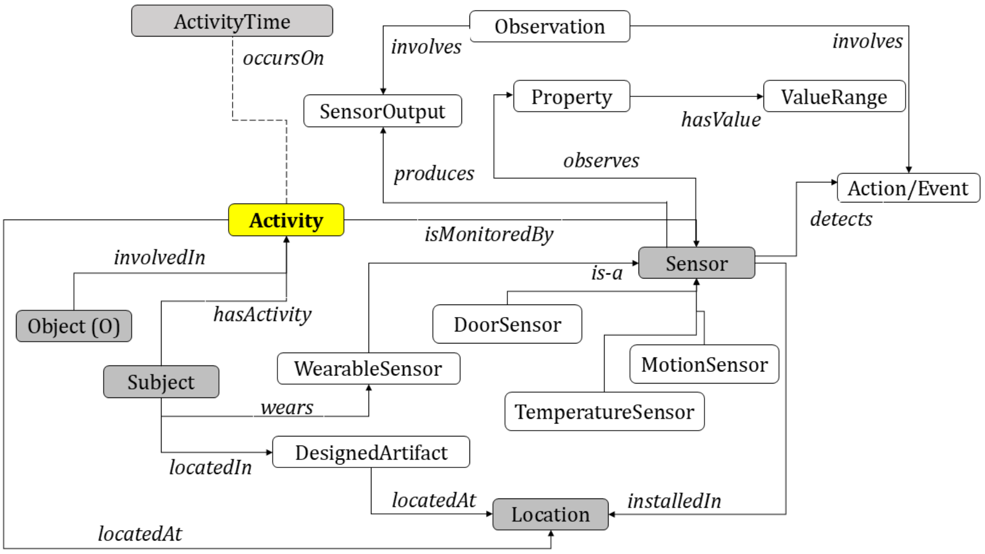
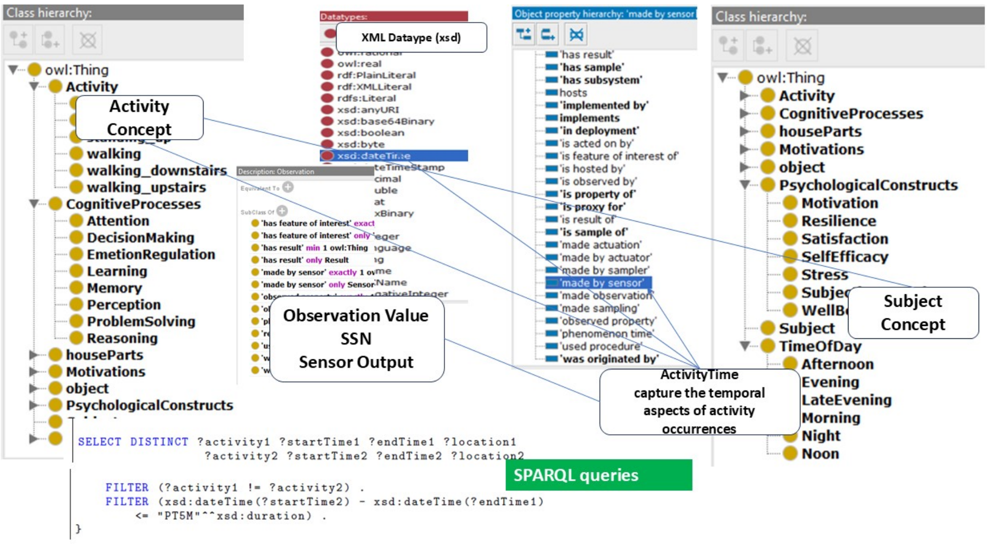


FIGURE 3.5: Main components of the HBOnt ontology to represent and organize knowledge about human activities.

The *locatedAt* relation denotes spatial proximity; for example, *locatedAt*(*Person*, *Place*, *Time Interval*) indicates that a subject is present at a specific location during a particular time interval. The *Sensor* concept encompasses two key properties: *observes*, which characterizes sensor attributes such as precision, resolution, and response time, and *produces*, which describes the sensor's output (e.g., temperature, brightness, position). Each temperature sensor (*TemperatureSensor*) possesses a unique identifier (Sensor ID), and the *locatedAt* property enables inferring a subject's presence at a location during a defined time interval.

The *Observation* concept captures contextual sensor data, linking to sensors via the *observed* property as defined in the SSN ontology. Our comprehensive taxonomy-driven semantic model incorporates essential contextual concepts including *locatedAt*, *installedIn*, and *observed.* Physical devices such as temperature and motion sensors are represented as *TemperatureSensor* and *MotionSensor* classes, respectively. Each sensor is uniquely identified by its Sensor ID, which supports precise inference of a subject's presence within a particular space.



```
SELECT DISTINCT ?activity1 ?startTime1 ?endTime1 ?location1
                ?activity2 ?startTime2 ?endTime2 ?location2

    FILTER (?activity1 != ?activity2) .
    FILTER (xsd:dateTime(?startTime2) - xsd:dateTime(?endTime1)
        <= "PT5M"^^xsd:duration) .
}
```

FIGURE 3.6: Structure of the ontology, adding semantic contextual information for each recorded activity occurrence, and modelling the relations of the concept Observation of the HBOnt ontology and the Subject and the Activity concepts.

3.4.1.2 Human Behavior Analysis

Our SPARQL queries are developed to investigate human behaviors and their interrelations within the cognitive psychology framework. These queries aim to extract pertinent information from the ontology to elucidate why individuals perform certain activities and to explore the interactions among activities, contexts, and motivations. Each query addresses a specific facet of human behavior, collectively contributing to a comprehensive analysis of cognitive processes involved in activity sequences.

Retrieve Activities of a Subject

```
SELECT DISTINCT ?activity ?startTime ?endTime
```

```
WHERE {
    <http://www.co-ode.org/ontologies/ont.owl#Mary> ex:hasActivity ?activity .
    ?activity ex:hasStartTime ?startTime ;
              ex:hasEndTime ?endTime .
}
```

This query retrieves specific tasks along with their start and end times for the subject. By identifying these behaviors, we gain insight into the exact steps an individual performed during a defined time interval. Analyzing the sequence and duration of these activities is crucial for understanding cognitive patterns and behavioral motivations.

Filter Activities by Duration

```
SELECT DISTINCT
    (STRAFTER(STR(?activity), "http://www.co-ode.org/ontologies/ont.owl#") AS ?activityName)
    ?startTime ?endTime
WHERE {
    {
        SELECT ?activity ?startTime ?endTime
        WHERE {
            <http://www.co-ode.org/ontologies/ont.owl#Mary> ex:hasActivity ?activity .
            ?activity ex:hasStartTime ?startTime ;
                      ex:hasEndTime ?endTime .
        }
    }
    FILTER ((xsd:dateTime(?endTime) - xsd:dateTime(?startTime)) > "PT6M"^^xsd:duration) .
}
```

Building on the previous query, this one filters activities to include only those lasting longer than a specified duration. Focusing on extended tasks allows for the examination of patterns that may reflect deeper cognitive engagement or personal preferences.

Identify Activity Sequences with Contextual Locations

```

SELECT DISTINCT ?activity1 ?startTime1 ?endTime1 ?location1
                 ?activity2 ?startTime2 ?endTime2 ?location2
```

```
WHERE {
    <http://www.co-ode.org/ontologies/ont.owl#Mary> ex:hasActivity ?
   activity1 .
    ?activity1 ex:hasStartTime ?startTime1 ;
               ex:hasEndTime ?endTime1 ;
               ont:hasActivityLocation ?location1 .

    <http://www.co-ode.org/ontologies/ont.owl#Mary> ex:hasActivity ?
   activity2 .
    ?activity2 ex:hasStartTime ?startTime2 ;
               ex:hasEndTime ?endTime2 ;
               ont:hasActivityLocation ?location2 .

    FILTER (?activity1 != ?activity2) .
    FILTER (xsd:dateTime(?startTime2) - xsd:dateTime(?endTime1) <= "PT5M
   "^^xsd:duration) .
}
```

This query retrieves the start and end times along with the locations of activity pairs performed by the subject 'Mary.' By analyzing the temporal proximity and contextual settings of these activities, we can infer potential relationships and underlying reasons for transitions between them.

The structured human behavioral ontology developed provides a practical approach for transforming raw sensor data into a well-organized model that represents temporal, spatial, and event-based relationships among various human activities. The use of SPARQL-based ontological queries enables the examination of specific facets of human behavior, contributing to a comprehensive understanding of cognitive processes and activity patterns. These queries facilitate the extraction of pertinent information from the ontology, aiding in the interpretation of why individuals perform certain activities and how different activities, contexts, and motivations interact.

3.5 Modeling Human Cognition in Smart Environments: A Hybrid Deep Learning and Ontology-Based Approach to Activity Motivation Inference

While Human Activity Recognition has made significant progress in identifying *what* activities occur from sensor data, a crucial challenge remains: understanding the deeper motivations or intentions behind those activities. This distinction is essential for building intelligent, adaptive, and truly human-centric AI systems, since identical actions may reflect very different needs, emotional states, or even emergencies. We propose a novel hybrid framework that overcomes this limitation by combining deep learning with ontological reasoning. This integration allows us to dynamically infer the diverse motivations underlying human behaviors in smart environments—marking an important step toward cognitive HAR.

For example, a system that merely detects "preparing a meal" cannot distinguish whether the user is cooking a routine family dinner, grabbing a late-night snack due to insomnia, or preparing medication. Although the physical activity may appear identical, the underlying purpose is drastically different. Capturing such nuances is essential for smart environments to offer personalized assistance, timely interventions, and meaningful automation that responds to actual user needs.

Over the last decade, HAR has become a vibrant research domain, fueled by the widespread adoption of smart devices and ambient sensors [199]. Many solutions have leveraged machine learning techniques—ranging from decision trees and support vector machines to advanced deep learning architectures like convolutional and recurrent neural networks [128], as well as more recent hybrid models [107, 200]. While these methods have achieved impressive accuracy in classifying human activities, they often lack transparency and struggle to incorporate contextual or intentional reasoning [201].

To address this, context-aware systems have attempted to enhance activity recognition by incorporating auxiliary information such as location, time of day, and environmental states [168]. However, such systems often treat context as an external feature rather than formalizing it into a structured reasoning framework. Ontology-based approaches bridge this gap by offering a semantic representation of human behavior, explicitly modeling relationships among entities and enabling logical inference through rule-based reasoning [144].

Previous research has explored hybrid systems that combine machine learning with ontological reasoning to improve the robustness and interpretability of HAR [202]. However, these efforts have typically focused on improving classification performance or system adaptability, rather than explicitly addressing the question of human intent. Our work advances this line of research by targeting motivation inference directly, drawing from cognitive psychology theories and supported by a formal knowledge representation framework. This allows our system to explain not only *what* activity was performed, but also *why* it was performed.

By integrating GENA—a general-purpose semantic reasoning engine—with deep activity recognition models, we introduce a practical and powerful pathway for embedding cognition-inspired logic into HAR pipelines. Additionally, our inference rules are crafted based on real-world patterns of human behavior, enriched with contextual information such as health conditions, temporal cues, and object interactions. This enables our system to move beyond simple classification toward a deeper understanding of human intentions and needs.

3.5.1 Theoretical Foundations

Understanding the motivations behind human actions is key to pushing HAR beyond mere behavior classification. Traditional HAR systems are adept at identifying observable activities, but they often fall short when it comes to interpreting the deeper reasons or goals that drive those actions. To bridge this gap, our approach is grounded in well-established theories from cognitive psychology [150] and enhanced with a semantic ontology specifically designed to capture context and motivation.

Human behavior is inherently goal-driven. Cognitive psychology suggests that people rarely act without reason; most actions are taken to fulfill specific needs, desires, or intentions [203]. Theories such as the Theory of Planned Behavior and goal-setting theory emphasize that both internal factors—like health status, emotional state, or fatigue—and external stimuli—such as location, time of day, or social context—combine to influence how, when, and why people act.

Take the example of walking: a person might walk for exercise, to answer a ringing phone, or to join a conversation. Even though the physical action is the same, the *meaning* and *purpose* behind it are entirely different. These subtle but important distinctions often go unnoticed by deep learning models, which focus strictly on patterns in sensor data. To uncover the motivations behind actions, we need a semantic layer that can represent human goals, contextual settings, and individual characteristics—all of which shape behavior in meaningful ways.

3.5.2 Proposed hybrid deep learning and ontology-based approach to activity motivation inference

In this section, we present our proposed framework describing its main components.

Our proposed framework integrates a deep learning-based HAR pipeline with a semantic reasoning module grounded in a motivation-aware ontology. The architecture, illustrated in Figure 3.7, is composed of three principal components:

1. A **Deep Learning Module** responsible for recognizing low-level activities from sensor data.
2. A **Motivation Ontology** that models contextual, environmental, and cognitive aspects of human behavior.

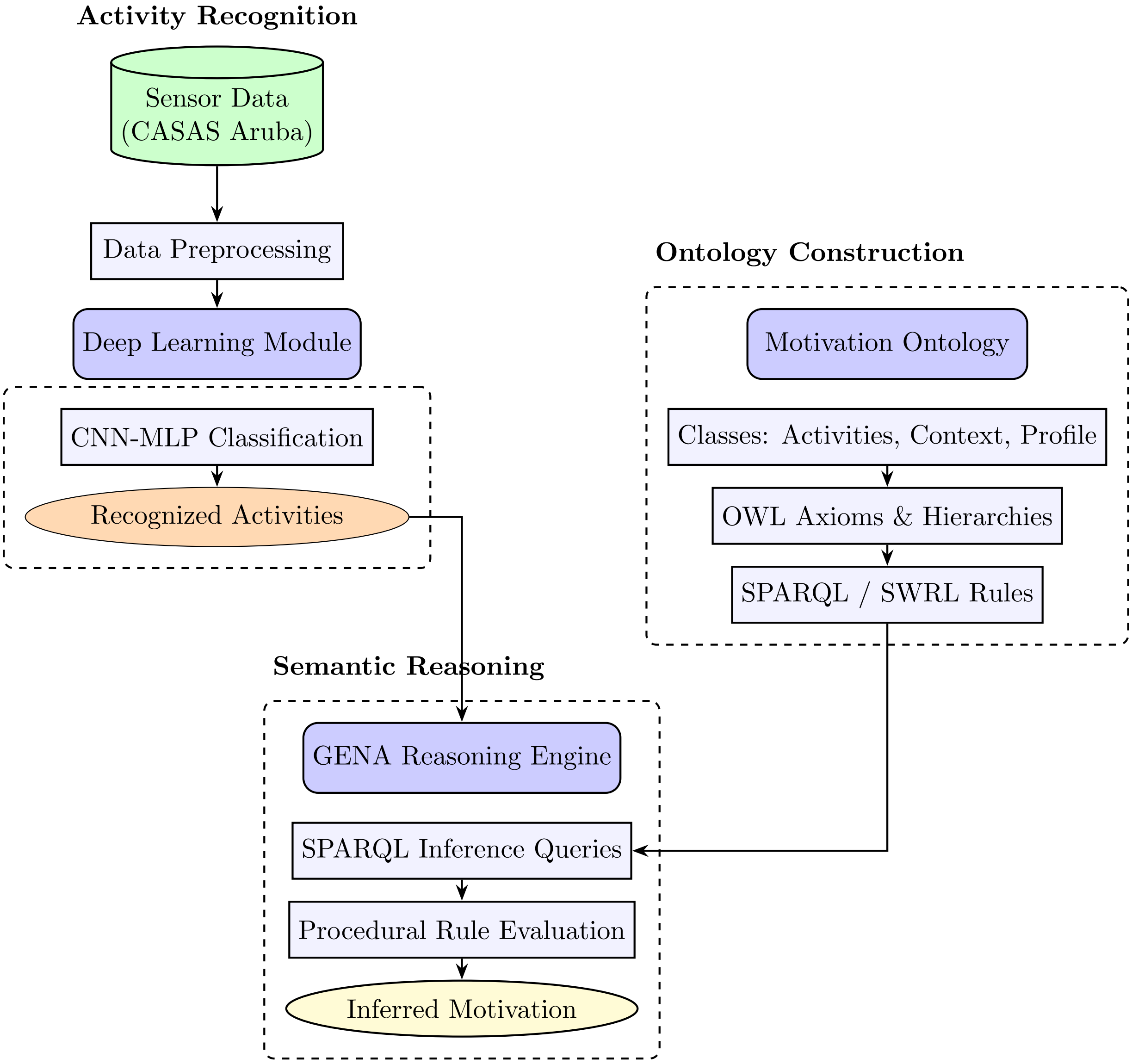


Figure 3.7: Architecture for motivation-aware human activity recognition using deep learning and ontology-based reasoning.

3. A **Semantic Reasoning Engine** powered by GENA, used to infer the underlying purpose or motivation behind detected activities.

3.5.2.1 Deep Learning Module

The HAR module utilizes a hybrid Convolutional Neural Network–Multi-Layer Perceptron (CNN-MLP) model [200], trained on the CASAS Aruba dataset [61].

This model achieves high classification accuracy and forms the first step in our hybrid framework, translating raw sensor events into semantically recognized activities. The architecture of our proposed CNN-MLP module is detailed in section 3.3.

3.5.2.2 Ontological Module

To enable motivation-aware reasoning, we developed an ontology that formalizes the concepts necessary to link observed activities with plausible human intentions. The ontology functions as a shared vocabulary and inference framework, enabling high-level semantic interpretation of low-level sensor outputs.

The core classes and properties modeled in the ontology include:

- **RecognizedActivity:** Captures atomic or composite human activities (e.g., *Walking*, *Cooking*, *Sleeping*) as detected by a HAR system.
- **TimeContext:** Encodes temporal dimensions (e.g., *Morning*, *Mealtime*, *Night*) to facilitate reasoning about routine behaviors or time-sensitive activities.
- **LocationContext:** Describes the physical setting in which an activity takes place (e.g., *Kitchen*, *Bedroom*, *Living Room*), helping disambiguate behavior.
- **ObjectInteraction:** Represents interactions with physical objects (e.g., *Refrigerator*, *Phone*, *MedicineBox*), which provide cues for inferring the intent behind the activity [202].
- **HumanProfile:** Encodes personal characteristics of the observed individual (e.g., chronic conditions, age group, cognitive state), which influence behavior interpretation.
- **InferredPurpose:** Represents semantically derived motivations or goals based on the combination of recognized activity, context, and personal traits. Examples include:
  - *ManageHealthCondition*
  - *RespondToCallOrVisitor*
  - *SocialInteraction*
  - *PrepareMeal*
  - *RestOrSleep*

These concepts are connected through object properties such as *hasTimeContext*, *hasLocation*, *involvesObject*, and *performedBy*. Reasoning is carried out using custom semantic rules, expressed in SWRL-like syntax [204], and implemented within the GENA framework [205].

The ontology aligns with best practices in knowledge representation by following modular, reusable design principles and by partially reusing concepts from foundational ontologies such as DOLCE [206] and FOAF [207] where applicable. This structure supports a richer understanding of human activities, moving beyond surface-level classification to deeper cognitive modeling grounded in context and purpose. The main concepts of our ontology are presented in Table 3.1.

The semantic layer is operationalized through the GENA inference engine, which supports both SPARQL-based querying and custom semantic rule execution. Motivational inferences are triggered by matching ontological assertions with rule-based conditions.

Table 3.1: Structure of the proposed ontology for modeling human activity motivations, showing key classes, their instances or subclasses, and the associated object and data properties used to represent semantic relationships between activities, context, and human factors.

Class	Subclass/Instance	Object Property	Data Property
Activity	RecognizedActivity	canTakePlaceDuring	hasCurrentTime
	Purpose	hasActor	hasDuration
	Motivation	hasLocation	hasTimeGranularity
	Interaction	hasArtifact	*N/A*
Location	LivingRoom	contains	isInside
	Kitchen	isUpon	measuredTemperature
	Office	locatedIn	isWearing
	Bedroom	isInside	*N/A*
Object	RemoteControl	isUsedIn	*N/A*
	Book	isManipulatedBy	*N/A*
	Phone	isUsedIn	*N/A*
Human Profile	HealthCondition	hasInteractionType	*N/A*
	Occupation	canPerformActivity	*N/A*
Time Context	Mealtime	*N/A*	*N/A*
	LeisureTime	*N/A*	*N/A*

These rules are either encoded using SPARQL CONSTRUCT queries or implemented in the form of procedural logic within the GENA reasoning framework.

3.5.2.3 Inference of Human Purpose Based on Recognized Activities

Understanding *why* an individual engages in a specific activity provides a deeper layer of human behavior modeling that extends beyond simple recognition. We incorporate semantic reasoning to infer the purpose behind detected activities using predefined rules based on context, time, location, and user profile information. These rules are grounded in commonsense knowledge and cognitive psychology principles and are implemented using ontology-based reasoning and SPARQL rule-based inference. Below, we detail some rules used for inferring human purpose with high or medium confidence, along with the motivations behind each inference.

This pair of rules showcases how the same activities related to consumption can have drastically different motivations, highlighting the system's ability to discern routine behavior from critical health monitoring.

Routine Consumption Purpose Inference Rule

IF Recognized_Activity is (Sitting/Standing)
AND Current_Location is_in (Kitchen/Dining_Room)
AND Object_Interaction indicates (Using_Cutlery/Picking_Up_Glass)
AND Time_Context is (Morning_Mealtime/Evening_Mealtime)
THEN Infer Purpose is **Eat/Drink WITH High Confidence**.
Motivation: Alignment of activity, context, and environment with routine food consumption behavior during typical meal hours.

Health-Related Consumption Purpose Inference Rule

IF Recognized_Activity is (Sitting/Standing)
AND Current_Location is_in (Kitchen/Bedroom)
AND Object_Interaction indicates (Using_Pill_Bottle/Measuring_Spoon)
AND Time_Context is (Not_Mealtime/Unusual_Time)
AND Human_Profile includes (KnownMedicalCondition:Chronic_Disease)
THEN Infer Purpose is **Medication_Administration/Health_Management WITH High Confidence**.
Motivation: The specific object interaction combined with the occupant's health profile at an atypical time points to a non-routine, health-critical consumption event.

These rules demonstrate the critical ability to distinguish between a casual act of eating or drinking and a necessary health-related consumption. For smart home systems, this means differentiating between simply providing meal reminders versus alerting a caregiver if a medication dosage is missed or taken at an irregular time, directly contributing to proactive health management and safety.

Another set of rules illustrates how an activity like "lying down" can be interpreted very differently, enabling the system to identify potential distress rather than just leisure.

Routine Relaxation/Entertainment Purpose Inference Rule

IF Recognized_Activity is (Sitting/Lying_Down)
AND Current_Location is_in (Living_Room/Bedroom)
AND Object_Interaction indicates (Using_Remote_Control/Book/Phone)
AND Device_State is (TV_On/Music_Playing)
THEN Infer Purpose is **Relax/Entertainment WITH High Confidence**.
Motivation: Typical passive behavior combined with entertainment-related interactions during leisure hours.

Prolonged Inactivity/Potential Distress Purpose Inference Rule

IF Recognized_Activity is (Lying_Down)
AND Current_Location is_in (Living_Room/Any_Non_Bedroom_Location)
AND Time_Context is (Daytime)
AND No_Movement_Detected_For_Duration is (Longer_than_X_minutes)
AND Object_Interaction indicates (No_Interaction_with_any_device_or_object)
AND Human_Profile indicates (Age_Group:Elderly / MedicalCondition:Frailty)
THEN Infer Purpose is **Prolonged_Inactivity/Potential_Distress WITH High Confidence**.
Motivation: Unexplained prolonged inactivity, especially in a non-sleep area during active hours and for a vulnerable individual, strongly suggests a potential issue requiring attention.

This rule pair is crucial for distinguishing between deliberate rest or entertainment and a potentially problematic state. It allows a smart home to avoid false alarms during routine relaxation while providing timely alerts when an occupant is unresponsive or in distress, significantly enhancing the safety and reliability of assistive technology by focusing on true deviations from healthy patterns.

These rules are applied during the reasoning phase after activity recognition and context gathering. In the implementation, relevant properties and instances from the ontology are queried using SPARQL, and the resulting inferences are used to determine the likely human purpose behind activity sequences. This high-level semantic interpretation facilitates enhanced human behavior modeling and can support applications in healthcare, smart environments, and cognitive assistance systems.

3.6 Conclusion

This chapter has comprehensively outlined the key contributions of this research, highlighting the novel advancements made in the semantic modeling and analysis of human activity recognition within smart home environments. By developing a structured human behavioral ontology and integrating it with sensor data, this work provides a robust framework that enhances the understanding of complex human behaviors through formal semantic representation. The design and implementation of tailored SPARQL queries further demonstrate the capability of the ontology to support sophisticated reasoning and cognitive analysis.

Moreover, the integration of real-world datasets such as CASAS Aruba and Orange4Home validates the practical applicability and effectiveness of the proposed framework in capturing temporal, spatial, and contextual nuances of human activities. These contributions not only bridge the gap between raw sensor data and high-level behavioral interpretation but also lay a strong foundation for future advancements, including the incorporation of deep learning methods and personalized behavior modeling.

The research contributions detailed in this chapter represent a significant step toward intelligent, context-aware smart home systems that can support personalized health, well-being, and quality of life improvements. The groundwork established here opens numerous avenues for ongoing research and innovation in activity recognition, semantic modeling, and cognitive behavior analysis. The next chapter presents the implementation and experimental results of the proposed contributions.

Chapter 4

Evaluation and Experimental results

4.1 Introduction

This chapter presents the implementation details and experimental evaluation of the proposed hybrid system. The system combines a Convolutional Neural Network–MultiLayer Perceptron (CNN-MLP) architecture for automatic feature extraction from raw sensor data with ontology-based reasoning to enhance semantic interpretability and contextual understanding of the human behavior.

The performance of the system is evaluated using several benchmark datasets, including both smartphone-based and smart home sensor environments. This diversity enables a comprehensive assessment of the system's accuracy, robustness, and generalizability across different settings.

In addition, the chapter outlines how integrating data-driven learning methods with semantic reasoning can support more structured interpretations of human behavior, contributing to more informed activity analysis.

4.2 Proposed CNN-MLP deep model for sensor based human activity recognition

During the model architecture development, the Keras library in Python was employed as a high-level neural network API, which provides compatibility with both TensorFlow and Theano backends. For this work, TensorFlow was chosen as the backend owing to its capability to efficiently handle accelerated computations and leverage GPU resources for optimized performance. All simulations were executed on a computing system running Microsoft Windows 10,

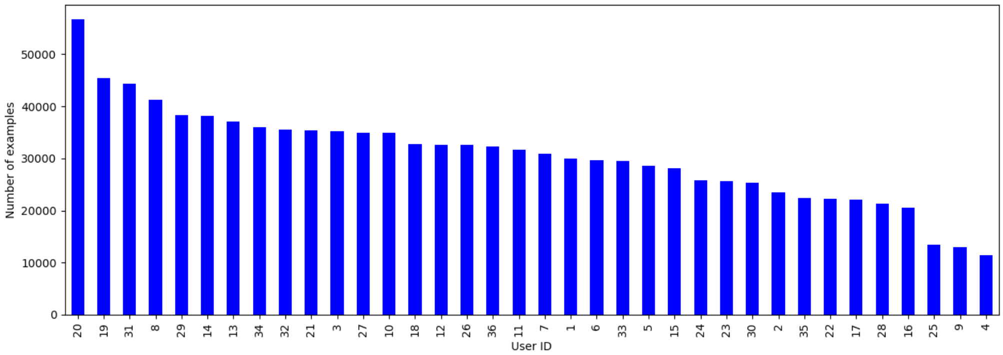


FIGURE 4.1: Distribution of the examples by activity types in WISDM dataset.

configured with an Intel Core i5 processor operating at 1.90 GHz, 8 GB of RAM, and a 64-bit architecture.

4.2.1 Datasets used for CNN-MLP Proposed Model

To evaluate the proposed method, we use four widely-adopted human activity recognition datasets: UCI HAR [208], WISDM [209], PAMAP2 [210], and CASAS Aruba [61]. These datasets represent two major sensor categories—smartphone-based (UCI HAR, WISDM, PAMAP2) and ambient smart home-based (CASAS Aruba)—enabling a broad evaluation of the model's adaptability across diverse sensor modalities and real-world environments.

4.2.1.1 WISDM Dataset

The complete WISDM dataset consists of 1,098,209 recorded instances encompassing six activity categories: *WALKING*, *JOGGING*, *STANDING*, *SITTING*, *WALKING UPSTAIRS*, and *WALKING DOWNSTAIRS*. As depicted in Figure 4.2, the dataset exhibits a class imbalance, with *WALKING* comprising 38.6% of the total samples, while *STANDING* represents only 4.4%. Data acquisition involved thirty-six individuals, each carrying an Android smartphone positioned in the front pocket of their leg during the execution of daily activities. The primary sensor used was a triaxial accelerometer, sampling at a frequency of 20 Hz. Furthermore, the smartphone was equipped with an embedded motion sensor. Figure 4.1 illustrates the distribution of training samples across different user_ID values.

4.2.1.2 UCI HAR Dataset

The UCI Human Activity Recognition (HAR) dataset comprises sensor readings obtained from thirty participants aged between 19 and 48. Each subject was instructed to perform a series

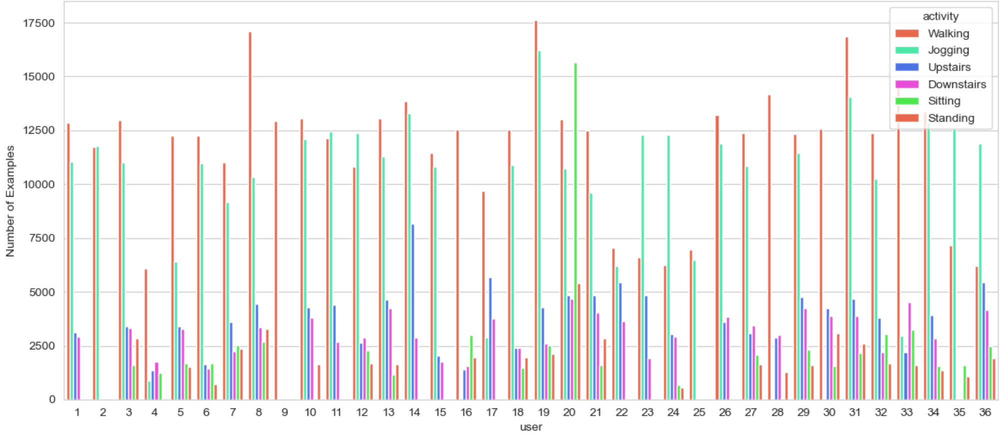


Figure 4.2: Distribution of training examples by user_ID in WISDM dataset.

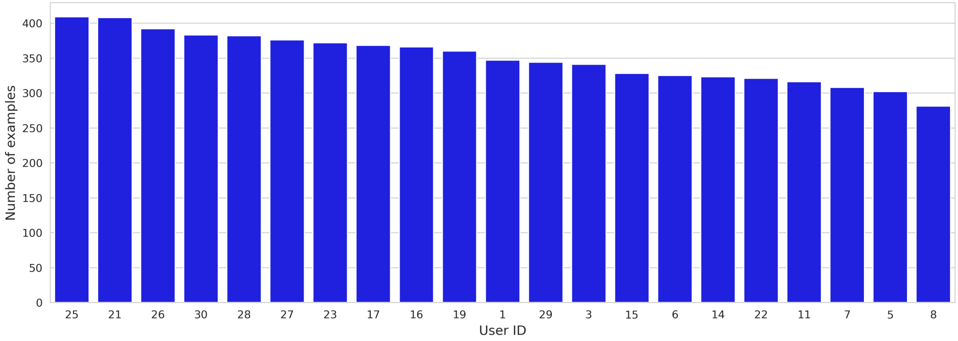


Figure 4.3: Distribution of Examples by User ID in the UCI HAR Dataset.

of predefined physical activities while carrying a Samsung Galaxy S II smartphone affixed to their waist. This device, equipped with inertial measurement units, recorded data corresponding to six primary activities: *WALKING*, *WALKING_UPSTAIRS*, *WALKING_DOWNSTAIRS*, *SITTING*, *STANDING*, and *LAYING*. Although the dataset also includes transitional activities—such as transitions from standing to sitting or from lying to standing—these instances were excluded from the present analysis due to their limited sample size.

The activity annotations were manually derived by synchronizing the sensor data with video recordings, thereby ensuring accurate ground truth labeling. The recorded signals include three-axis accelerometer and gyroscope data, sampled at a uniform frequency of 50 Hz. In total, the dataset contains 748,406 time-stamped data points. Owing to its structured collection protocol and controlled conditions, this dataset serves as a robust benchmark for evaluating smartphone-based activity recognition methodologies. Figure 4.3 presents the distribution of samples by subject ID, while Figure 4.4 illustrates the proportion of activity types recorded for each user in the UCI HAR dataset.

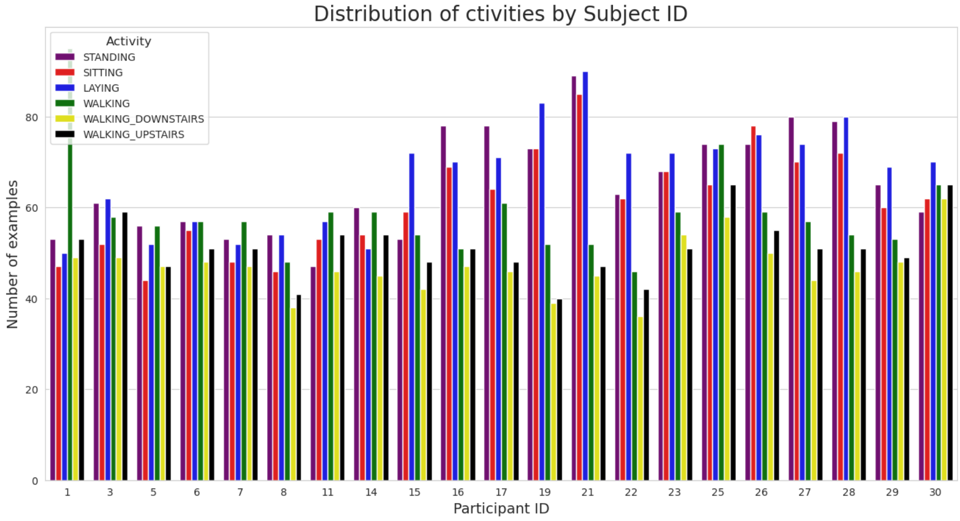


FIGURE 4.4: Distribution of Examples by User ID and Activity Type in the UCI HAR Dataset

4.2.1.3 PAMAP2 Dataset

The PAMAP2 dataset contains sensor data collected from nine individuals engaged in twelve different activities of daily living, encompassing both household tasks and physical exercises, including soccer and Nordic walking. The data acquisition spanned approximately ten hours and employed Inertial Measurement Units (IMUs) affixed to the participants' hand, chest, and ankle. These IMUs captured a wide range of signals, including those from accelerometers, gyroscopes, magnetometers, as well as temperature sensors and heart rate monitors. The heterogeneous nature of the sensor modalities in this dataset renders it particularly valuable for investigating the role and effectiveness of various sensor types in human activity recognition. Figure 4.5 displays the sample distribution across subject identifiers, while Figure 4.6 shows a heatmap representing the correlation matrix, which emphasizes the statistical interdependencies among the recorded sensor signals.

This analysis also yields several applied observations. For instance, a notable correlation emerges between the hand-worn accelerometer and the temperature sensor data, consistently observed across the three accelerometers positioned on the hand. This finding indicates the potential utility of accelerometer data for approximating hand temperature in certain contexts. In addition, the magnetometer measurements obtained from the chest demonstrate a strong association with heart rate values, a plausible result given the close anatomical placement of the sensors. These patterns may be leveraged to enhance the reliability of heart rate estimation methodologies.

4.2.1.4 CASAS-Aruba Dataset

In the domain of smart home research—particularly in activity recognition and ambient intelligence—the CASAS Aruba dataset is a well-established and extensively used benchmark.

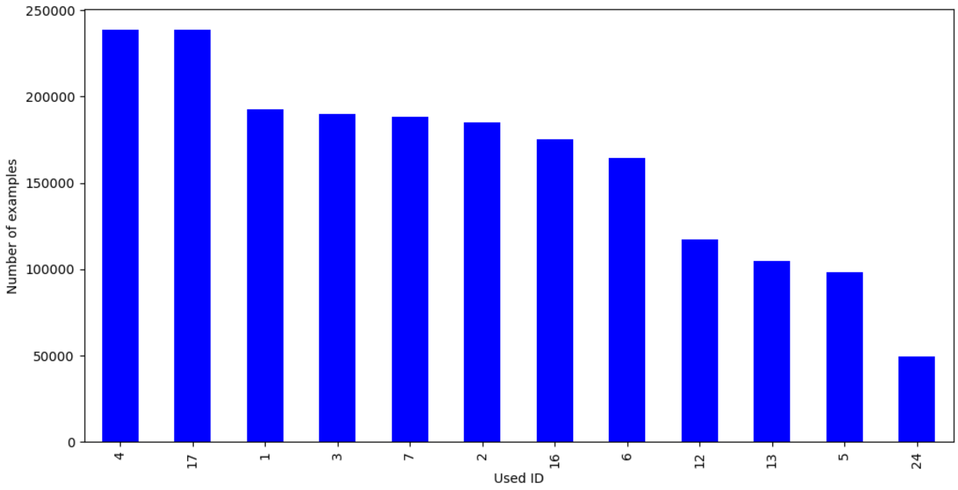


FIGURE 4.5: Distribution of the examples by user ID in PAMAP2 dataset.

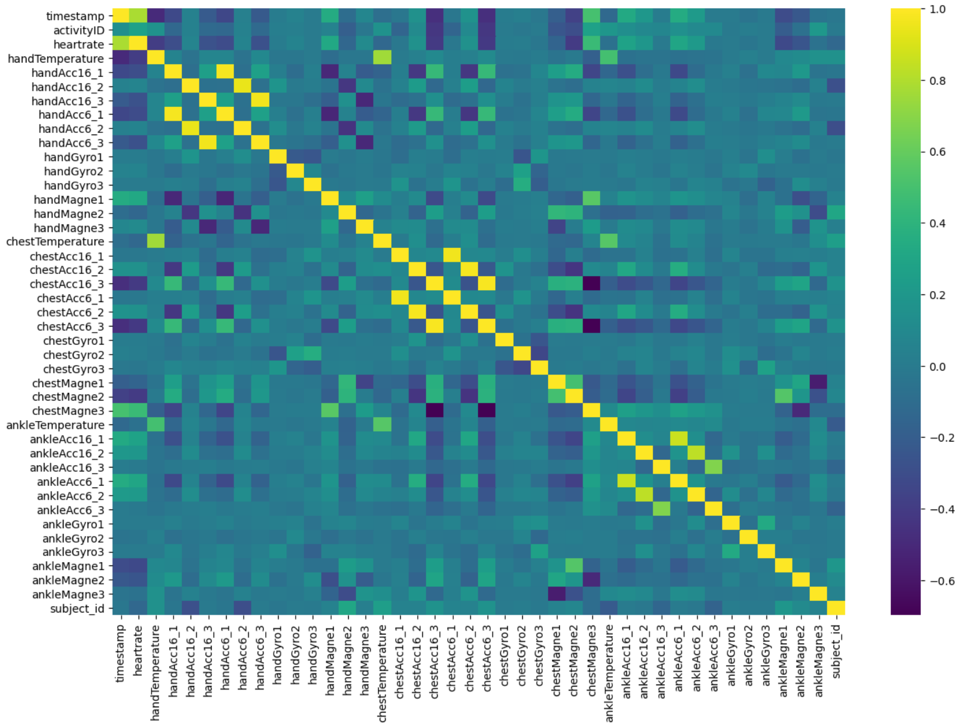


FIGURE 4.6: Heatmap in PAMAP2 dataset.

It forms part of a broader dataset collection curated by the Center for Advanced Studies in Adaptive Systems (CASAS) at Washington State University, offering researchers and system developers access to realistic sensor-based activity data.

The Aruba dataset was collected over a span of seven months from the home of a single adult female occupant, capturing her routine daily activities along with occasional interactions involving visiting family members. It encompasses annotated data for eleven discrete activities: *Meal_Preparation*, *Relax*, *Eating*, *Work*, *Sleeping*, *Wash_Dishes*, *Bed_To_Toilet*, *Enter_Home*, *Leave_Home*, *Housekeeping*, *Respirate*, and *Other*. The *Other* category corresponds to ambient sensor events that occurred outside the scope of the predefined activity labels. As is characteristic of naturalistic environments, the dataset presents a substantial class imbalance, with approximately 50% of instances categorized as *Other*, representing periods with unannotated or ambiguous activity data.

The smart environment included a network of forty sensors, comprising four binary door sensors (D001–D004), five temperature sensors (T001–T005), and thirty-one passive infrared (PIR) motion detectors labeled M001 through M031. This diverse sensor configuration facilitates the evaluation of model performance in real-world ambient intelligence contexts, where data sparsity and sensor heterogeneity are significant challenges. Figure 4.7 illustrates a representative sample of the Aruba dataset, accompanied by a view of the dataset's column structure as described in Table 4.1.

A summary of the annotated activities, including their respective labels, identifiers, and the frequency of their occurrence, is provided in Table 4.2.

```
2011-05-08 19:07:36.51009 T002 24
2011-05-08 19:07:36.640848 T004 22.5
2011-05-08 19:12:39.810776 T005 24
2011-05-08 19:12:39.948323 T003 23
2011-05-08 19:12:40.098513 T002 23
2011-05-08 19:15:30.102992 D004 OPEN Enter_Home begin
2011-05-08 19:15:30.756739 M030 ON
2011-05-08 19:15:34.794553 M030 OFF
2011-05-08 19:15:35.649254 D004 CLOSE Enter_Home end
2011-05-08 19:15:36.935262 M022 ON
2011-05-08 19:15:40.070464 M022 OFF
2011-05-08 19:15:45.57804 M019 ON
2011-05-08 19:15:46.35326 M018 ON
2011-05-08 19:15:46.971172 M018 OFF
```

Figure 4.7: Excerpt from the CASAS Aruba dataset illustrating raw sensor event logs, including timestamps, sensor IDs, sensor values, and annotated activity labels.

Table 4.1: Sample data from the CASAS Aruba dataset

Date	**Time**	**Sensor**	**Value**	**Activity**	**Log**
2010-11-04	00:03:50.209589	M003	ON	Sleeping	begin
2010-11-04	00:03:57.399391	M003	OFF	Sleeping	begin
2010-11-04	00:15:08.984841	T002	21.5	Sleeping	begin
2010-11-04	00:30:19.185547	T003	21	Sleeping	begin
2010-11-04	00:30:19.385336	T004	21	Sleeping	begin

Table 4.2: Distribution of activity instances in the CASAS Aruba dataset, showing the total number of recorded events per activity label.

id	Activity	Number of Events
1	BedtoToilet	1330
2	Eating	16037
3	EnterHome	2018
4	Housekeeping	10583
5	LeaveHome	1922
6	MealPreparation	285149
7	Relax	354585
8	Resperate	542
9	Sleeping	32682
10	WashDishes	10464
11	Work	16321
12	Otheractivity	871320

Table 4.3 presents a summary of essential statistics for the different activities recorded in the dataset. It includes details such as the type of activity, the count of occurrences for each activity, the average duration, and the sensors used to collect the related data.

Table 4.3: Overview of annotated activities in the CASAS Aruba dataset, including the number of instances, average duration per activity, and the types of sensors involved in their detection.

Activity	**Number of Instances**	**Avg. Duration (minutes)**	**Sensors Involved**
Bed to Toilet	150	5	Motion, Door
Cook	300	45	Motion, Temperature
Eat	200	30	Motion
Enter Home	100	1	Door
Leave Home	100	1	Door
Personal Hygiene	250	20	Motion, Light
Relax	400	60	Motion
Sleep	500	480	Motion
Work	350	120	Motion, Light
Total	**2350**	/	/

4.2.2 Data Preprocessing

This section outlines the preprocessing steps applied to our datasets to ready the data for model training.

4.2.2.1 Smartphone-based Datasets

Robust data preprocessing is a critical prerequisite for ensuring that the model can effectively interpret and learn from raw sensor data. In the case of the WISDM dataset, a sequence of preprocessing steps was applied, including label encoding, linear interpolation, normalization, segmentation, and one-hot encoding.

To satisfy the model's requirement for numerical input, the categorical activity labels located in the `activity` column were encoded into a new numeric column named `activityLabel`. This transformation involved assigning each activity a unique integer code, as detailed in Table 4.4.

Table 4.4: Mapping of activity classes to numerical labels used for classification in smartphone-based human activity recognition datasets.

Activity	Label
Downstairs	0
Jogging	1
Sitting	2
Standing	3
Upstairs	4
Walking	5

To handle missing data entries (NaN), linear interpolation is utilized as a technique to estimate intermediate values based on neighboring observations. Although the dataset contains relatively few missing values, this approach contributes to constructing a more continuous and complete dataset suitable for model training.

Subsequently, feature normalization is applied to rescale the input variables of the training set to a standardized range between 0 and 1. This transformation is performed using the following normalization equation:

$$Y^{(i)}_{\text{normalized}} = \frac{Y^{(i)} - Y^{(i)}_{\text{min}}}{Y^{(i)}_{\text{max}} - Y^{(i)}_{\text{min}}} \quad \text{for} \quad i = 1, 2, \ldots, n \tag{4.1}$$

In this equation, n denotes the number of channels, while Y_{max} and Y_{min} correspond to the maximum and minimum values of the i-th channel, respectively.

To prepare the data for model input, the dataset is reshaped so that each sample has a fixed length of 80 time steps. We define a function named *segments* which processes the dataframe along with label names and three input parameters. Subsequently, the *x_train* and *y_train* functions separate the features from the labels.

Important variables such as *time_period*, *sensors*, and *num_classes* are retained for later use. The *reshape* function converts the two-dimensional data into a list format suitable for the model, and all data points are cast to the *float32* data type to ensure model compatibility.

As the final step in the preprocessing pipeline, one-hot encoding is employed to transform categorical class labels into a binary vector representation. In this format, each label is represented by a vector where a single element corresponding to the class is assigned a value of 1 (indicating the active class), while all remaining elements are set to 0. The dimensionality of the resulting vector equals the total number of distinct classes. Within our workflow, the one-hot encoded labels are stored in the variable *y_train_hot*, thereby completing the data preprocessing phase.

The UCI HAR dataset also underwent a comprehensive preprocessing phase. Initially, the dataset was inspected for duplicate entries and missing values; however, neither duplicates nor null entries were identified. The training set was extracted from the files `X_train.txt` and `subject_train.txt`, while the testing set was obtained from `X_test.txt` and `subject_test.txt`. Both subsets were subsequently exported and stored as individual CSV files. Additionally, the feature names were sanitized to eliminate any superfluous or non-standard characters.

For the PAMAP2 dataset, preprocessing commenced with the importation of data from multiple files, each representing a specific activity. Each file was loaded into a Pandas `DataFrame`, where sensor readings and associated activity labels were consolidated into structured columns. These were then combined with corresponding metadata columns—such as `timestamp`, `activityID`, and `heartrate`—to produce a well-organized dataset. This process involved the assembly of relevant column headers into a predefined list used during the data-loading procedure.

Subsequent to data importation, a multi-step cleaning process was implemented. First, orientation-related features recorded by body-worn IMUs were excluded, as they were not required for the intended analysis. Second, instances corresponding to transitional or undefined activities (e.g., those labeled with activity ID 0) were removed, as they do not contribute to the primary recognition task. Third, all non-numeric fields were converted into numeric format to ensure consistency across all variables. Finally, any residual missing values were addressed through interpolation techniques to reconstruct continuous and complete data sequences.

Upon completion of the preprocessing pipeline, the resulting datasets were deemed ready for subsequent modeling and analysis. Emphasis was placed on maintaining data integrity and consistency throughout the process to support accurate activity recognition.

Following the data-cleaning steps, an analysis of activity label distributions was conducted to assess the diversity and frequency of recorded activities across different subjects.

For each of the three datasets, the data was partitioned into training and testing subsets using an 80/20 split ratio. Specifically, 80% of the samples were allocated to model training, while the remaining 20% were reserved for testing. This standard partitioning approach facilitates reliable assessment of the model's generalization capability on previously unseen data.

4.2.2.2 CASAS Aruba dataset

The raw dataset is initially structured with the columns `date`, `time`, `sensor`, `value`, `activity`, and `log`, establishing a consistent framework for downstream analysis. To ensure temporal relevance, rows are filtered based on specific date thresholds, enabling the inclusion or exclusion of entries captured before or after designated time periods.

Activity labels are refined to enhance clarity and reliability. Missing values in the `activity` column are imputed using corresponding entries from the `log` column. The string "end" and any residual missing values are replaced with "Other," while instances of "begin" are substituted with "None," standardizing the label set for accurate representation.

To uphold data integrity, the `value` column is inspected for inconsistencies and erroneous entries. Corrections are implemented by mapping specific values to binary representations, thereby improving uniformity. Additionally, temperature sensor data is excluded by removing rows associated with sensor identifiers starting with "C" or "T," thus concentrating the analysis on motion-relevant sensors.

Categorical attributes such as `sensor_id` and `activity` are encoded numerically through label encoding, converting textual data into a format amenable to machine learning models. The function *markdown_activities* is utilized to eliminate rows corresponding to irrelevant or extraneous activity labels, ensuring the dataset is focused on the primary activities of interest.

In the final preprocessing stage, the data is segmented into overlapping sliding windows, facilitating temporal feature extraction. Each window aggregates sensor readings across a fixed interval, enabling the capture of time-dependent patterns and trends. Frequency statistics of sensor activations within each window are computed to detect anomalies and assess temporal dynamics. Additionally, mutual information is calculated between sensor features and activity labels to evaluate the predictive relevance of individual sensors.

Upon completion of feature extraction and weighting, the dataset is reorganized into input-target pairs consisting of sliding window features and their corresponding activity labels. This structured format supports efficient training and testing of activity recognition models.

The inherent variability introduced by sensor type, placement, and signal properties presents challenges in sensor-based Human Activity Recognition (HAR). This work evaluates four heterogeneous datasets—UCI HAR, WISDM, PAMAP2, and CASAS Aruba—each differing in sensing modalities and environmental context. These datasets predominantly utilize accelerometers and gyroscopes to record motion signals. Key challenges include variation in sensor placement (e.g., waist, pocket, wrist), inconsistencies in sampling frequency, and signal noise due to natural movement artifacts. Convolutional Neural Network (CNN) layers are employed to extract spatial features from raw sensor inputs, capturing motion-specific signatures, while Multilayer Perceptron (MLP) layers model complex, nonlinear relationships between features, thereby improving classification performance. Batch normalization is integrated to reduce the impact of hardware differences and sampling irregularities, thus enhancing generalization across devices.

Unlike wearable-based HAR, the CASAS Aruba dataset relies on ambient sensors such as motion detectors, door sensors, and temperature sensors. Whereas smartphone-based HAR identifies activities through motion intensities, smart home HAR depends on sequential activation patterns across environmental sensors. To enable compatibility with deep learning models, event-driven sensor data is transformed into fixed-length time-series windows, allowing CNN layers to capture spatial dependencies and MLP layers to incorporate contextual information.

Recognition accuracy is influenced by discrepancies in sensor precision and deployment. Smartphone sensors may vary in sensitivity and sampling rates across devices, while ambient sensors in smart homes can yield missing values due to latency or missed activations. To mitigate such issues, preprocessing methods such as normalization and resampling are applied consistently across datasets.

4.2.3 Evaluation metrics

To assess model performance, we employed classification reports in conjunction with the confusion matrix. A stratified 10-fold cross-validation approach was adopted to ensure robust and balanced evaluation across all activity classes.

The confusion matrix offers a detailed overview of classification performance by comparing predicted labels against actual ground truth values for each category. This matrix provides critical insights into the model's strengths and limitations across different classes. It is composed of four fundamental elements:

- **True Positive (TP):** Represents the count of positive records that were correctly identified as positive by the model.
- **True Negative (TN):** Indicates the number of negative records that were correctly identified as negative by the model.

- **False Positive (FP):** Refers to the count of negative records that were incorrectly predicted as positive by the model.
- **False Negative (FN):** Represents the number of positive records that were incorrectly predicted as negative by the model.

Table 4.5 illustrates the standard structure of a confusion matrix, serving as a generic representation of its components.

The classification report encompasses key evaluation metrics, including accuracy, precision, recall, and F1 score, which collectively provide a holistic measure of model performance. These metrics enable rigorous comparison across various classification approaches by offering detailed insight into the effectiveness of each model.

- **Precision**: Measures the percentage of correctly identified positive cases out of all cases predicted as positive. It is calculated as:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4.2}$$

- **Recall**: Represents the proportion of actual positive cases that the model correctly identified. It is calculated as:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4.3}$$

- **F1 Score**: The harmonic mean of precision and recall, given by:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.4}$$

- **Accuracy**: A commonly used metric for evaluating overall model performance, reflecting the proportion of correct predictions among total predictions. It is calculated as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.5}$$

Table 4.5: Simplified Binary Classification Confusion Matrix.

	Positive Class	Negative Class
Predicted Positive	True Positive (TP)	False Positive (FP)
Predicted Negative	False Negative (FN)	True Negative (TN)

4.2.4 Results and analysis of the human activity recognition module

This section details the evaluation outcomes of the proposed deep learning architecture across diverse datasets. The analysis encompasses both smartphone-based datasets and smart home sensor data, emphasizing critical observations, encountered challenges, and comparative performance results. For clarity and coherence, the evaluation is organized into two subsections, each dedicated to addressing the unique characteristics and requirements of the respective dataset types.

4.2.4.1 Smartphone-based datasets (UCI HAR, WISDM, and PAMAP2)

The proposed model exhibited outstanding performance across all activity categories.

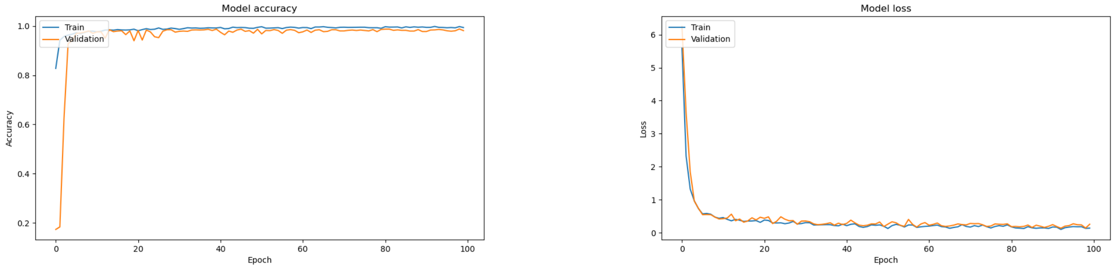
In the case of the *WALKING* activity from the UCI HAR dataset, the model achieved perfect scores in precision, recall, and F1-score (each reaching 100%), indicating highly accurate detection for this class. For the *WALKING_UPSTAIRS* activity, the model yielded a precision of 97.67% and a recall of 94.10%, resulting in an F1-score of 95.85%, reflecting strong performance with a slight decrease in recall. In the *WALKING_DOWNSTAIRS* class, the precision reached 94.66%, while recall was 97.64%, producing an F1-score of 96.12%, which demonstrates reliable classification of downward walking patterns.

For the *SITTING* category, the model attained a precision of 100% and a recall of 97.97%, leading to an F1-score of 98.97%, confirming high accuracy in identifying seated postures. The *STANDING* class achieved precision and recall values of 99.29% and 98.93%, respectively, corresponding to an F1-score of 99.11%, further evidencing the model's effectiveness. Regarding the *LAYING* activity, the precision was 97.17% with perfect recall (100%), yielding an F1-score of 98.56%, indicative of consistent and accurate recognition of this activity.

The overall accuracy achieved by the model was 98.06%. The macro-averaged metrics included a precision of 98.23%, recall of 98.11%, and an F1-score of 98.10%. The weighted averages for precision, recall, and F1-score were 98.09%, 98.06%, and 98.06%, respectively. These results underscore the model's robustness and generalization capability in recognizing a diverse set of human motion patterns.

Figure 4.8 displays the training accuracy and loss curves for the UCI HAR dataset over 100 epochs. We observe that the model converges after approximately 50 epochs. Furthermore, Figure 4.9 presents the corresponding confusion matrices for both epoch configurations, offering detailed insights into the model's classification accuracy and the nature of misclassifications.

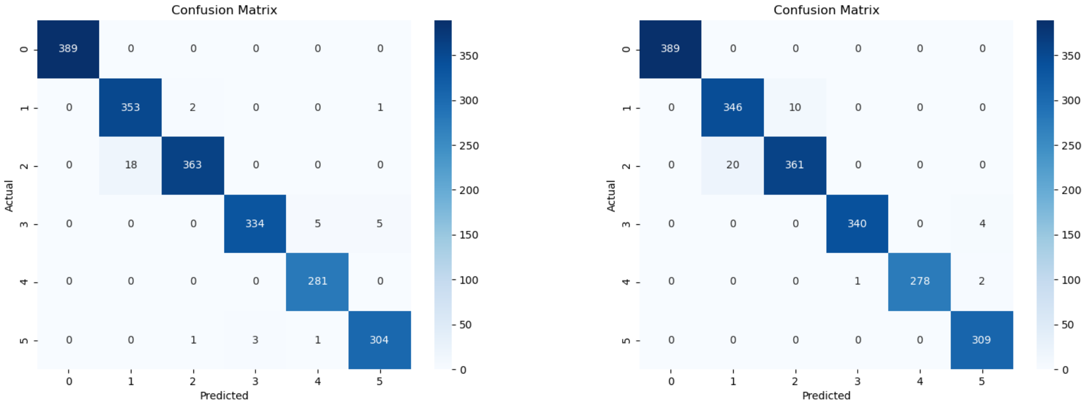
Within the WISDM dataset, notably high values of F1-score, precision, and recall were observed for the *WALKING* and *SITTING* activities, reflecting the model's strong capability in identifying these frequently occurring behaviors. These results suggest that the model effectively learns the distinctive sensor signatures associated with these specific activities. Similarly, the *STANDING* and *WALKING_UPSTAIRS* categories also demonstrate elevated performance metrics, indicating consistent and reliable recognition across these classes. Overall, the model



(a) Accuracy

(b) Loss

FIGURE 4.8: Accuracy and loss curves for UCI HAR with 100 epochs



(a) 50 epochs

(b) 100 epochs

FIGURE 4.9: Confusion matrix for the proposed CNN-MLP model with different epoch values for the UCI dataset

achieved a weighted average accuracy of 97.14%, while the macro-averaged precision, recall, and F1-score reached 96.26%, 95.69%, and 95.97%, respectively, confirming the balanced nature of the model's performance across different activity types.

Comparable to the UCI HAR dataset, WISDM encompasses a range of physical activities, each associated with distinct motion profiles and sensor placements. The exceptional results in *WALKING* and *SITTING* are likely attributable to the strong and consistent sensor patterns exhibited during these activities. In contrast, classification challenges in other categories may stem from sensor noise, ambiguous transitions between activities, or subtle movement characteristics.

The classification reports from both datasets reveal performance variations that can be attributed to the specific properties of each dataset. The model trained on UCI HAR demonstrates superior performance in activities such as *LAYING* and *SITTING*, likely due to consistent sensor patterns linked to these behaviors. Conversely, the WISDM-based model excels in

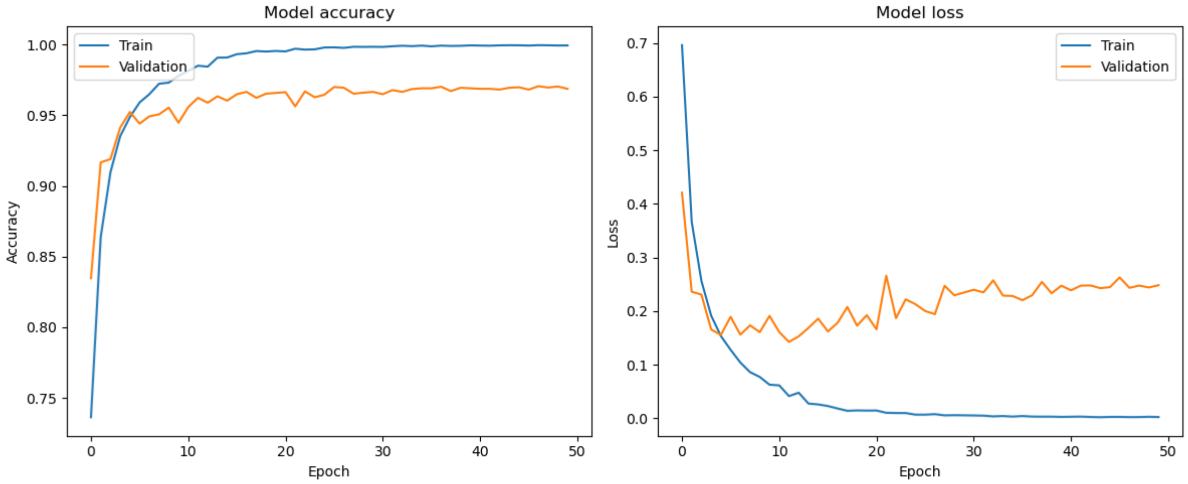


FIGURE 4.10: Wisdm curves for 50 epochs

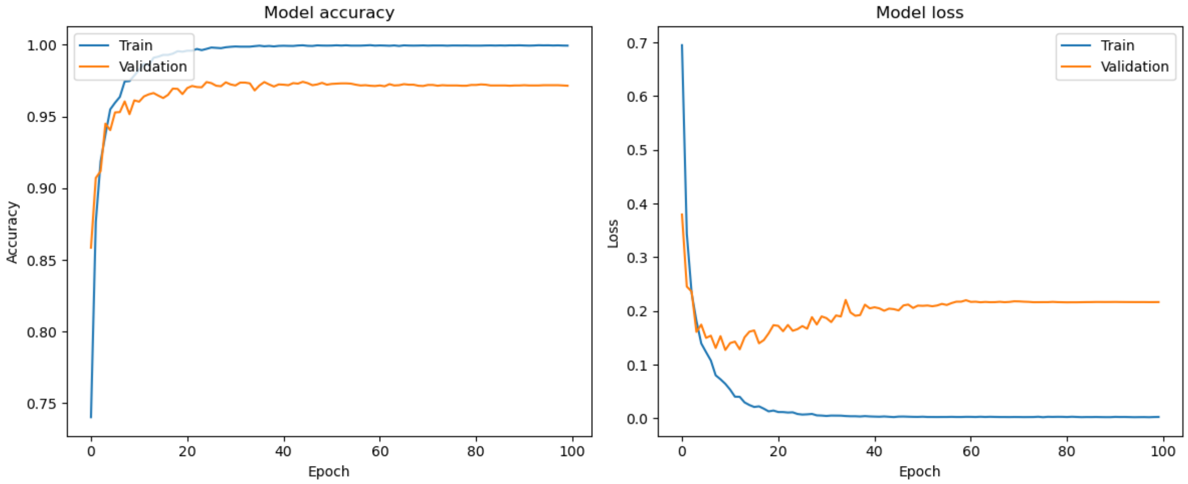
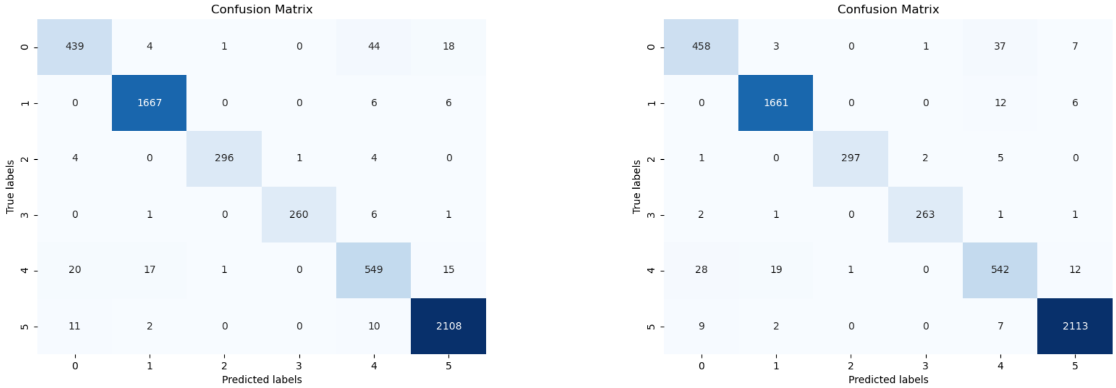


FIGURE 4.11: Wisdm curves for 100 epochs

identifying *WALKING* and *SITTING*, indicating the model's adaptability to the unique context and characteristics of each dataset. These differences highlight the necessity of tailoring recognition models to account for the distinct complexities inherent in individual datasets.

Figures 4.10 and 4.11 present the training accuracy and loss curves for the WISDM dataset using two different epoch configurations. Additionally, Figure 4.12 displays the confusion matrices generated for each epoch setting, offering deeper insight into the classification results and error distribution.

The proposed deep learning model exhibited outstanding performance on the PAMAP2 dataset, underscoring both its robustness and efficacy. This dataset, characterized by its balanced composition, encompasses a wide spectrum of activities including physical exercises and routine household tasks. To facilitate thorough temporal analysis, the data was downsampled to 33.3 Hz and segmented using a sliding window method with substantial overlap.



	0	1	2	3	4	5
0	439	4	1	0	44	18
1	0	1667	0	0	6	6
2	4	0	296	1	4	0
3	0	1	0	260	6	1
4	20	17	1	0	549	15
5	11	2	0	0	10	2108

	0	1	2	3	4	5
0	458	3	0	1	37	7
1	0	1661	0	0	12	6
2	1	0	297	2	5	0
3	2	1	0	263	1	1
4	28	19	1	0	542	12
5	9	2	0	0	7	2113

(a) 50 epochs

(b) 100 epochs

FIGURE 4.12: Confusion matrix for the proposed CNN-MLP model with different epoch values for the WISDM dataset

During training, the model demonstrated rapid convergence, with accuracy steadily increasing over 50 epochs, and both training and validation loss decreasing markedly. By the end of the training phase, the model showed excellent generalization, achieving a validation accuracy of 99.74%.

The classification report offers a detailed evaluation of the model's performance across all activities. Near-perfect values in precision, recall, and F1-score attest to the model's strong ability to accurately classify diverse activity types. Specifically, the activities *lying*, *sitting*, *standing*, *walking*, *running*, *cycling*, and *Nordic_walking* achieved 100% in all three metrics, reflecting flawless recognition. Other activities such as *ascending_stairs*, *descending_stairs*, *vacuum_cleaning*, *ironing*, and *rope_jumping* also attained very high scores, with precision, recall, and F1-score values ranging between 99% and 100%.

The correlation matrix heatmap provided additional insights into inter-sensor relationships. It was observed that gyroscope data had weak correlations with other sensor modalities, suggesting a limited contribution to model performance. In contrast, a strong correlation was identified between the chest magnetometer and heart rate sensors, as well as between the hand accelerometers and temperature sensors, likely due to their anatomical proximity.

The proposed deep learning architecture achieved exceptional results on the PAMAP2 dataset, delivering near-perfect accuracy and precision. These findings validate the model's architectural design and its suitability for human activity recognition applications. Figure 4.14 presents the training accuracy and loss curves for the PAMAP2 dataset, while Figure 4.13 shows the confusion matrix resulting from the CNN-MLP model applied to the PAMAP2 dataset.

A comparative evaluation of the model's performance across all three datasets is summarized in Table 4.6. This table consolidates key evaluation metrics, providing a comprehensive view of

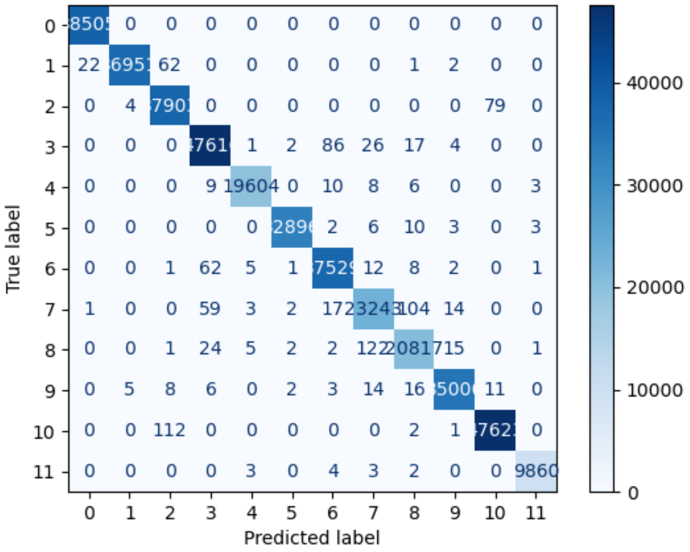


FIGURE 4.13: PAMAP confusion matrix

the model's effectiveness in diverse human activity recognition scenarios.

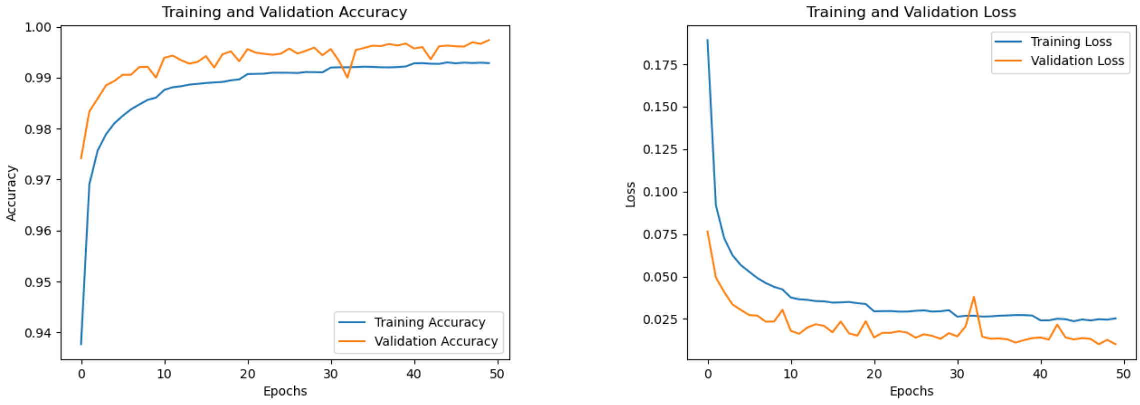
TABLE 4.6: Performance of the proposed CNN-MLP model across different epoch settings (50 and 100) on three benchmark datasets. The table reports precision, recall, F-score, and accuracy for each configuration.

Dataset	Metrics				
	Epochs	Precision	Recall	F-Score	Accuracy
WISDM	50	96.26	95.69	95.97	97.14
	100	96.16	95.90	95.76	97.05
UCI HAR	50	98.23	98.11	98.10	98.06
	100	98.36	98.35	98.35	98.35
PAMAP2	50	99.83	99.83	99.83	100
	100	98.36	98.35	98.35	98.35

4.2.4.2 CASAS Aruba dataset

The CNN-MLP architecture was evaluated on the CASAS Aruba dataset under two experimental configurations: one incorporating the *"other"* class, which encompasses unidentified or unannotated activities, and the other excluding it.

In the inclusive scenario, the model achieved an accuracy of 78.88%, whereas accuracy improved to 92.95% when the *"other"* class was excluded. This considerable enhancement underscores the impact of removing unlabeled or ambiguous data on the model's overall accuracy and reliability in activity recognition tasks.



(a) Accuracy curve

(b) Loss curve

Figure 4.14: Accuracy and loss curves for PAMAP2 dataset

Excluding the *"other"* category notably improved the F1-score and precision for specific activities such as *Meal_Preparation* and *Eating*. This improvement suggests that eliminating irrelevant or noisy data enables the model to focus more effectively on the distinguishing features of well-defined activities, thereby enhancing classification reliability.

Despite these changes, the model consistently maintained strong performance for activities such as *Relax* and *Sleeping*, achieving high F1-scores and precision in both scenarios. However, the classification accuracy for activities like *Enter_Home* and *Leave_Home* declined when the *"other"* class was omitted. This reduction in performance implies that the model relied on contextual cues present in the *"other"* class to differentiate these transition-based activities effectively.

Table 4.7 presents a detailed summary of precision and F1-score values for each activity across both scenarios, accompanied by relevant insights and interpretative commentary.

The CNN-MLP model demonstrates variable effectiveness in recognizing activities within the CASAS Aruba dataset. Notably, the exclusion of the *"other"* class leads to substantial improvements in overall accuracy and the recognition performance of several activities, emphasizing the critical role of preprocessing techniques in optimizing model outcomes for human activity recognition tasks.

Figures 4.17 and 4.18 illustrate the training accuracy/loss curves and the confusion matrix, respectively, for the model when trained on the Aruba dataset without excluding the unlabeled data. Conversely, Figures 4.15 and 4.16 present the corresponding results after removing the unlabeled data from the training process.

While the proposed CNN-MLP model demonstrates robust performance across multiple datasets, several limitations persist. One prominent challenge is the model's dependence on fully labeled datasets, which may not always be feasible in practical, real-world scenarios. Furthermore, although the model shows improved generalization capabilities, effective domain adaptation across heterogeneous sensor environments remains a difficult task. Despite its relative computational efficiency compared to earlier deep learning models, real-time deployment in resource-limited settings may necessitate additional optimization. Additionally, the model predominantly targets static activity recognition; future extensions could focus on capturing temporal dependencies and modeling transitions between sequential activities more effectively.

Table 4.7: Detailed results for CNN-MLP Model on CASAS Aruba Dataset

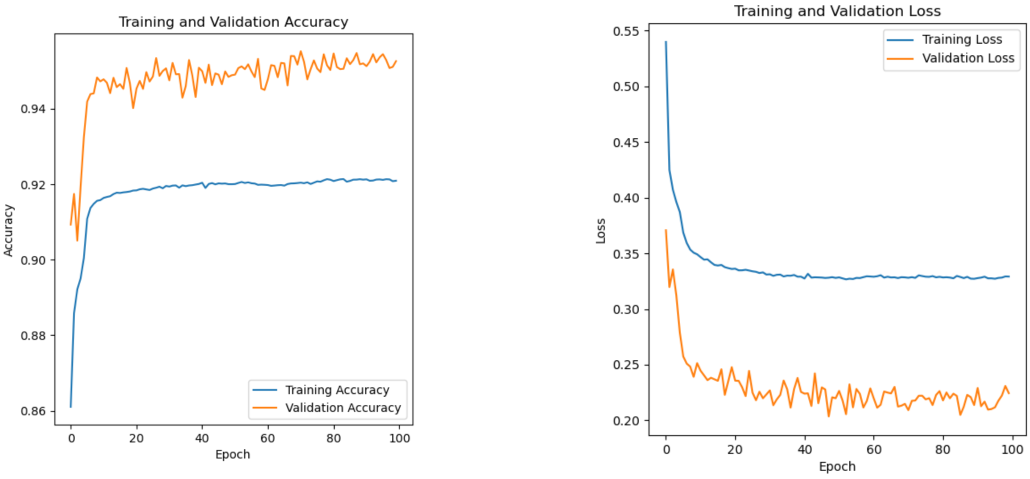
Activity	**Including "Other" Class**	**Excluding "Other" Class**	**Key Observations**
Overall Accuracy	78.88%	92.95%	Accuracy improves significantly when excluding the "other" class.
Bed_to_Toilet	Precision: 0.00 F1-score: 0.00	Precision: 0.00 F1-score: 0.00	The model fails to identify this activity in both cases due to low sample support.
Eating	Precision: 0.63 F1-score: 0.16	Precision: 0.81 F1-score: 0.79	Performance improves significantly when excluding the "other" class.
Enter_Home	Precision: 0.41 F1-score: 0.54	Precision: 0.00 F1-score: 0.00	The model shows better performance when including the "other" class.
Housekeeping	Precision: 0.00 F1-score: 0.00	Precision: 0.42 F1-score: 0.02	Performance remains poor in both cases.
Leave_Home	Precision: 0.00 F1-score: 0.00	Precision: 0.50 F1-score: 0.59	Significant improvement when excluding the "other" class.
Meal_Preparation	Precision: 0.65 F1-score: 0.71	Precision: 0.91 F1-score: 0.94	The model performs very well for this activity, with improvements when excluding the "other" class.
Other	Precision: 0.82 F1-score: 0.82	-	The "other" class shows strong performance when included.
Relax	Precision: 0.87 F1-score: 0.85	Precision: 0.96 F1-score: 0.96	The model performs well in both cases, with slight improvements when excluding the "other" class.
Respirate	Precision: 0.00 F1-score: 0.00	Precision: 0.00 F1-score: 0.00	The model fails to recognize this activity in both cases due to extremely low sample support.
Sleeping	Precision: 0.80 F1-score: 0.77	Precision: 0.87 F1-score: 0.91	Significant improvement when excluding the "other" class.
Wash_Dishes	Precision: 0.00 F1-score: 0.00	Precision: 0.00 F1-score: 0.00	The model fails to recognize this activity in both cases.
Work	Precision: 0.52 F1-score: 0.62	Precision: 0.87 F1-score: 0.88	The model performs moderately well when including the "other" class, with significant improvements when excluding it.

Table 4.8 illustrates the influence of dropout and L2 regularization on mitigating overfitting across four benchmark datasets: UCI HAR, WISDM, PAMAP2, and CASAS Aruba. The

(a) 50 epochs

(b) 100 epochs

FIGURE 4.15: Confusion matrices of Aruba dataset excluding unlabeled data



(a)

(b)

FIGURE 4.16: Training and validation accuracy and loss curves for the Aruba dataset excluding unlabeled data

Average Confusion Matrix

True label \ Predicted label	Bed_to_Toilet	Eating	Enter_Home	Housekeeping	Leave_Home	Meal_Preparation	Other	Relax	Respirate	Sleeping	Wash_Dishes	Work
Bed_to_Toilet	0	0	0.2	0	0	0	0.8	0	0	0	0	0.0035
Eating	0	0.038	0	0	0	0.048	0.9	0.01	0	0.0016	0	0
Enter_Home	0	0	0.79	0	0	0.002	0.21	0	0	0.002	0	0
Housekeeping	0	9.4e-05	0	0	0	0.13	0.81	0.051	0	0.0098	0	0.00038
Leave_Home	0	0	0.72	0	0	0	0.27	0	0	0.0035	0	0
Meal_Preparation	0	8.3e-05	0	0	0	0.79	0.21	0.0063	0	0.00017	0	0.00019
Other	0	0.00028	9.6e-05	0	0	0.12	0.82	0.047	0	0.016	0	0.0034
Relax	0	0	0	0	0	0.01	0.16	0.82	0	0.0026	0	0.00086
Respirate	0	0	0	0	0	0	0.99	0	0	0.0039	0	0.0019
Sleeping	0	0	0	0	0	0.00037	0.2	0.0024	0	0.79	0	0.0001
Wash_Dishes	0	0	0	0	0	0.86	0.14	0.0003	0	0	0	0
Work	0	0	6.2e-05	0	0	0.00043	0.67	0.00062	0	0.0039	0	0.32

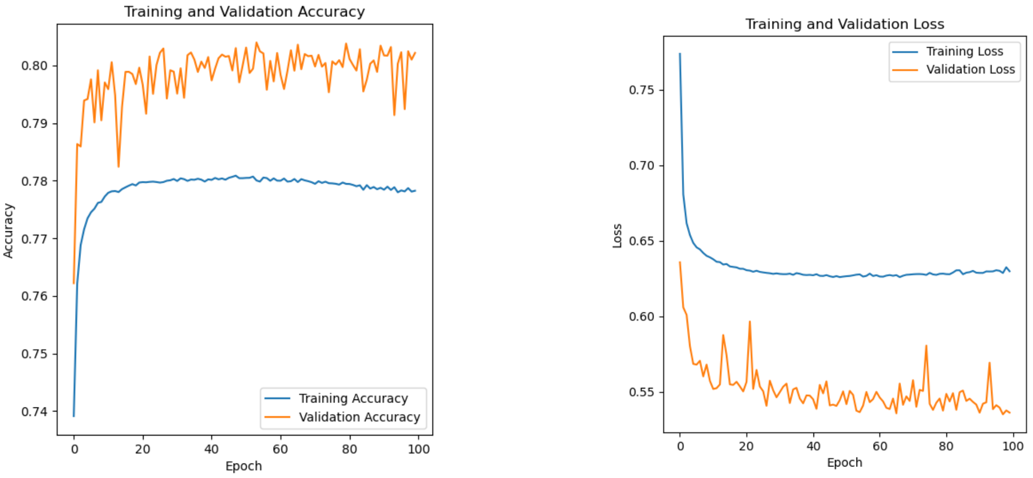
(a) 50 epochs

Average Confusion Matrix

True label \ Predicted label	Bed_to_Toilet	Eating	Enter_Home	Housekeeping	Leave_Home	Meal_Preparation	Other	Relax	Respirate	Sleeping	Wash_Dishes	Work
Bed_to_Toilet	0	0	0.021	0	0	0	0.98	0	0	0	0	0.0035
Eating	0	0.098	0	0	0	0.06	0.83	0.011	0	0.0014	0	0
Enter_Home	0	0	0.77	0	0	0.0039	0.2	0.0098	0	0.0079	0	0.002
Housekeeping	0	0.00066	0	0	0	0.13	0.79	0.065	0	0.009	0	0.0011
Leave_Home	0	0	0.71	0	0	0	0.28	0.0018	0	0.0035	0	0.0035
Meal_Preparation	0	0.00023	0	0	0	0.83	0.16	0.012	0	0.00017	0	0.00029
Other	0	0.00081	0.00011	0	0	0.13	0.79	0.061	0	0.013	0	0.008
Relax	0	1.1e-05	0	0	0	0.011	0.13	0.85	0	0.0022	0	0.0016
Respirate	0	0	0	0	0	0	0.97	0	0	0.0019	0	0.033
Sleeping	0	0	0	0	0	0.0011	0.24	0.00067	0	0.75	0	0.00018
Wash_Dishes	0	0.0001	0	0	0	0.89	0.11	0.0013	0	0.0001	0	0
Work	0	0	6.2e-05	0	0	0.00056	0.39	0.00031	0	0.0032	0	0.61

(b) 100 epochs

Figure 4.17: Confusion matrices for the Aruba dataset without excluding unlabeled data



(a)

(b)

Figure 4.18: Training and validation accuracy and loss curves for the Aruba dataset without excluding unlabeled data.

Table 4.8: Effect of Dropout and L2 Regularization on Overfitting across Different Datasets. The table compares training and validation accuracies under four configurations: no regularization, L2 only, Dropout only, and a combination of L2 and Dropout. The values in parentheses indicate the gap between training and validation accuracy, reflecting the degree of overfitting.

Dataset	**No Regularization**		**L2 Only**		**Dropout Only**		**L2 + Dropout**	
	Train Acc	Val Acc	Train Acc	Val Acc	Train Acc	Val Acc	Train Acc	Val Acc
UCI HAR	98.4%	92.5% (5.9%)	96.2%	93.2% (3.0%)	95.7%	93.6% (2.1%)	95.8%	98.35% (2.5%)
WISDM	96.5%	85.7% (10.8%)	94.1%	86.9% (7.2%)	93.4%	87.5% (5.9%)	99.2%	97.05% (2.15%)
PAMAP2	97.1%	89.1% (8.0%)	95.6%	90.4% (5.2%)	94.8%	90.9% (3.9%)	99.72%	98.35% (1.37%)
CASAS Aruba	95.9%	78.2% (17.7%)	92.8%	80.1% (12.7%)	91.3%	81.0% (10.3%)	98.46%	92.95% (5.51%)

results compare training and validation accuracy under four configurations: no regularization, application of L2 regularization only, dropout only, and a combination of both (L2 + Dropout).

In the absence of regularization, training accuracy remains consistently high for all datasets; however, a notable decline in validation accuracy indicates substantial overfitting. For example, in the CASAS Aruba dataset, the difference between training and validation accuracy reaches 17.7%, reflecting pronounced overfitting. The inclusion of L2 regularization introduces a modest decrease in training accuracy due to penalization of large weights, but yields an improvement in validation accuracy, thereby narrowing the overfitting gap. Dropout, when used independently, also proves effective in reducing overfitting. This is evident in the UCI HAR dataset, where validation accuracy improves from 92.5% to 93.6%, and the gap between training and validation accuracy is reduced from 5.9% to 2.1%.

The integration of both L2 regularization and dropout consistently results in superior performance across all datasets, achieving the lowest overfitting margins and enhanced generalization. For instance, in the PAMAP2 dataset, this combined strategy reduces the overfitting gap to 1.37% and elevates validation accuracy to 98.35%, outperforming configurations utilizing a single regularization method. These findings affirm that combining dropout and L2 regularization effectively strengthens model robustness by limiting neuron co-adaptation and constraining weight magnitudes.

While the datasets employed in academic research provide controlled environments for model evaluation, real-world applications introduce additional complexities, such as sensor noise, unanticipated activities, and environmental variability. To address these challenges, the proposed model integrates dropout and L2 regularization, which help prevent overfitting and enhance robustness to novel or noisy data. The hybrid CNN-MLP architecture is adept at capturing both spatial and contextual features, thereby ensuring adaptability to diverse sensor inputs.

To examine scalability, we assessed the model's behavior across datasets of varying sizes. Results indicate that the architecture maintains consistent inference times, demonstrating its suitability for larger-scale applications. Moreover, regularization mechanisms—including L2 penalties, dropout, and batch normalization—contribute to robust generalization, mitigating the risk of overfitting to specific subjects or sensor configurations.

The model also supports real-time inference by processing sequential input streams efficiently, making it well-suited for continuous monitoring in domains such as smart homes and healthcare. Profiling experiments confirm that inference latency remains within operational thresholds required for real-time applications, even as the dataset size increases.

Table 4.9: Computational complexity comparison of the proposed model across datasets.

Dataset	**No. of Samples**	**Feature Dim.**	**Training Time (min)**	**Inference Time (ms/sample)**
WISDM	1,098,207	9	75.3	1.12
UCI HAR	10,299	561	22.7	0.94
PAMAP2	2,844,187	52	142.5	1.45
CASAS Aruba	1,554,298	17	89.2	1.30

To evaluate the computational efficiency of the proposed model, we conducted a comparative analysis across multiple datasets, considering metrics such as training time, inference time, total number of samples, and feature dimensionality (Table 4.9). It is important to note that feature extraction and segmentation strategies substantially influence both the resulting sample count and the dimensionality of the input features. The findings reveal that although extensive datasets like PAMAP2 necessitate longer training durations due to their size and complexity, inference times remain consistently low across all datasets. This outcome underscores the suitability of the proposed architecture for real-time applications, where rapid and efficient inference is critical.

The proposed CNN-MLP architecture strikes an effective balance between classification accuracy and computational efficiency, positioning it as a strong candidate for real-time Human Activity Recognition (HAR) applications. One of the principal challenges in developing deep learning models for HAR lies in managing the trade-off between model complexity and execution speed. While highly complex architectures may yield greater accuracy, they often incur substantial computational costs, rendering them unsuitable for deployment on resource-constrained edge devices.

To evaluate this trade-off, we examine three critical computational metrics: the total number of trainable parameters, the number of floating-point operations (FLOPs), and the average inference time per sample. The CNN-MLP model comprises **2.1 million parameters** and performs approximately **320 million FLOPs** per inference. These values are significantly lower than those required by standard CNN-LSTM configurations, while still achieving superior classification performance. The model's low inference latency further reinforces its applicability in real-time scenarios.

This reduction in computational burden is primarily attributed to several key architectural optimizations:

- **Global Average Pooling (GAP):** Instead of fully connected layers with a large number of parameters, we employ GAP to reduce dimensionality while preserving important spatial information.
- **Efficient Feature Extraction:** The CNN component extracts hierarchical spatial features, reducing the dependency on large recurrent layers that typically increase computational load.
- **Regularization and Dropout:** To prevent overfitting while maintaining efficiency, we incorporate dropout layers, ensuring the model generalizes well without excessive complexity.
- **Layer Reduction Strategy:** Unlike multi-layered deep networks, our model maintains a compact structure, limiting the depth while retaining sufficient feature extraction capacity.

Despite the applied optimizations, the proposed model consistently delivers high accuracy scores, surpassing traditional CNN and CNN-LSTM architectures while retaining computational efficiency. These characteristics render it particularly suitable for real-time HAR applications, especially in edge computing contexts where constraints on latency and energy consumption are paramount.

The results underscore the potential of CNN-MLP architectures as robust and efficient alternatives to conventional CNN-LSTM models. By offering a favorable trade-off between predictive performance and computational overhead, the proposed approach demonstrates its practicality for deployment in resource-limited environments.

The following section provides a comparative performance analysis, benchmarking the proposed model against recent state-of-the-art approaches to further assess its effectiveness.

4.2.5 Comparative results

Table 4.10 provides a comparative analysis of our proposed model against several baseline approaches on the WISDM, UCI HAR, and PAMAP2 datasets. Statistical significance is evaluated using p-values derived from paired t-tests, while the 95% confidence intervals reflect variability in model performance.

Hurtado et al. [211] introduced a semi-supervised Encoder-Decoder CNN aimed at healthcare-oriented HAR tasks, leveraging unlabeled data to enhance generalization. However, their method lacks comprehensive benchmarking on widely accepted datasets. In contrast, our CNN-MLP model demonstrates superior generalization capabilities, achieving higher accuracy across all three datasets. Seelwal et al. [212] presented a CNN-based model that attained an accuracy of 87.85% on the WISDM dataset, though it did not incorporate architectural enhancements. Our approach significantly improves upon this, reaching 96.9% accuracy on the same dataset, underscoring the effectiveness of our hybrid design.

Kaya et al. [213] employed a 1D-CNN architecture with extensive hyperparameter tuning, achieving 97.8% accuracy on WISDM and 90.27% on PAMAP2. Our model surpasses these results, attaining 99.74% accuracy on PAMAP2, indicating superior adaptability. Additionally, Rong et al. [214] proposed the MULTI-SCALE TIME SEGMENTS ATTENTION (MTSA) method to improve temporal segmentation efficiency without increasing computational cost. While effective, their approach emphasizes time-series segmentation over network-level optimization. In contrast, our CNN-MLP architecture yields consistently higher accuracy across all datasets, highlighting its architectural refinement and broader applicability.

Table 4.10: Comparison of our model with baseline models on WISDM, UCI HAR, and PAMAP2 datasets (Accuracy in %), including statistical significance.

Method	**WISDM**			**UCI HAR**			**PAMAP2**		
	Accuracy (%)	p-value	95% CI	Accuracy (%)	p-value	95% CI	Accuracy (%)	p-value	95% CI
CNN [84, 215]	93.32	0.0	[92.83, 93.81]	92.71	0.0	[92.20, 93.22]	91.16	0.0	[90.60, 91.72]
Multi-input CNN LSTM [216]	95.54	4.59×10^{-7}	[95.14, 95.94]	95.13	0.0	[94.71, 95.55]	94.04	0.0	[93.58, 94.50]
CNN-based [212]	82.27	0.0	[81.52, 83.02]	-	-	-	-	-	-
Encoder-Decoder CNNs[211]	66.7	0.0	[65.78, 67.62]	-	-	-	-	-	-
1D-CNN[213]	97.8	7.42×10^{-5}	[97.51, 98.09]	-	-	-	90.27	0.0	[89.69, 90.85]
MTSA [214]	**98.52**	0.0	**[98.28, 98.76]**	**97.18**	0.8648	**[96.86, 97.50]**	93.66	0.0	[93.18, 94.14]
CNN-LSTM [103, 217, 218]	86.3	0.0	[85.63, 86.97]	92.13	0.0	[91.60, 92.66]	92.81	0.0	[92.30, 93.32]
Proposed Model	**96.9**	**-**	**[96.56, 97.24]**	**97.14**	**-**	**[96.81, 97.47]**	**99.74**	**-**	**[99.64, 99.84]**

Table 4.11 provides a comparative overview of different models evaluated on the CASAS Aruba dataset, emphasizing both their classification accuracy and the methodological innovations introduced by each approach.

Table 4.11: Comparison of the performance of our model on the Aruba dataset with other models, including statistical significance

Work	**Excluding "Other"**	**Without Excluding "Other"**	**p-value**	**95% Confidence Interval**
[219]	91.88%	/	0.0043	[91.72, 92.04]
[220]	87%	69%	<0.0001	[86.82, 87.18] / [68.74, 69.26]
[221]	91%	/	0.00000038	[90.84, 91.16]
[222]	89%	/	<0.0001	[88.84, 89.16]
[176]	/	**80.87**%	<0.0001	[80.72, 81.02]
[223]	92.4%	/	0.136	[92.24, 92.56]
[224]	83.2%	/	<0.0001	[83.04, 83.36]
Our model	**92.95**	**78.88**	**-**	**[92.80, 93.10] / [78.72, 79.04]**

Several studies have proposed varied strategies to address human activity recognition (HAR) in smart home environments. Wen et al. [219] utilized a combination of labeled and unlabeled data to improve recognition performance, attaining an accuracy of 91.88% when the "Other" category was excluded. Ashry et al. [221] developed CHARM-Deep, a deep learning-based framework tailored for optimized feature engineering, achieving 91% accuracy. Yala et al. [220] addressed the challenges of streaming and highly imbalanced datasets, reporting an accuracy of 87% without "Other" and 69% when it was included, thus underscoring the difficulty of managing ambiguous or infrequent activity types. Boralessa et al. [222] proposed a deep neural network (DNN) approach that encodes binary sensor signals as video sequences, yielding an accuracy of 89%.

Contemporary approaches have increasingly explored advanced architectures to enhance recognition capabilities. Chen et al. [176] introduced AttCLHAR, a self-supervised model that integrates SimCLR, attention mechanisms, and sharpness-aware minimization, achieving 80.87%. Plötz et al. [223] incorporated graph-based modeling within deep learning pipelines and reported a classification accuracy of 92.4%. Ramadan et al. [224] proposed a real-time HAR system that enhances activity alignment through the Needleman-Wunsch algorithm, achieving an accuracy of 83.2%.

In contrast, the proposed CNN-MLP model surpasses these prior methods, attaining 92.95% accuracy when the "Other" category is excluded and maintaining a robust performance of 78.88% even when it is included. By leveraging convolutional layers for feature extraction and a multilayer perceptron for classification, the model effectively captures both spatial and temporal dependencies, demonstrating resilience in real-world HAR scenarios.

To further improve generalization, future enhancements may include transfer learning, allowing the model to quickly adapt to new activity types with limited labeled data. In addition, incorporating online learning techniques would enable the system to update its predictions dynamically as new activity patterns are observed. Ongoing research will focus on evaluating the model's effectiveness under real-world conditions and implementing domain adaptation strategies to enhance its resilience in evolving environments.

Scalability is another critical consideration for deploying the proposed model in practical scenarios, particularly in settings involving multiple users, larger datasets, or increased complexity. The model is engineered to retain computational efficiency as data volume increases, employing lightweight convolutional layers and global average pooling to minimize resource consumption without compromising predictive performance. Future development will focus on optimizing the model further through techniques such as model pruning and quantization, facilitating deployment on edge devices and enabling scalability to larger environments with constrained computational resources.

4.3 Implementation Details of the Stream Reasoning Approach on Human Behavior for Medication Risk Detection using the LARS Framework

This section presents the implementation methodology for our stream reasoning approach, designed to detect medication-related risks such as missed or excessive doses in elderly individuals. The logical rules were formulated to infer the medication adherence behavior of an older adult—referred to here as “Mary”—and to identify deviations that may indicate potential health hazards.

```
wellDone :- havMedication [10 sec].
havMedication at T :-  drug(C), hav(C) at T in [3 #], not hav(C) [2 #].
havMedication at T :- hav(1) at T in [3 #], hav(2) at T in [3 #], hav(3) at T in [3 #].
havMedication at T :-  hav(1) at T in [1 #], hav(2) in [2 #], hav(3) in [3 #].
havMedication at T :- hav(3) at T in [3 #], not hav(3) [2 #], hav(2) at T in [2 #], not hav(2) [1 #], hav(1) [1 #]
```

Table 4.12 outlines the primary terms utilized in the formulation of our logical rules, including their respective definitions and the specific conditions under which each inference is activated.

Leveraging this reasoning logic, the system evaluates whether Mary has taken each prescribed medication within the specified temporal window. The rules are designed to verify the intake of the three medications both independently and in sequential order.

These rules are executed over a configurable 10-second sliding window. The initial rule infers `wellDone` if at least one medication intake event is detected within the preceding 10 seconds.

The second rule employs a tuple-based windowing mechanism over the three most recent medication intake events. It verifies whether a `hav(C)` atom—indicating consumption of a given drug—appears within the last three but not among the last two events. This pattern implies that the detected medication is the third in the sequence, and the corresponding timestamp is annotated with the `havMedication` atom.

- **Missed dose risk** The system identifies a missed dose scenario when one or more medications are not taken on time. The following rules address this risk:

```
missedDose:- lessThanThree [10 sec].
lessThanThree at T:- not wellDone, drug(C), hav(C) at T in [3 #], not hav(C) [1 #].
lessThanThree at T:- not wellDone, drug(C), hav(C) at T in [2 #].
```

```
noMed:- not wellDone, not missedDose.
```

The first rule asserts the presence of a missed dose if `lessThanThree` is detected in the previous 10 seconds. The next two rules evaluate whether the individual took fewer than the prescribed three medications — either only one or two — within that timeframe. If none were taken, the `noMed` atom is inferred.

- **Overdose risk and Prevention** An overdose is inferred if the individual consumes more than the prescribed three medications within the valid time window.

```
overDose:- moreThanThree[10 sec].
moreThanThree at T:- not wellDone, drug(C), hav(C) at T in [4 #],
   not hav(C) [3 #].
```

Here, `overDose` is inferred if a fourth dose is observed in the last 10 seconds. The rule detects if a drug is taken as a fourth entry among the last four events — but not in the last three — and if `wellDone` was not already satisfied.

Table 4.12: Used Terms, their Descriptions and Conditions

Result	**Description**	**Conditions**
noMed	Mary didn't have any of her drugs yet	• No actions of the type hav(drug) recorded
wellDone	Mary has all of her drugs in time	• Three actions of type hav(drug) have been detected • All drugs were taken in the correct order • All drugs were taken in the correct time
missedDose	A missed dose risk is detected	• One or more of the drugs was not taken
overDose	An overdose risk is detected	• One additional dose was taken
lessThanThree	Less than three medications are taken	• Only one (the first, second or third) medication is taken • Only two of the medications are taken
moreThanThree	Mary has more than three doses of medications	• A fourth (or more) dose of medication (1, 2 or 3) is taken within the specified time window

Table 4.13: Results of Experiments of the Stream Reasoning Approach

Experiment	Entered Actions with Drug Reference	Results		Inference Time (sec)	
		t<Window Size	t≥Window Size	Window Size of 10 sec	Window Size of 20 sec
1	hav(Drug X) (any of the three drugs)	lessThanThree, missedDose	noMed	5	10
2	hav(1), hav(2)	lessThanThree, missedDose	noMed	7	15
3	hav(1), hav(2), hav(3), hav(3)	wellDone, overDose	noMed	8	12
4	hav(1), hav(2), hav(3)	wellDone	noMed	6	14
5	hav(2), hav(3)	lessThanThree, missedDose	noMed	9	18
6	no action recorded	missedDose	noMed	4	8
7	hav(1), hav(2), hav(3), hav(4) (drug 4 is taken after the specified time window size)	wellDone	noMed	11	22

CQL query language

Continuous Query Language (CQL) is an extension of SQL that incorporates novel operators specifically tailored for handling streaming data. It introduces two essential categories of operators: Stream-to-Relation and Relation-to-Relation operators.

Stream-to-Relation operators facilitate the application of windowing functions to incoming data streams. These operators define mappings between execution times and the set of tuples valid within a specific window, forming a relation without retaining timestamp information. In contrast, Relation-to-Relation operators, akin to those found in relational algebra and traditional SQL, are employed to manipulate these derived relations in a more conventional manner [225].

CQL distinguishes itself through its specialized syntax and operator set, enabling continuous querying and real-time analytics. By utilizing window operations, CQL allows for refined analysis such as examining the most recent events — e.g., the last 10 cases or events within the last 10 minutes. Intermediate query results are stored in relational structures, processed using projection and selection operators modeled after SQL, and subsequently converted back into streams using stream-to-relation transformations.

The following CQL query illustrates how our framework retrieves information in the context of our experimental scenario:

```
SELECT ACTION, DRUG_REFERENCE
FROM MEDICATION_STREAM[PARTITION BY PATIENT_ID ROWS 1]
WHERE PATIENT_ID = 'Mary'
    AND ACTION = 'hav'
```

```
    AND DRUG_REFERENCE IN (1, 2, 3)
    AND TIME >= CURRENT_TIMESTAMP - INTERVAL '10' SECOND
LIMIT 1;
```

In this query:

`MEDICATION_STREAM[PARTITION BY PATIENT_ID ROWS 1]` retrieves the latest row for each unique patient ID in the `MEDICATION_STREAM` stream. The condition `PATIENT_ID = 'Mary'` filters the results only to include entries related to Mary. The condition `ACTION = 'hav'` selects rows with the action 'hav'. The condition `DRUG_REFERENCE IN (1, 2, 3)` selects rows with the specified drug references. `TIME >= CURRENT_TIMESTAMP - INTERVAL '10' SECOND` filters the results to include entries within the last 10 seconds. `LIMIT 1` returns only the latest relevant row.

This query effectively retrieves Mary's latest intake event for any of the three prescribed medications within a 10-second window.

For the implementation, we adopted Ticker [226], a stream reasoning engine compatible with the LARS framework. Ticker, developed in Scala, supports both time-based and tuple-based windows. The rule set was authored in a file named `program.LARS` using the IntelliJ IDEA 2022.1.2 development environment.

Results of the Stream Reasoning Approach

Our framework outputs the label `wellDone` when it confirms that all three medications have been correctly administered within a 10-second window.

Missed Dose Risk

The system is capable of detecting the likelihood of a missed dose when not all prescribed medications are taken within the expected timeframe. During experimentation, a simulated input stream was used with events `hav(1)`, `hav(2)`, and `hav(3)`, where each `hav(x)` denotes the administration of medication `x`.

After each recorded intake event, the system provides feedback to the patient. Initially, it signals `noMed` to indicate that no medication has been taken. As medications are consumed, it updates the status to `missedDose` until all three medications are registered within the 10-second window. Upon successful completion, the system outputs `wellDone`, confirming proper adherence to the prescription.

If, for instance, the sequence `hav(1)` and `hav(3)` is received and no further medications are taken before the time window expires, the system identifies a potential missed dose.

Overdose Risk

Additionally, the system is designed to identify overdose risks — situations where an extra dose is administered beyond the prescribed three within the same temporal window.

To evaluate this feature, a test stream with events `hav(1)`, `hav(2)`, and `hav(3)` was used. A window of 10 seconds was configured, within which the reasoning engine applied the predefined rules to detect overdosing behavior.

The program flags a potential overdose only if a fourth `hav(C)` event occurs within this 10-second period. If an additional medication dose is observed after the time window has ended, it is not considered an overdose.

Table 4.13 outlines the experimental scenarios and the corresponding outputs.

The scalability of this approach is of paramount importance. As the volume and complexity of real-time data streams continue to grow, our method's modular and adaptive architecture ensures it remains efficient and accurate. Its capacity to handle continuous data and incorporate new technologies seamlessly makes it suitable for diverse domains such as smart healthcare and urban monitoring.

Compared to traditional static methods, which often struggle with dynamic data interpretation, our solution stands out for its ability to apply temporal logic to real-time behavior tracking. This allows the construction of sophisticated cognitive models to understand human behavior patterns with greater precision. Particularly in healthcare applications, the ability to rapidly detect anomalies such as missed doses or overdoses supports timely interventions and enhances patient safety. The integration of LARS for reasoning over time-aware data streams provides significant advantages in terms of responsiveness and accuracy, making this approach a valuable tool for advancing human activity monitoring.

This framework presents a stream reasoning approach leveraging the LARS framework, tailored to detect medication-related risks in elderly individuals. The primary objective is to monitor behavioral patterns and infer threats such as missed or excess doses in real-time, particularly beneficial for seniors living independently. The proposed reasoning model is implemented using LARS syntax and temporal constructs to ensure effective and accurate detection. To the best of our knowledge, this represents the first deployment of LARS for behavioral analysis over continuous data streams.

4.4 Proposed Ontological Querying-based Cognitive Perspective framework for Activity Recognition.

This section presents the performance evaluation of the proposed framework, which integrates ontology-based reasoning with real-world sensor data. Ontology construction was carried out using Protégé, while the implementation was performed in Java utilizing the GENA API. The Human Behavior Ontology (HBOnt) was derived from the dataset and interrogated through SPARQL queries using Apache Jena. Experimental analysis was conducted on a system equipped with an Intel i5 processor and 8 GB of RAM. The study specifically investigates behavioral patterns related to the *sleep* activity. By analyzing activities occurring immediately before and after sleep, the framework reveals how routine behaviors and environmental contexts can influence cognitive dimensions such as emotional well-being and stress regulation.

4.4.1 Datasets description

To assess the effectiveness of the proposed approach, two datasets were employed: the CASAS Aruba dataset (Sections 4.2.1.4 and 4.2.2.2) and the Orange4Home dataset [1]. These datasets were chosen for their structured time-series format and their relevance to diverse scenarios in activity and behavior recognition.

Due to the substantial volume of data in the CASAS Aruba dataset (1,048,576 rows), preprocessing was essential to manage missing entries, eliminate irrelevant information, and prepare the data for semantic mapping within the ontology. To ensure computational feasibility, only the initial four days of the dataset were utilized, encompassing a total of 20,897 events.

The Orange4Home dataset was selected for its structured temporal nature and multimodal complexity, making it well-suited for evaluating behavior recognition using semantic reasoning. This dataset includes approximately 180 hours of data recorded in a two-story smart home (Figure 4.19), where a single resident performed daily activities across four consecutive work weeks. It features 17 annotated activities: *Entering, Living, Preparing, Cooking, Washing the dishes, Eating, Watching TV, Computing, Using the toilet, Going up, Going down, Using the sink, Showering, Dressing, Reading, and Napping.* The dataset's multimodal nature—comprising wearable, object-based, and ambient sensors—offers a rich testbed for evaluating ontology-based inference methods. A sliding window of 3 seconds was adopted for segmentation, striking a balance between granularity and contextual awareness. Table 4.14 provides an overview of the dataset's characteristics.

To prepare the Orange4Home dataset for model input, a structured preprocessing pipeline was implemented. Initially, the 17 activities were label-encoded to assign unique numerical identifiers. Missing sensor readings were addressed using forward-fill interpolation. All continuous sensor values were normalized to the [0,1] range. Sensor fusion was then applied to integrate data from various sensor modalities into a unified feature matrix. The dataset was segmented into overlapping 3-second windows, with each window assigned the dominant activity label. Finally, one-hot encoding was applied to convert categorical activity labels into binary vectors suitable for classification tasks.

For evaluation purposes, the dataset was partitioned into training (70%), validation (15%), and testing (15%) subsets. The training set facilitated model learning, the validation set supported hyperparameter tuning, and the test set was reserved for assessing final model performance.

TABLE 4.14: Summary of Orange4Home dataset characteristics

Property	**Details**
Sensor Types	Wearable, object, and ambient sensors
Data Types	Binary, integer, real number, categorical
Number of Activities	17 annotated activities
Participants	Single occupant
Preprocessing	Normalization, segmentation into fixed-size windows
Recording Duration	180 hours over four consecutive work weeks

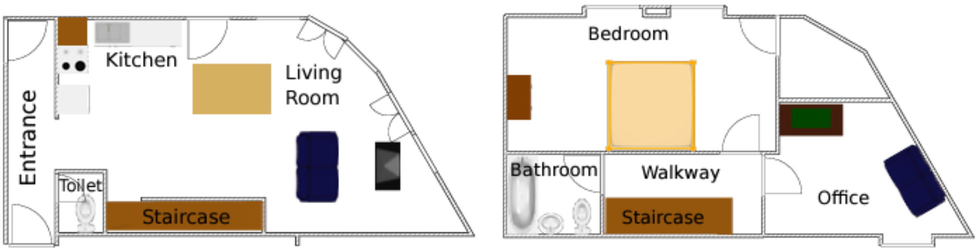


FIGURE 4.19: Smart Home Layout Used for Orange4Home Dataset Collection [1]

4.4.2 Results and discussion of the proposed framework

The findings of this study indicate that consistent sleep schedules are positively associated with cognitive mechanisms involved in stress regulation, corroborating prior research emphasizing the significance of sleep regularity in cognitive health. Furthermore, sleep quality appears to be influenced by both the nature of preceding activities—whether calming or stimulating—and external environmental factors such as ambient noise. These observations align with theories in cognitive psychology that posit behavior and cognition are shaped by habitual routines and environmental context.

Utilizing SPARQL queries enables an in-depth examination of individual behavioral patterns. For example, analysis of the subject "Mary" reveals a structured daily routine characterized by sequential activities including meal preparation, relaxation, and sleep. Query outputs demonstrate that certain activities, such as meal preparation, may embody personal values like creativity or caregiving. Additionally, a consistent preference for relaxation prior to sleep underscores the role of unwinding routines in supporting rest. The examination of activity transitions further highlights the presence of a balanced and repetitive structure in daily life, reinforcing the cognitive psychological view that behavioral regularity contributes to psychological well-being.

From the perspective of cognitive theory, both the duration and temporal placement of activities can reflect underlying behavioral routines, physiological states, mental needs, and reactions to external stimuli. In contrast to brief, impulsive actions, **Query 1** was designed to identify occurrences of the *Relax* activity that exhibit prolonged engagement. Such sustained periods of relaxation may signify intentional rest or focused disengagement, potentially indicating a higher level of cognitive involvement and reflective behavior.

- **Results for Query 1:**
  Activity: *Relax_2024-07-04_1*
  Start Time: *2024-07-04T09:29:23*
  End Time: *2024-07-04T09:34:05*
  Activity: *Relax_2024-07-06_1*
  Start Time: *2024-07-06T11:11:08*
  End Time: *2024-07-06T11:25:23*

Prolonged relaxation periods reflect cognitive recovery strategies consistent with theories of resource allocation and cognitive equilibrium. In the observed instances, repeated or extended rest phases suggest that the individual may be engaging in deliberate recovery following episodes of exertion.

Moreover, the short transitions between tasks highlighted in **Query 3** can be interpreted as manifestations of goal-oriented behavior, wherein each activity is directed toward fulfilling a specific internal need or responding to an external stimulus. This interpretation aligns with cognitive psychological frameworks that conceptualize motivation as an organized sequence of actions guided by perceived demands or objectives.

- **Results for Query 3:**
  Start Time: *2024-07-06T11:11:08*
  End Time: *2024-07-06T11:25:23*
  Activity Name: *Relax_2024-07-06_1*

Start Time: *2024-07-06T22:27:35*
End Time: *2024-07-06T22:36:08*
Activity Name: *Relax_2024-07-06_6*

The observed transitions from *Relax* to other activities such as *Meal Preparation* may reflect a structured adherence to predefined routines or schedules, often driven by physiological needs (e.g., hunger). This suggests that certain actions are prioritized in response to immediate, time-sensitive demands. Such behavioral patterns are consistent with motivational theories, which posit that individuals regulate the interplay between leisure and responsibilities by engaging in specific tasks based on necessity or habitual behavior.

As illustrated in **Query 3**, changes in spatial positioning and activity type offer valuable insights into adaptive behavior in response to environmental stimuli.

- **Results for Query 3:**
  Start Time: *2024-07-06T18:29:36*
  End Time: *2024-07-06T18:41:18*
  Activity Name: *Relax_2024-07-06_2*
  Location:*Living*
  Start Time: *2024-07-04T16:36:10*
  End Time: *2024-07-04T17:06:00*
  Activity Name: *Meal_Preparation_2024-07-04_8* Location:*Kitchen*

In this instance, transitions between tasks—such as shifting from *Relax* in one location to *Meal Preparation* in another—exemplify cognitive flexibility. This behavioral adjustment, prompted by contextual cues and environmental resources, indicates situational awareness and the capacity to modify actions accordingly. Such flexibility is a core component of cognitive adaptability, where actions remain both goal-directed and responsive to situational changes.

These examples illustrate a subset of our findings, demonstrating the utility of ontology-based querying in inferring motivational states, evaluating behavioral patterns, and detecting anomalies. The framework enhances our capacity to identify effective routines, recognize deviations that may signal cognitive fatigue, and interpret behaviors within the context of established motivational theories.

A notable advantage of our ontology-based approach is its inherent flexibility in handling novel situations. Ontologies, by design, are extensible; new concepts and relationships can be seamlessly integrated. This is particularly beneficial in dynamic environments with evolving datasets and behavioral contexts. Unlike conventional deep learning models—which often require extensive retraining to incorporate new activities—our method allows for the addition of new classes or properties to the ontology, thereby supporting scalability and adaptability.

Interpretability is a critical requirement in Human Activity Recognition (HAR), particularly when the objective extends beyond classification to understanding human intent. Our ontology-based reasoning approach provides high transparency, as each inference is supported by explicitly defined rules and semantic relationships. This contrasts with deep learning models, which are frequently described as “black boxes” due to their limited explainability regarding internal decision processes. The transparency of our method enables traceable and explainable reasoning, making it particularly suitable for applications where interpretability is essential.

The structural design of our ontology captures not only discrete actions but also the sequential and relational dependencies among them, offering a more comprehensive model of human behavior. This enables the inference of higher-level motivations behind action sequences—for instance, deducing the intent of “preparing to sleep” based on preceding actions such as “brushing teeth” and “changing clothes.” Standard machine learning models typically do not support this form of hierarchical and relational reasoning, limiting their capacity to uncover deeper behavioral insights.

A comparative overview of these key characteristics is presented in Table 4.16, which highlights the advantages of our ontology-based approach over widely used deep learning models in the context of HAR.

Table 4.15: Comparative Analysis of Query Efficiency (CASAS Aruba and Orange4Home)

Dataset	**Number of Activities**	**Number of SPARQL Queries Run**	**Average Query Time (s)**	**Max Query Time (s)**	**Query Complexity**
CASAS Aruba	11	150	0.25	0.35	Medium
Orange4Home	17	200	0.20	0.30	Medium

Table 4.16: Comparison of the proposed ontology-based approach with different types of deep learning models in HAR

Metric	**Our Proposed Approach**	**Deep Learning Models**
Accuracy of Inference	High correctness in composite activities	High in predefined labels [227]
Context Awareness	Detailed spatiotemporal context	Limited to training data context [228]
Flexibility and Adaptability	Easily extensible with new concepts	Requires retraining for new activities [229]
Interpretability	High, with transparent rules	Low, often a ”black box” [230]
Query Execution Efficiency	Rapid, supports real-time inference	Slower, may require preprocessing [231]
Behavioral Understanding	Hierarchical and sequence-based	Limited to labeled data [232]

Table 4.15 compares the query efficiency of the proposed ontology-based reasoning framework when applied to the **CASAS Aruba** and **Orange4Home** datasets, focusing on its capacity to handle large-scale data effectively.

For the **CASAS Aruba** dataset, which includes 11 annotated activities, a total of **150 SPARQL queries** were executed. The system achieved an **average query execution time of 0.25 seconds**, with a **maximum query time of 0.35 seconds**. The overall query complexity for this dataset is classified as **medium**, indicating that while the queries are moderately intricate, the system processes them efficiently.

In comparison, the **Orange4Home** dataset—comprising 17 annotated activities—underwent **200 SPARQL queries**. The system recorded an **average query time of 0.20 seconds** and a **maximum of 0.30 seconds**, maintaining a similar classification of **medium** complexity. Despite the increased number of activities, the framework demonstrated consistent and efficient query handling.

Although the method exhibits generally strong performance, several practical considerations must be addressed. The overall success of the framework can be influenced by sensor data quality and the system's ability to adapt across diverse and dynamic real-world environments. Moreover, to align with evolving behavior patterns, the ontology may require periodic updates. In highly dynamic settings, additional optimization techniques could help preserve real-time performance.

Nevertheless, the results from both datasets confirm the system's capacity for efficient reasoning under moderate complexity, with query times consistently remaining within acceptable thresholds. This highlights the framework's suitability for real-time applications in smart environments, especially those requiring fast and reliable processing of extensive and diverse activity data.

In conclusion, the effectiveness of the proposed semantic framework has been validated through empirical experimentation using the CASAS Aruba and Orange4Home datasets. The findings demonstrate clear advantages over traditional approaches in modeling and interpreting the nuanced patterns of human behavior within smart home contexts.

4.5 Modeling human cognition: Hybrid Deep Learning and Ontology-Based Approach to Activity Motivation Inference

In this framework, we use our CNN-MLP model proposed in [200] as a human activity recognition module (Section 4.2). Training of the deep learning was performed with a batch size of 128 and 100 epochs, using early stopping based on validation loss. A learning rate scheduling strategy was applied: the learning rate was initially set to 0.001 and decayed exponentially after 10 epochs (Algorithm 1).

4.5.1 Ontology and Inference

The semantic layer was built using an OWL2 ontology developed in Protégé. The ontology encodes key classes such as:

- `RecognizedActivity`
- `TimeContext`, `LocationContext`, `ObjectInteraction`
- `InferredPurpose` (e.g., SocialInteraction, ManageHealth)

The ontology was loaded and queried using the GENA framework. SPARQL CONSTRUCT queries and custom Java-based rule processors were used to apply the motivational reasoning rules. Inference was triggered dynamically as new activity events were recognized by the deep learning module.

This integrated approach enables context-sensitive interpretation of sensor data, enhancing the system's ability to explain not just *what* the user is doing, but *why* (Algorithm 2).

Algorithm 1 Activity Recognition using CNN-MLP

```
        Sensor time-series data $X$, Labels $Y$
        Normalize input data $X$
        Segment data into fixed-length windows
        Define CNN-MLP model with:Conv1D $\rightarrow$ ReLU $\rightarrow$ BatchNorm
        GlobalAveragePooling
        Dense layer with L2 regularization
        Dropout layer
        Output Dense layer with Softmax
    Train model on $(X, Y)$ using cross-entropy loss
    Optimize using Adam optimizer with learning rate scheduling
    Return trained model $\mathcal{M}$
```

4.5.2 Results and Discussion

To validate our proposed framework, we used CASAS Aruba dataset (Section 4.2.1.4 and Section 4.2.2.2). The results of the deep learning module on the CASAS Aruba dataset are detailed in Section 4.2.4.2. The motivational inference layer was evaluated using manually annotated test cases based on real-world scenarios. In over **90%** of the test cases, the inferred purposes aligned with ground truth human annotations, indicating high semantic interpretability. Table 4.17 presents the inference performance results.

A key contribution of our framework is the semantic reasoning layer, which enhances deep learning with symbolic inference. This addition enables the system not only to recognize *what*

Algorithm 2 Motivation Inference using Ontology and Rules

```
Recognized activity $A$, Location $L$, Time $T$, Ontology $\mathcal{O}$, Rule set $\mathcal{R}$ Load ontology $\mathcal{O}$
and rules $\mathcal{R}$ using GENA Construct SPARQL query with input parameters $(A, L, T)$ Execute
inference engine over $\mathcal{O}$ rule $r_i$ in $\mathcal{R}$ $r_i$ matches activity and context Infer motivational
purpose $M_i$ Assign confidence score $C_i$ Return top-$k$ motivations with highest $C$
```

activity is occurring but also to explain *why*—a crucial feature for real-world, human-centered applications. For example, recognizing the activity ”Lying” on a bed at night, in a low-light setting and aligned with a regular sleep routine, leads to the inference: *“Purpose: Sleeping — consistent with typical behavior.”* However, detecting ”Lying” on the bathroom floor during the day, particularly for an elderly user with a known medical condition, triggers a different and urgent interpretation: *“Purpose: Fall Requiring Assistance — unusual location and time, possible emergency.”* This ability to disambiguate identical activities based on contextual, temporal, and personal cues demonstrates the power of our hybrid approach. It brings interpretability and transparency to HAR systems, enabling intelligent, proactive, and trustworthy responses in sensitive environments such as healthcare and assistive living.

TABLE 4.17: Motivational Rule Accuracy

Motivational Rule	**Correct Inference (%)**	**Incorrect Inference (%)**
Navigate to Location	97.2	2.8
Eat/Drink	94.6	5.4
Personal Hygiene	95.5	4.5
Relax/Entertainment	91.3	8.7
Work/Study	92.9	7.1
Exercise/Train	96.4	3.6
Prepare Meal/Snack	93.0	7.0
Social Interaction	87.5	12.5
Respond to Need	98.2	1.8
Manage Health Condition	95.3	4.7

4.6 Conclusion

This chapter presented a comprehensive evaluation of the proposed hybrid framework for Human Activity Recognition (HAR), highlighting its implementation details, experimental setup, and performance across multiple benchmark datasets. The framework successfully combines deep learning and symbolic reasoning to achieve accurate, explainable, and context-aware activity recognition in both smartphone-based and smart home environments.

At the core of the framework is a novel hybrid deep learning model, designed to extract rich temporal features from sensor data. The model integrates Convolutional Neural Networks (CNN) with Multi-Layer Perceptrons (MLP), enabling effective recognition of complex human activities. Applied to datasets such as UCI HAR, WISDM, PAMAP2, Orange4Home, and CASAS Aruba, this architecture consistently achieved high accuracy, precision, recall, and F1-scores, outperforming several baseline models. The use of rigorous evaluation protocols, including 10-fold cross-validation, confusion matrices, and detailed performance tables, reinforces the robustness and reliability of the results.

Beyond accurate activity classification, this work contributes to the semantic understanding of human behavior. A key component of the framework is the ontology-based reasoning layer, developed in Protégé and powered by the GENA engine. By formulating SPARQL queries and logical rules, the system is capable of high-level inferences, such as identifying abnormal behavior patterns or detecting health-related risks.

Another major contribution lies in the design of an ontological querying framework grounded in cognitive psychology. This framework enables the interpretation of human activities from a cognitive perspective, incorporating theoretical concepts such as intentions, goals, and decision-making processes. Through the modeling of these cognitive aspects, the system extends beyond mere recognition to infer *why* certain actions occur. This is exemplified in the final contribution: modeling human cognition in smart environments. Here, sensor data is linked to symbolic representations of cognitive states, allowing the system to hypothesize about user motivations in various real-life scenarios.

Together, these contributions address the main objectives of the thesis: to develop a robust, interpretable, and cognitively informed system for human activity recognition. The hybrid deep learning and ontology-based architecture proves to be a powerful solution for achieving accurate and meaningful interpretations of human behavior in ubiquitous computing environments.

Chapter 5

General Conclusion

The core ambition of this work has been to push HAR beyond simple pattern recognition—toward systems that can also interpret human behavior from a cognitive and motivational perspective. To achieve this, we developed a hybrid framework that integrates deep learning techniques with semantic reasoning and psychological modeling. This combination not only improves activity recognition performance but also enhances the system's ability to make sense of human actions in context.

Fulfillment of the Thesis Objectives Through Key Contributions

From the beginning, this thesis aimed to create an intelligent system that could do more than recognize physical actions—it sought to understand the reasons behind them. By addressing this complex challenge, we contribute a significant step forward in HAR research, bridging the gap between raw data processing and human-centered interpretation.

The system we propose rests on three main pillars: deep learning, semantic ontologies, and cognitive psychology. At its core, the architecture combines the powerful classification abilities of convolutional and feedforward neural networks with the interpretive strengths of symbolic reasoning. Rather than simply identifying "what" action is happening, the system is designed to explore "why" it is happening—what it might tell us about the person's internal state, intention, or context.

A crucial part of our approach was the use of cognitively inspired ontologies. These ontologies encode psychological constructs such as routine, memory, goal orientation, and motivation. By formalizing these concepts, the system is able to reason semantically about observed behaviors, interpret them within a broader context, and make inferences that align with cognitive principles.

In parallel, the system's deep learning modules were extensively tested on a range of benchmark datasets—including both wearable sensor-based datasets (UCI HAR, WISDM, PAMAP2) and smart home-based environments (CASAS Aruba, Orange4Home). The results confirmed that the hybrid model is not only accurate and generalizable but also enhanced by the inclusion of domain knowledge, improving interpretability without reducing performance.

What distinguishes this work most is its effective unification of symbolic and sub-symbolic paradigms. This integration leads to a more nuanced HAR system—one that offers statistical precision, semantic richness, and psychological awareness. These capabilities are especially important for applications that require explainability, personalization, and sensitivity to user needs, such as elderly care, mental health monitoring, or context-aware smart environments.

Perspectives and Future Directions

The proposed framework reveals important challenges and opportunities for future development. One of the main limitations lies in the manual effort required to align ontologies with sensor data. This process is time-consuming and not easily scalable. Future work should therefore explore ways to automate ontology construction, populate semantic models from real-world data, and adapt knowledge structures dynamically as user behavior evolves.

Another key challenge is dealing with uncertainty, ambiguity, and novelty in behavior sequences. While this work introduced basic mechanisms for temporal reasoning and contextual awareness, more advanced tools—such as probabilistic models, belief revision frameworks, and adaptive inference engines—are needed to ensure the system can function robustly in dynamic, real-world settings. Moreover, integrating multimodal sources of information (e.g., audio, video, physiological signals) could provide a richer foundation for both learning and reasoning.

A final but critical consideration concerns the ethical dimension of interpreting human cognition. Inferring mental states from behavior brings up serious concerns about privacy, consent, and algorithmic bias. Any deployment of such systems must be accompanied by transparent design practices, explainable AI methods, and strong safeguards to protect user autonomy and dignity.

This project represents an initial step in the broader and promising direction of incorporating cognitive psychology into Human Activity Recognition. While the proposed framework introduces novel mechanisms for understanding the motivations behind human actions, it also opens the door for deeper exploration into the cognitive and psychological underpinnings of behavior in smart environments. By moving HAR research toward the study of human cognition, this work lays the groundwork for future systems that are not only context-aware and intelligent, but also capable of interpreting human behavior through the lens of psychological theory. As such, it contributes to a growing shift in HAR—from recognizing actions to understanding the human mind behind them.

Publications

- A. Nadia, S. Lyazid, K. Okba, C. Abdelghani. *Stream Reasoning Approach on Human Behavior for Medication Risks Detection using LARS Framework.* **5th International Conference on Pattern Analysis and Intelligent Systems (PAIS)**, 2023.
- A. Nadia, S. Lyazid, K. Okba, C. Abdelghani. *A CNN-MLP Deep Model for Sensor-Based Human Activity Recognition.* **15th International Conference on Innovations in Information Technology (IIT)**, 2023.
- N. Agti, L. Sabri, O. Kazar, A. Chibani. *A Novel Deep Learning Model for Smartphone-Based Human Activity Recognition.* **International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services**, 2023.
- N. Agti, L. Sabri, O. Kazar. *Hybrid Deep Learning Model for Human Activity Recognition Using Smartphone and Smart Home Data.* **Journal of Ambient Intelligence and Smart Environments**, vol. 17, no. 2, pp. 1–27, 2025.
- N. Agti, L. Sabri, O. Kazar. *A Framework for an Ontological Querying-based Cognitive Perspective for Activity Recognition.* **International Journal of Computers and Their Applications**, vol. 32, no. 2, pp. 112–123, 2025.

Bibliography

[1] Julien Cumin, Grégoire Lefebvre, Fano Ramparany, and James L. Crowley. A dataset of routine daily activities in an instrumented home. In *Ubiquitous Computing and Ambient Intelligence: 11th International Conference, UCAmI 2017*, pages 413–425, 2017.

[2] Ian W Eisenberg, Patrick G Bissett, Jessica R Canning, Jesse Dallery, A Zeynep Enkavi, Susan Whitfield-Gabrieli, Oscar Gonzalez, Alan I Green, Mary Ann Greene, Michaela Kiernan, et al. Applying novel technologies and methods to inform the ontology of self-regulation. *Behaviour research and therapy*, 101:46–57, 2018.

[3] Juan-Miguel López-Gil, Rosa Gil, and Roberto García. Web ontologies to categorialy structure reality: Representations of human emotional, cognitive, and motivational processes. *Frontiers in psychology*, 7:551, 2016.

[4] Bénédicte Batrancourt, Michel Dojat, Bernard Gibaud, and Gilles Kassel. A multilayer ontology of instruments for neurological, behavioral and cognitive assessments. *Neuroinformatics*, 13:93–110, 2015.

[5] Charmi Jobanputra, Jatna Bavishi, and Nishant Doshi. Human activity recognition: A survey. *Procedia Computer Science*, 155:698–703, 2019.

[6] Liming Chen, Jesse Hoey, Chris D Nugent, Diane J Cook, and Zhiwen Yu. Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6):790–808, 2012.

[7] Liming Chen and Chris D Nugent. *Human activity recognition and behaviour analysis.* Springer, 2019.

[8] Mohd Halim Mohd Noor, Zoran Salcic, I Kevin, and Kai Wang. Enhancing ontological reasoning with uncertainty handling for activity recognition. *Knowledge-Based Systems*, 114:47–60, 2016.

[9] Liming Chen and Chris Nugent. Ontology-based activity recognition in intelligent pervasive environments. *International Journal of Web Information Systems*, 5(4):410–430, 2009.

[10] Daniele Riboni and Claudio Bettini. Context-aware activity recognition through a combination of ontological and statistical reasoning. In *International Conference on Ubiquitous Intelligence and Computing*, pages 39–53. Springer, 2009.

[11] Qin Ni, Ana Belén García Hernando, and Iván Pau de la Cruz. A context-aware system infrastructure for monitoring activities of daily living in smart home. *Journal of Sensors*, 2016, 2016.

[12] Serge Thomas Mickala Bourobou and Younghwan Yoo. User activity recognition in smart homes using pattern clustering applied to temporal ann algorithm. *Sensors*, 15(5):11953–11971, 2015.

[13] Natalia Díaz Rodríguez, Manuel P Cuéllar, Johan Lilius, and Miguel Delgado Calvo-Flores. A fuzzy ontology for semantic modelling and recognition of human behaviour. *Knowledge-Based Systems*, 66:46–60, 2014.

[14] Vito Magnanimo, Matteo Saveriano, Silvia Rossi, and Dongheui Lee. A bayesian approach for task recognition and future human activity prediction. In *The 23rd IEEE international symposium on robot and human interactive communication*, pages 726–731. IEEE, 2014.

[15] Qingjuan Li, Huansheng Ning, Tao Zhu, Shan Cui, and Liming Chen. A hybrid approach to inferring the internet of things for complex activity recognition. *EURASIP Journal on Wireless Communications and Networking*, 2019(1):1–11, 2019.

[16] Prabhat Kumar and S Suresh. Recurrenthar: A novel transfer learning-based deep learning model for sequential, complex, concurrent, interleaved, and heterogeneous type human activity recognition. *IETE Technical Review*, 40(3):312–333, 2023.

[17] Prabhat Kumar and S Suresh. Deep learning models for recognizing the simple human activities using smartphone accelerometer sensor. *IETE Journal of Research*, 69(8):5148–5158, 2023.

[18] Shilpa Ankalaki. Simple to complex, single to concurrent sensor based human activity recognition: Perception and open challenges. *IEEE Access*, 2024.

[19] Abisek Dahal, Soumen Moulik, and Rohan Mukherjee. Stack-har: Complex human activity recognition with stacking-based ensemble learning framework. *IEEE Sensors Journal*, 2025.

[20] Lu Bai, Chris Yeung, Christos Efstratiou, and Moyra Chikomo. Motion2vector: Unsupervised learning in human activity recognition using wrist-sensing data. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, pages 537–542, 2019.

[21] Fuqiang Gu, Kourosh Khoshelham, Shahrokh Valaee, Jianga Shang, and Rui Zhang. Locomotion activity recognition using stacked denoising autoencoders. *IEEE Internet of Things Journal*, 5(3):2085–2093, 2018.

[22] Jiwei Wang, Yiqiang Chen, Yang Gu, Yunlong Xiao, and Haonan Pan. Sensorygans: An effective generative adversarial framework for sensor-based human activity recognition. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2018.

[23] Ali A Alani, Georgina Cosma, and Aboozar Taherkhani. Classifying imbalanced multimodal sensor data for human activity recognition in a smart home using deep learning. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2020.

[24] Rene Grzeszick, Jan Marius Lenk, Fernando Moya Rueda, Gernot A Fink, Sascha Feldhorst, and Michael ten Hompel. Deep neural network based human activity recognition for the order picking process. In *Proceedings of the 4th international Workshop on Sensor-based Activity Recognition and Interaction*, pages 1–6, 2017.

[25] Kaixuan Chen, Lina Yao, Dalin Zhang, Xianzhi Wang, Xiaojun Chang, and Feiping Nie. A semisupervised recurrent convolutional attention model for human activity recognition. *IEEE transactions on neural networks and learning systems*, 31(5):1747–1756, 2019.

[26] Yu Guan and Thomas Plötz. Ensembles of deep lstm learners for activity recognition using wearables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(2):1–28, 2017.

[27] Praneeth Vepakomma, Debraj De, Sajal K Das, and Shekhar Bhansali. A-wristocracy: Deep learning on wrist-worn sensing for recognition of user complex activities. In *2015 IEEE 12th International conference on wearable and implantable body sensor networks (BSN)*, pages 1–6. IEEE, 2015.

[28] Liangying Peng, Ling Chen, Zhenan Ye, and Yi Zhang. Aroma: A deep multi-task learning based simple and complex human activity recognition method using wearable sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2 (2):1–16, 2018.

[29] Weihao Cheng, Sarah M Erfani, Rui Zhang, and Ramamohanarao Kotagiri. Predicting complex activities from ongoing multivariate time series. In *IJCAI*, pages 3322–3328, 2018.

[30] Jun-Ho Choi and Jong-Seok Lee. Confidence-based deep multimodal fusion for activity recognition. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, pages 1548–1556, 2018.

[31] Shehroz S Khan and Babak Taati. Detecting unseen falls from wearable devices using channel-wise ensemble of autoencoders. *Expert Systems with Applications*, 87:280–290, 2017.

[32] Elnaz Soleimani and Ehsan Nazerfard. Cross-subject transfer learning in human activity recognition systems using generative adversarial networks. *Neurocomputing*, 426:26–34, 2021.

[33] Ling Chen, Yi Zhang, and Liangying Peng. Metier: A deep multi-task learning based activity and user recognition model using wearable sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(1):1–18, 2020.

[34] Yanyi Zhang, Xinyu Li, Jianyu Zhang, Shuhong Chen, Moliang Zhou, Richard A Farneth, Ivan Marsic, and Randall S Burd. Car-a deep learning structure for concurrent activity recognition. In *2017 16th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 299–300. IEEE, 2017.

[35] Silvia Rossi, Roberto Capasso, Giovanni Acampora, and Mariacarla Staffa. A multimodal deep learning network for group activity recognition. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–6. IEEE, 2018.

[36] Dipanwita Thakur, Antonella Guzzo, and Giancarlo Fortino. Intelligent adaptive real-time monitoring and recognition system for human activities. *IEEE Transactions on Industrial Informatics*, 2024.

[37] ukasz Czekaj, Mateusz Kowalewski, Jakub Domaszewicz, Robert Kitłowski, Mariusz Szwoch, and Włodzisław Duch. Real-time sensor-based human activity recognition for efitness and ehealth platforms. *Sensors*, 24(12):3891, 2024.

[38] Mohammad Malekzadeh, Richard G Clegg, Andrea Cavallaro, and Hamed Haddadi. Mobile sensor data anonymization. In *Proceedings of the international conference on internet of things design and implementation*, pages 49–58, 2019.

[39] Dalin Zhang, Lina Yao, Kaixuan Chen, Guodong Long, and Sen Wang. Collective protection: Preventing sensitive inferences via integrative transformation. In *2019 IEEE International Conference on Data Mining (ICDM)*, pages 1498–1503. IEEE, 2019.

[40] Kun Wang, Jun He, and Lei Zhang. Attention-based convolutional neural network for weakly labeled human activities' recognition with wearable sensors. *IEEE Sensors Journal*, 19(17):7598–7604, 2019.

[41] Ali Akbari, Jian Wu, Reese Grimsley, and Roozbeh Jafari. Hierarchical signal segmentation and classification for accurate activity recognition. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, pages 1596–1605, 2018.

[42] Yong Zhang, Yu Zhang, Zhao Zhang, Jie Bao, and Yunpeng Song. Human activity recognition based on time series analysis using u-net. *arXiv preprint arXiv:1809.08113*, 2018.

[43] Darpan Triboan, Liming Chen, Feng Chen, and Zumin Wang. Semantic segmentation of real-time sensor data stream for complex activity recognition. *Personal and Ubiquitous Computing*, 21(3):411–425, 2017.

[44] Darpan Triboan, Liming Chen, Feng Chen, and Zumin Wang. A semantics-based approach to sensor data segmentation in real-time activity recognition. *Future Generation Computer Systems*, 93:224–236, 2019.

[45] Dalin Zhang, Lina Yao, Kaixuan Chen, Sen Wang, Pari Delir Haghighi, and Caley Sullivan. A graph-based hierarchical attention model for movement intention detection from eeg signals. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(11): 2247–2253, 2019.

[46] Mohd Halim Mohd Noor, Zoran Salcic, and Kevin I-Kai Wang. Ontology-based sensor fusion activity recognition. *Journal of Ambient Intelligence and Humanized Computing*, 11(8):3073–3087, 2020.

[47] Toan H Vu, An Dang, Le Dung, and Jia-Ching Wang. Self-gated recurrent neural networks for human activity recognition on wearable devices. In *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*, pages 179–185, 2017.

[48] Zhan Yang, Osolo Ian Raymond, Chengyuan Zhang, Ying Wan, and Jun Long. Dfternet: Towards 2-bit dynamic fusion networks for accurate human activity recognition. *IEEE Access*, 6:56750–56764, 2018.

[49] Dheeru Dua and Casey Graff. UCI machine learning repository, 2017. URL `https://archive.ics.uci.edu/ml/index.php`.

[50] Human activity recognition using smartphones. `https://archive.ics.uci.edu/dataset/240/human+activity+recognition+using+smartphones`, 2012. Accessed: 2023-10-11.

[51] Wisdm smartphone and smartwatch activity and biometrics dataset. `https://archive.ics.uci.edu/dataset/507/wisdm+smartphone+and+smartwatch+activity+and+biometrics+dataset`, 2019. Accessed: 2023-10-11.

[52] Pamap2 physical activity monitoring. `https://archive.ics.uci.edu/dataset/231/pamap2+physical+activity+monitoring`, 2012. Accessed: 2023-10-11.

[53] Mhealth dataset. `https://archive.ics.uci.edu/dataset/319/mhealth+dataset`, 2014. Accessed: 2023-09-23.

[54] Thomas Stiefmeier, Daniel Roggen, Gerhard Tröster, Georg Ogris, and Paul Lukowicz. Wearable activity tracking in car manufacturing. *IEEE Pervasive Computing*, 7(2):42–50, May 2008. doi: 10.1109/MPRV.2008.36.

[55] O. Catal, C. Tunca, H. Ertan, and T. Ozcelebi. Heterogeneity activity recognition dataset. `https://archive.ics.uci.edu/dataset/505/heterogeneity+activity+recognition`, 2018. UCI Machine Learning Repository.

[56] Ricardo Chavarriaga, Hesam Sagha, Alberto Calatroni, Sundara Tejaswi Digumarti, Gerhard Tröster, José del R. Millán, and Daniel Roggen. The opportunity challenge: A benchmark database for on-body sensor-based activity recognition. *Pattern Recognition Letters*, 34(15):2033–2042, 2013. ISSN 0167-8655. doi: 10.1016/j.patrec.2012.12.014. URL `https://www.sciencedirect.com/science/article/pii/S0167865512004205`. Smart Approaches for Human Action Recognition.

[57] Casas datasets. `https://casas.wsu.edu/datasets/`. Accessed: 2023-12-11.

[58] Emmanuel Munguia Tapia, Stephen S Intille, and Kent Larson. Activity recognition in the home using simple and ubiquitous sensors. In *International conference on pervasive computing*, pages 158–175. Springer, 2004.

[59] Nobuo Kawaguchi, Hodaka Watanabe, Tianhui Yang, Nobuhiro Ogawa, Yohei Iwasaki, Katsuhiko Kaji, Tsutomu Terada, Kazuya Murao, Hisakazu Hada, Sozo Inoue, et al. Hasc2012corpus: Large scale human activity corpus and its application. In *Proceedings of the second international workshop of mobile sensing: from smartphones and wearables to big data*, pages 10–14, 2012.

[60] Daniel Roggen, Alberto Calatroni, Mirco Rossi, Thomas Holleczek, Kilian Förster, Gerhard Tröster, Paul Lukowicz, David Bannach, Gerald Pirkl, Alois Ferscha, et al. Collecting complex activity datasets in highly rich networked sensor environments. In *2010 Seventh international conference on networked sensing systems (INSS)*, pages 233–240. IEEE, 2010.

[61] Diane J Cook. Learning setting-generalized activity models for smart spaces. *IEEE Intelligent Systems*, 2010(99):1, 2010.

[62] Activities of daily living (adls) recognition using binary sensors. `https://archive.ics.uci.edu/dataset/271/activities+of+daily+living+adls+recognition+using+binary+sensors`, 2013. Accessed: 2023-10-05.

[63] Jorge-L Reyes-Ortiz, Luca Oneto, Albert Samà, Xavier Parra, and Davide Anguita. Transition-aware human activity recognition using smartphones. *Neurocomputing*, 171: 754–767, 2016.

[64] Enrique A de la Cal, Mirko Fáñez, Mario Villar, Jose R Villar, and Víctor M González. A low-power har method for fall and high-intensity adls identification using wrist-worn accelerometer devices. *Logic Journal of the IGPL*, 31(2):375–389, 2023.

[65] Po Yang, Congmin Yang, Vitaveska Lanfranchi, and Fabio Ciravegna. Activity graph based convolutional neural network for human activity recognition using acceleration and gyroscope data. *IEEE Transactions on Industrial Informatics*, 18(10):6619–6630, 2022.

[66] Mitchell Webber and Raul Fernandez Rojas. Human activity recognition with accelerometer and gyroscope: A data fusion approach. *IEEE Sensors Journal*, 21(15):16979–16989, 2021.

[67] Phuc Huu Truong, Sujeong You, Sang-Hoon Ji, and Gu-Min Jeong. Wearable system for daily activity recognition using inertial and pressure sensors of a smart band and smart shoes. *International Journal of Computers Communications & Control*, 14(6):726–742, 2019.

[68] Ömer Faruk İnce, Ibrahim Furkan Ince, Mustafa Eren Yıldırım, Jang Sik Park, Jong Kwan Song, and Byung Woo Yoon. Human activity recognition with analysis of angles between skeletal joints using a rgb-depth sensor. *ETRI journal*, 42(1):78–89, 2020.

[69] Xi Chen, Julien Cumin, Fano Ramparany, and Dominique Vaufreydaz. Towards llm-powered ambient sensor based multi-person human activity recognition. In *2024 IEEE 30th International Conference on Parallel and Distributed Systems (ICPADS)*, pages 609–616. IEEE, 2024.

[70] Leo Breiman. Random forests. *Machine learning*, 45:5–32, 2001.

[71] Sarah Fallmann and Johannes Kropf. Human activity recognition of continuous data using hidden markov models and the aspect of including discrete data. In *2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld)*, pages 121–126. IEEE, 2016.

[72] Lukas Köping, Kimiaki Shirahama, and Marcin Grzegorzek. A general framework for sensor-based human activity recognition. *Computers in biology and medicine*, 95:248–260, 2018.

[73] KG Manosha Chathuramali and Ranga Rodrigo. Faster human activity recognition with svm. In *International conference on advances in ICT for emerging regions (ICTer2012)*, pages 197–203. IEEE, 2012.

[74] Chunyu Hu, Yiqiang Chen, Lisha Hu, and Xiaohui Peng. A novel random forests based class incremental learning method for activity recognition. *Pattern Recognition*, 78:277–290, 2018.

[75] Pichleap Sok, Ting Xiao, Yohannes Azeze, Arun Jayaraman, and Mark V Albert. Activity recognition for incomplete spinal cord injury subjects using hidden markov models. *IEEE Sensors Journal*, 18(15):6369–6374, 2018.

[76] Charissa Ann Ronao and Sung-Bae Cho. Recognizing human activities from smartphone sensors using hierarchical continuous hidden markov models. *International Journal of Distributed Sensor Networks*, 13(1):1550147716683687, 2017.

[77] Bilal M'hamed Abidine, Lamya Fergani, Belkacem Fergani, and Mourad Oussalah. The joint use of sequence features combination and modified weighted svm for improving daily activity recognition. *Pattern Analysis and Applications*, 21(1):119–138, 2018.

[78] Zhangjie Chen and Ya Wang. Infrared–ultrasonic sensor fusion for support vector machine–based fall detection. *Journal of Intelligent Material Systems and Structures*, 29(9):2027–2039, 2018.

[79] Carlos Avilés-Cruz, Andrés Ferreyra-Ramírez, Arturo Zúñiga-López, and Juan Villegas-Cortéz. Coarse-fine convolutional deep-learning strategy for human activity recognition. *Sensors*, 19(7):1556, 2019.

[80] Yuqing Chen and Yang Xue. A deep learning approach to human activity recognition based on single accelerometer. In *2015 IEEE international conference on systems, man, and cybernetics*, pages 1488–1492. IEEE, 2015.

[81] Md Abdullah Al Hafiz Khan, Nirmalya Roy, and Archan Misra. Scaling human activity recognition via deep learning-based domain adaptation. In *2018 IEEE international conference on pervasive computing and communications (PerCom)*, pages 1–9. IEEE, 2018.

[82] Bandar Almaslukh, Abdel Monim Artoli, and Jalal Al-Muhtadi. A robust deep learning approach for position-independent smartphone-based human activity recognition. *Sensors*, 18(11):3726, 2018.

[83] Ran Zhu, Zhuoling Xiao, Ying Li, Mingkun Yang, Yawen Tan, Liang Zhou, Shuisheng Lin, and Hongkai Wen. Efficient human activity recognition solving the confusing activities via deep ensemble learning. *Ieee Access*, 7:75490–75499, 2019.

[84] Shaohua Wan, Lianyong Qi, Xiaolong Xu, Chao Tong, and Zonghua Gu. Deep learning models for real-time human activity recognition with smartphones. *Mobile Networks and Applications*, 25(2):743–755, 2020.

[85] Yin Tang, Qi Teng, Lei Zhang, Fuhong Min, and Jun He. Layer-wise training convolutional neural networks with smaller filters for human activity recognition using wearable sensors. *IEEE Sensors Journal*, 21(1):581–592, 2020.

[86] Wenbin Gao, Lei Zhang, Qi Teng, Jun He, and Hao Wu. Danhar: Dual attention network for multimodal human activity recognition using wearable sensors. *Applied Soft Computing*, 111:107728, 2021.

[87] Chi Ian Tang, Ignacio Perez-Pozuelo, Dimitris Spathis, Soren Brage, Nick Wareham, and Cecilia Mascolo. Selfhar: Improving human activity recognition through self-training with unlabeled data. *arXiv preprint arXiv:2102.06073*, 2021.

[88] Nafiul Rashid, Berken Utku Demirel, and Mohammad Abdullah Al Faruque. Ahar: Adaptive cnn for energy-efficient human activity recognition in low-power edge devices. *IEEE Internet of Things Journal*, 2022.

[89] Xin Cheng, Lei Zhang, Yin Tang, Yue Liu, Hao Wu, and Jun He. Real-time human activity recognition using conditionally parametrized convolutions on mobile and wearable devices. *IEEE Sensors Journal*, 2022.

[90] Bin Xie and Qing Zhang. Deep filtering with dnn, cnn and rnn. *arXiv preprint arXiv:2112.12616*, 2021.

[91] Sungpil Woo, Jaewook Byun, Seonghoon Kim, Hoang Minh Nguyen, Janggwan Im, and Daeyoung Kim. Rnn-based personalized activity recognition in multi-person environment using rfid. In *2016 IEEE International Conference on Computer and Information Technology (CIT)*, pages 708–715. IEEE, 2016.

[92] Francisco Javier Ordóñez and Daniel Roggen. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1):115, 2016.

[93] Marcus Edel and Enrico Köppe. Binarized-blstm-rnn based human activity recognition. In *2016 International conference on indoor positioning and indoor navigation (IPIN)*, pages 1–7. IEEE, 2016.

[94] Abdulmajid Murad and Jae-Young Pyun. Deep recurrent neural networks for human activity recognition. *Sensors*, 17(11):2556, 2017.

[95] Jun He, Qian Zhang, Liqun Wang, and Ling Pei. Weakly supervised human activity recognition from wearable sensors by recurrent attention learning. *IEEE Sensors Journal*, 19(6):2287–2297, 2018.

[96] Fadi Al Machot, Suneth Ranasinghe, Johanna Plattner, and Nour Jnoub. Human activity recognition based on real life scenarios. In *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 3–8. IEEE, 2018.

[97] Yu Zhao, Rennong Yang, Guillaume Chevalier, Ximeng Xu, and Zhenxing Zhang. Deep residual bidir-lstm for human activity recognition using wearable sensors. *Mathematical Problems in Engineering*, 2018, 2018.

[98] LuKun Wang and RuYue Liu. Human activity recognition based on wearable sensor using hierarchical deep lstm networks. *Circuits, Systems, and Signal Processing*, 39(2):837–856, 2020.

[99] Xiaokang Zhou, Wei Liang, I Kevin, Kai Wang, Hao Wang, Laurence T Yang, and Qun Jin. Deep-learning-enhanced human activity recognition for internet of healthcare things. *IEEE Internet of Things Journal*, 7(7):6429–6438, 2020.

[100] Jian Sun, Yongling Fu, Shengguang Li, Jie He, Cheng Xu, and Lin Tan. Sequential human activity recognition based on deep convolutional network and extreme learning machine using wearable sensors. *Journal of Sensors*, 2018, 2018.

[101] Tongtong Su, Huazhi Sun, Chunmei Ma, Lifen Jiang, and Tongtong Xu. Hdl: Hierarchical deep learning model based human activity recognition using smartphone sensors. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2019.

[102] Kun Xia, Jianguang Huang, and Hanyu Wang. Lstm-cnn architecture for human activity recognition. *IEEE Access*, 8:56855–56866, 2020.

[103] Ronald Mutegeki and Dong Seog Han. A cnn-lstm approach to human activity recognition. In *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, pages 362–366. IEEE, 2020.

[104] Abdu Gumaei, Mohammad Mehedi Hassan, Abdulhameed Alelaiwi, and Hussain Alsalman. A hybrid deep learning model for human activity recognition using multimodal body sensing data. *IEEE Access*, 7:99152–99160, 2019.

[105] Han Zou, Yuxun Zhou, Jianfei Yang, Hao Jiang, Lihua Xie, and Costas J Spanos. Deepsense: Device-free human activity recognition via autoencoder long-term recurrent convolutional network. In *2018 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2018.

[106] Yuta Yuki, Junto Nozaki, Kei Hiroi, Katsuhiko Kaji, and Nobuo Kawaguchi. Activity recognition using dual-convlstm extracting local and global features for shl recognition

challenge. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, pages 1643–1651, 2018.

[107] Nadia Agti, Lyazid Sabri, Okba Kazar, and Abdelghani Chibani. A novel deep learning model for smartphone-based human activity recognition. In *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*, pages 231–243. Springer, 2023.

[108] Nadia Agti, Lyazid Sabri, Okba Kazar, and Abdelghani Chibani. A cnn-mlp deep model for sensor-based human activity recognition. In *2023 15th International Conference on Innovations in Information Technology (IIT)*, pages 121–126, 2023.

[109] Naoharu Yamada, Kenji Sakamoto, Goro Kunito, Yoshinori Isoda, Kenichi Yamazaki, and Satoshi Tanaka. Applying ontology and probabilistic model to human activity recognition from surrounding things. *IPSJ Digital Courier*, 3:506–517, 2007.

[110] Juan Ye, Graeme Stevenson, and Simon Dobson. Usmart: An unsupervised semantic mining activity recognition technique. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 4(4):1–27, 2014.

[111] Liming Chen, Chris Nugent, and George Okeyo. An ontology-based hybrid approach to activity modeling for smart homes. *IEEE Transactions on human-machine systems*, 44 (1):92–105, 2013.

[112] KS Gayathri, Susan Elias, and S Shivashankar. Composite activity recognition in smart homes using markov logic network. In *2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom)*, pages 46–53. IEEE, 2015.

[113] Daniele Riboni, Claudio Bettini, Gabriele Civitarese, Zaffar Haider Janjua, and Rim Helaoui. Fine-grained recognition of abnormal behaviors for early detection of mild cognitive impairment. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 149–154. IEEE, 2015.

[114] Simin Ahmadi-Karvigh, Ali Ghahramani, Burcin Becerik-Gerber, and Lucio Soibelman. Real-time activity recognition for energy efficiency in buildings. *Applied energy*, 211: 146–160, 2018.

[115] KS Gayathri, KS Easwarakumar, and Susan Elias. Probabilistic ontology based activity recognition in smart homes using markov logic network. *Knowledge-Based Systems*, 121: 173–184, 2017.

[116] Gabriele Civitarese, Claudio Bettini, Timo Sztyler, Daniele Riboni, and Heiner Stuckenschmidt. Nectar: Knowledge-based collaborative active learning for activity recognition. In *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 1–10. IEEE, 2018.

[117] Roghayeh Mojarad, Ferhat Attal, Abdelghani Chibani, Sandro Rama Fiorini, and Yacine Amirat. Hybrid approach for human activity recognition by ubiquitous robots. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5660–5665. IEEE, 2018.

[118] Gabriele Civitarese, Riccardo Presotto, and Claudio Bettini. Hybrid data-driven and context-aware activity recognition with mobile devices. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, pages 266–267, 2019.

[119] Fernando Moya Rueda, Stefan Lüdtke, Max Schröder, Kristina Yordanova, Thomas Kirste, and Gernot A Fink. Combining symbolic reasoning and deep learning for human activity recognition. In *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 22–27. IEEE, 2019.

[120] Abdul Syafiq Abdull Sukor, Ammar Zakaria, Norasmadi Abdul Rahim, Latifah Munirah Kamarudin, Rossi Setchi, and Hiromitsu Nishizaki. A hybrid approach of knowledge-driven and data-driven reasoning for activity recognition in smart homes. *Journal of Intelligent & Fuzzy Systems*, 36(5):4177–4188, 2019.

[121] Gabriele Civitarese, Claudio Bettini, Timo Sztyler, Daniele Riboni, and Heiner Stuckenschmidt. newnectar: Collaborative active learning for knowledge-based probabilistic activity recognition. *Pervasive and Mobile Computing*, 56:88–105, 2019.

[122] Gabriele Civitarese, Riccardo Presotto, and Claudio Bettini. Context-driven active and incremental activity recognition. *arXiv preprint arXiv:1906.03033*, 2019.

[123] Roghayeh Mojarad, Ferhat Attal, Abdelghani Chibani, and Yacine Amirat. A context-aware hybrid framework for human behavior analysis. In *2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 460–465. IEEE, 2020.

[124] Claudio Bettini, Gabriele Civitarese, and Riccardo Presotto. Caviar: Context-driven active and incremental activity recognition. *Knowledge-Based Systems*, 196:105816, 2020.

[125] Daniele Riboni and Flavia Murru. Unsupervised recognition of multi-resident activities in smart-homes. *IEEE Access*, 8:201985–201994, 2020.

[126] Claudio Bettini, Gabriele Civitarese, and Michele Fiori. Explainable activity recognition over interpretable models. In *2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, pages 32–37. IEEE, 2021.

[127] Kaixuan Chen, Dalin Zhang, Lina Yao, Bin Guo, Zhiwen Yu, and Yunhao Liu. Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities. *ACM Computing Surveys (CSUR)*, 54(4):1–40, 2021.

[128] Nils Y Hammerla, Shane Halloran, and Thomas Plötz. Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880*, 2016.

[129] Ali Akbari and Roozbeh Jafari. Transferring activity recognition models for new wearable sensors with deep generative domain adaptation. In *Proceedings of the 18th International Conference on Information Processing in Sensor Networks*, pages 85–96, 2019.

[130] Gautham Krishna Gudur, Prahalathan Sundaramoorthy, and Venkatesh Umaashankar. Activeharnet: Towards on-device deep bayesian active learning for human activity recognition. In *The 3rd International Workshop on Deep Learning for Mobile Systems and Applications*, pages 7–12, 2019.

[131] Francesco Rea, Samia Nefti-Meziani, Umar Manzoor, and Steve Davis. Ontology enhancing process for a situated and curiosity-driven robot. *Robotics and Autonomous Systems*, 62(12):1837–1847, 2014.

[132] Zhi Zeng and Qiang Ji. Knowledge based activity recognition with dynamic bayesian network. In *European conference on computer vision*, pages 532–546. Springer, 2010.

[133] Liming Chen, Chris D Nugent, and Hui Wang. A knowledge-driven approach to activity recognition in smart homes. *IEEE Transactions on Knowledge and Data Engineering*, 24 (6):961–974, 2011.

[134] George Okeyo, Liming Chen, Hui Wang, and Roy Sterritt. A knowledge-driven approach to composite activity recognition in smart environments. In *International Conference on Ubiquitous Computing and Ambient Intelligence*, pages 322–329. Springer, 2012.

[135] Ihn-Han Bae. An ontology-based approach to adl recognition in smart homes. *Future Generation Computer Systems*, 33:32–41, 2014.

[136] Juan Ye, Graeme Stevenson, and Simon Dobson. Kcar: A knowledge-driven approach for concurrent activity recognition. *Pervasive and Mobile Computing*, 19:47–70, 2015.

[137] G. Meditskos, S. Dasiopoulou, and I. Kompatsiaris. Metaq: A knowledge-driven framework for context-aware activity recognition combining sparql and owl 2 activity patterns. *Pervasive and Mobile Computing*, 25:104–124, 2016.

[138] Asad Masood Khattak, Wajahat Ali Khan, Zeeshan Pervez, Farkhund Iqbal, and Sungyoung Lee. Towards a self adaptive system for social wellness. *Sensors*, 16(4):531, 2016.

[139] Simin Ahmadi-Karvigh, Burcin Becerik-Gerber, and Lucio Soibelman. A framework for allocating personalized appliance-level disaggregated electricity consumption to daily activities. *Energy and Buildings*, 111:337–350, 2016.

[140] Muhammad Safyan, Zia Ul Qayyum, Sohail Sarwar, Muddesar Iqbal, Raul Garcia Castro, and Anwer Al-Dulaimi. Ontology evolution for personalised and adaptive activity recognition. *IET Wireless Sensor Systems*, 9(4):193–200, 2019.

[141] KS Gayathri, KS Easwarakumar, and Susan Elias. Fuzzy ontology based activity recognition for assistive health care using smart home. *International Journal of Intelligent Information Technologies (IJIIT)*, 16(1):17–31, 2020.

[142] Yaqing Liu, Xiangxin Wang, Zhengguo Zhai, Rong Chen, Bin Zhang, and Yu Jiang. Timely daily activity recognition from headmost sensor events. *ISA transactions*, 94: 379–390, 2019.

[143] Diksha Hooda and Rinkle Rani. Ontology driven human activity recognition in heterogeneous sensor measurements. *Journal of Ambient Intelligence and Humanized Computing*, 11(12):5947–5960, 2020.

[144] Daniele Riboni and Claudio Bettini. Cosar: hybrid reasoning for context-aware activity recognition. *Personal and Ubiquitous Computing*, 15(3):271–289, 2011.

[145] Gabriele Civitarese, Timo Sztyler, Daniele Riboni, Claudio Bettini, and Heiner Stuckenschmidt. Polaris: Probabilistic and ontological activity recognition in smart-homes. *IEEE Transactions on Knowledge and Data Engineering*, 33(1):209–223, 2019.

[146] Roghayeh Mojarad, Ferhat Attal, Abdelghani Chibani, and Yacine Amirat. A context-aware approach to detect abnormal human behaviors. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 89–104. Springer, 2020.

[147] Nikos Katzouris, Georgios Paliouras, and Alexander Artikis. Online learning probabilistic event calculus theories in answer set programming. *Theory and Practice of Logic Programming*, 23(2):362–386, 2023.

[148] Pascal Nicolas, Laurent Garcia, Igor Stéphan, and Claire Lefèvre. Possibilistic uncertainty handling for answer set programming. *Annals of Mathematics and Artificial Intelligence*, 47:139–181, 2006.

[149] Claudio Bettini, Gabriele Civitarese, Davide Giancane, and Riccardo Presotto. Procaviar: Hybrid data-driven and probabilistic knowledge-based activity recognition. *IEEE Access*, 8:146876–146886, 2020.

[150] SK MANGAL and SHUBHRA MANGAL. *COGNITIVE PSYCHOLOGY*. PHI Learning Pvt. Ltd., 2024.

[151] George A Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

[152] John R Anderson, Michael Matessa, and Christian Lebiere. Act-r: A theory of higher level cognition and its relation to visual attention. In *Human Attention and Performance*, pages 145–190. MIT Press, 1996.

[153] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.

[154] Philip N Johnson-Laird. *Mental Models: Toward a Cognitive Science of Language, Inference, and Consciousness.* Harvard University Press, 1983.

[155] Thomas R Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993.

[156] Jerry A Fodor. *The Modularity of Mind: An Essay on Faculty Psychology.* MIT Press, 1983.

[157] Donald A Norman. Cognitive engineering. *User Centered System Design*, 31:61, 1986.

[158] Allen Newell and Herbert A Simon. Human problem solving. In *Prentice-Hall series in automatic computation.* Prentice-Hall, 1972.

[159] Daniel Kahneman. *Thinking, Fast and Slow.* Farrar, Straus and Giroux, 2011.

[160] Andy Clark. *Being There: Putting Brain, Body, and World Together Again.* MIT Press, 1998.

[161] Paul Dourish. *Where the Action Is: The Foundations of Embodied Interaction.* MIT Press, 2004.

[162] Rosalind W Picard. *Affective Computing.* MIT Press, 1997.

[163] Donald A Norman. *The Design of Everyday Things.* Basic Books, 2002.

[164] Fred H. Previc. Functional specialization in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications. *Behavioral and Brain Sciences*, 13(3):519–542, 1990. doi: 10.1017/S0140525X00080018.

[165] Edward L. Deci. *Conceptualizations of Intrinsic Motivation*, pages 23–63. Springer US, Boston, MA, 1975. ISBN 978-1-4613-4446-9. doi: 10.1007/978-1-4613-4446-9_2. URL https://doi.org/10.1007/978-1-4613-4446-9_2.

[166] John E Laird. An analysis and comparison of act-r and soar. *arXiv preprint arXiv:2201.09305*, 2022. URL https://arxiv.org/abs/2201.09305.

[167] Zohour Salem and Alex P Weiss. Improved spatiotemporal framework for human activity recognition in smart environment. *Sensors*, 23(1):132, 2023. doi: 10.3390/s23010132. URL https://www.mdpi.com/1424-8220/23/1/132.

[168] Anind K Dey. Understanding and using context. *Personal and ubiquitous computing*, 5 (1):4–7, 2001.

[169] Maryam Ziaeefard and Robert Bergevin. Semantic human activity recognition: A literature review. *Pattern Recognition*, 48(8):2329–2345, 2015.

[170] Arvind Arasu, Shivnath Babu, and Jennifer Widom. The cql continuous query language: semantic foundations and query execution. *The VLDB Journal*, 15:121–142, 2006.

[171] Harald Beck, Minh Dao-Tran, Thomas Eiter, and Christian Folie. Stream reasoning with lars. *KI-Künstliche Intelligenz*, 32:193–195, 2018.

[172] Piero Bonatti, Francesco Calimeri, Nicola Leone, and Francesco Ricca. Answer set programming. *A 25-Year Perspective on Logic Programming: Achievements of the Italian Association for Logic Programming, GULP*, pages 159–182, 2010.

[173] Agti Nadia, Sabri Lyazid, Kazar Okba, and Chibani Abdelghani. Stream reasoning approach on human behavior for medication risks detection using lars framework. In *2023 5th International Conference on Pattern Analysis and Intelligent Systems (PAIS)*, pages 1–7, 2023. doi: 10.1109/PAIS60821.2023.10322064.

[174] Jiangpeng Dai, Xiaole Bai, Zhimin Yang, Zhaohui Shen, and Dong Xuan. Perfalld: A pervasive fall detection system using mobile phones. In *2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pages 292–297. IEEE, 2010.

[175] Malik Tabish, Zahoor-ur-Rehman Tanooli, and Muhammad Shaheen. Activity recognition framework in sports videos. *Multimedia Tools and Applications*, pages 1–23, 2024.

[176] Hui Chen, Charles Gouin-Vallerand, Kévin Bouchard, Sébastien Gaboury, Mélanie Couture, Nathalie Bier, and Sylvain Giroux. Enhancing human activity recognition in smart homes with self-supervised learning and self-attention. *Sensors*, 24(3):884, 2024.

[177] Saisakul Chernbumroong, Shuang Cang, and Hongnian Yu. A practical multi-sensor activity recognition system for home-based care. *decision support systems*, 66:61–70, 2014.

[178] Junqi Guo, Xi Zhou, Yunchuan Sun, Gong Ping, Guoxing Zhao, and Zhuorong Li. Smartphone-based patients' activity recognition by using a self-learning scheme for medical monitoring. *Journal of medical systems*, 40:1–14, 2016.

[179] Jianguo Hao, Abdenour Bouzouane, and Sébastien Gaboury. Recognizing multi-resident activities in non-intrusive sensor-based smart homes by formal concept analysis. *Neurocomputing*, 318:75–89, 2018.

[180] Michail Kaseris, Ioannis Kostavelis, and Sotiris Malassiotis. A comprehensive survey on deep learning methods in human activity recognition. *Machine Learning and Knowledge Extraction*, 6(2):842–876, 2024.

[181] Babak Moradi, Mohammad Aghapour, and Afshin Shirbandi. Compare of machine learning and deep learning approaches for human activity recognition. In *2022 30th International Conference on Electrical Engineering (ICEE)*, pages 592–596. IEEE, 2022.

[182] Ankita, Shalli Rani, Himanshi Babbar, Sonya Coleman, Aman Singh, and Hani Moaiteq Aljahdali. An efficient and lightweight deep learning model for human activity recognition using smartphones. *Sensors*, 21(11):3845, 2021.

[183] D. Alghazzawi, O. Rabie, O. Bamasaq, A. Albeshri, and M. Z. Asghar. Sensor-based human activity recognition in smart homes using depthwise separable convolutions. *Human-centric Computing and Information Sciences*, 12(50), 2022.

[184] A. R. Javed, R. Faheem, M. Asim, T. Baker, and M. O. Beg. A smartphone sensors-based personalized human activity recognition system for sustainable smart cities. *Sustainable Cities and Society*, 71:102970, 2021.

[185] D. Das, Y. Nishimura, R. P. Vivek, N. Takeda, S. T. Fish, T. Ploetz, and S. Chernova. Explainable activity recognition for smart home systems. *ACM Transactions on Interactive Intelligent Systems*, 13(2):1–39, 2023.

[186] Y. Yu, K. Tang, and Y. Liu. A fine-tuning based approach for daily activity recognition between smart homes. *Applied Sciences*, 13(9):5706, 2023.

[187] A. K. Sahoo, V. Kompally, and S. K. Udgata. Wi-fi sensing based real-time activity detection in smart home environment. In *Proceedings of the IEEE Applied Sensing Conference (APSCON)*, pages 1–3, 2023.

[188] L. G. Fahad and S. F. Tahir. Activity recognition and anomaly detection in smart homes. *Neurocomputing*, 423:362–372, 2021.

[189] S. Mekruksavanich and A. Jitpattanakul. Lstm networks using smartphone data for sensor-based human activity recognition in smart homes. *Sensors*, 21(5):1636, 2021.

[190] Y. Li, G. Yang, Z. Su, S. Li, and Y. Wang. Human activity recognition based on multi-environment sensor data. *Information Fusion*, 91:47–63, 2023.

[191] Roghayeh Mojarad, Abdelghani Chibani, Ferhat Attal, Ghazaleh Khodabandelou, and Yacine Amirat. A hybrid and context-aware framework for normal and abnormal human behavior recognition. *Soft Computing*, 28(6):4821–4845, 2024.

[192] G. Meditskos, S. Dasiopoulou, V. Efstathiou, and I. Kompatsiaris. Sp-act: A hybrid framework for complex activity recognition combining owl and sparql rules. In *2013 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pages 25–30, 2013.

[193] M. Larhrib, M. Escribano, C. Cerrada, and J. J. Escribano. An ontological behavioral modeling approach with shacl, sparql, and rdf applied to smart grids. *IEEE Access*, 2024.

[194] Q. Ni, I. P. de la Cruz, and A. B. Garcia Hernando. A foundational ontology-based model for human activity representation in smart homes. *Journal of Ambient Intelligence and Smart Environments*, 8(1):47–61, 2016.

[195] K. Moulouel, A. Chibani, and Y. Amirat. Ontology-based hybrid commonsense reasoning framework for handling context abnormalities in uncertain and partially observable environments. *Information Sciences*, 631:468–486, 2023.

[196] K. Moulouel, A. Chibani, H. Abdelkawy, and Y. Amirat. Hybrid approach for anticipating human activities in ambient intelligence environments. In *2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)*, pages 2006–2011, 2022.

[197] Nadia Agti, Lyazid Sabri, and Okba Kazar. A framework for an ontological querying-based cognitive perspective for activity recognition. *International Journal of Computers and Their Applications*, 32(2):112–123, 2025.

[198] P. Compton et al. The ssn ontology of the w3c semantic sensor network incubator group. *Web Semantics: Science, Services and Agents on the World Wide Web*, 17:25–32, 2012.

[199] Oscar D Lara and Manuel A Labrador. A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys & Tutorials*, 15(3):1192–1209, 2013.

[200] Nadia Agti, Lyazid Sabri, and Okba Kazar. Hybrid deep learning model for human activity recognition using smartphone and smart home data. *Journal of Ambient Intelligence and Smart Environments*, page 18761364251339587, 2024.

[201] Andreas Holzinger, Chris Biemann, Constantinos S Pattichis, and Douglas B Kell. What do we need to build explainable ai systems for the medical domain? *arXiv preprint arXiv:1712.09923*, 2017.

[202] Liming Chen, Chris D Nugent, and Hui Wang. A knowledge-driven approach to activity recognition in smart homes. *IEEE Transactions on Knowledge and Data Engineering*, 24 (6):961–974, 2012.

[203] Icek Ajzen. The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2):179–211, 1991.

[204] Ian Horrocks, Peter F Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, and Mike Dean. Swrl: A semantic web rule language combining owl and ruleml. Technical report, W3C Member Submission, 2004.

[205] Anupriya Ranganathan and Roy H Campbell. A middleware for context-aware agents in ubiquitous computing environments. In *Middleware 2003*, pages 143–161. Springer, 2003.

[206] Aldo Gangemi, Nicola Guarino, Claudio Masolo, Alessandro Oltramari, and Luc Schneider. Sweetening ontologies with dolce. In *EKAW*, volume 2473, pages 166–181. Springer, 2002.

[207] Dan Brickley and Libby Miller. Foaf vocabulary specification 0.91, 2005. Available: http://xmlns.com/foaf/spec/.

[208] Jorge Reyes-Ortiz, Davide Anguita, Alessandro Ghio, Luca Oneto, and Xavier Parra. Human activity recognition using smartphones dataset, 2012. https://archive.ics.uci.edu/ml/datasets/human+activity+recognition+using+smartphones.

[209] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, 12(2):74–82, 2011.

[210] Attila Reiss and Didier Stricker. Introducing a new benchmarked dataset for activity monitoring. In *2012 16th international symposium on wearable computers*, pages 108–109. IEEE, 2012.

[211] Sandro Hurtado Requena, José Manuel García Nieto, Anton Popov, and Ismael Navas Delgado. Human activity recognition from sensorised patient's data in healthcare: A streaming deep learning-based approach. *IJIMAI*, 8(1):23–37, 2023.

[212] Pardeep Seelwal, Chetana Srinivas, et al. Human activity recognition using wisdm datasets. *Journal of Online Engineering Education*, 14(1s):88–94, 2023.

[213] Yasin Kaya and Elif Kevser Topuz. Human activity recognition from multiple sensors data using deep cnns. *Multimedia Tools and Applications*, 83(4):10815–10838, 2024.

[214] Hailong Rong, Hao Wang, Xiaohui Wu, Tianlei Jin, and Ling Zou. A multi-scale time segments attention mechanism for sensor-based human activity recognition. 2025.

[215] Andrey Ignatov. Real-time human activity recognition from accelerometer data using convolutional neural networks. *Applied Soft Computing*, 62:915–922, 2018.

[216] Nidhi Dua, Shiva Nand Singh, and Vijay Bhaskar Semwal. Multi-input cnn-gru based human activity recognition using wearable sensors. *Computing*, 103(7):1461–1478, 2021.

[217] Imen Trabelsi, Jules Françoise, and Yacine Bellik. Sensor-based activity recognition using deep learning: A comparative study. In *Proceedings of the 8th International Conference on Movement and Computing*, pages 1–8, 2022.

[218] Sravan Kumar Challa, Akhilesh Kumar, and Vijay Bhaskar Semwal. A multibranch cnn-bilstm model for human activity recognition using wearable sensor data. *The Visual Computer*, 38(12):4095–4109, 2022.

[219] Jiahui Wen and Mingyang Zhong. Activity discovering and modelling with labelled and unlabelled data in smart environments. *Expert Systems with Applications*, 42(14):5800–5810, 2015.

[220] Nawel Yala, Belkacem Fergani, and Anthony Fleury. Towards improving feature extraction and classification for activity recognition on streaming data. *Journal of Ambient Intelligence and Humanized Computing*, 8(2):177–189, 2017.

[221] Sara Ashry, Tetsuji Ogawa, and Walid Gomaa. Charm-deep: Continuous human activity recognition model based on deep neural network using imu sensors of smartwatch. *IEEE Sensors Journal*, 20(15):8757–8770, 2020.

[222] Kalhan Boralessa, Isibor Kennedy Ihianle, Pedro Machado, Salisu Wada Yahaya, and Ahmad Lotfi. Input-adaptation approach for human activity recognition. In *Proceedings of the 17th International Conference on PErvasive Technologies Related to Assistive Environments*, pages 369–374, 2024.

[223] Thomas Plötz et al. Using graphs to perform effective sensor-based human activity recognition in smart homes. *Sensors*, 24(12):3944, 2024.

[224] Abbas Ramadan, Farida Saïd, and Florent Frizon De Lamotte. Enhanced online segmentation and performance evaluation method for real-time activity recognition in smart homes. In *2024 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, pages 1–8. IEEE, 2024.

[225] Harald Beck Minh Dao-Tran Thomas and Eiter Michael Fink. Lars: A logic-based framework for analyzing reasoning over streams. 2015.

[226] Harald Beck, Thomas Eiter, and Christian Folie. Ticker: A system for incremental asp-based stream reasoning. *Theory and Practice of Logic Programming*, 17(5-6):744–763, 2017.

[227] S. Zhang, Y. Li, S. Zhang, F. Shahabi, S. Xia, Y. Deng, and N. Alshurafa. Deep learning in human activity recognition with wearable sensors: A review on advances. *Sensors*, 22 (4):1476, 2022.

[228] L. Miranda, J. Viterbo, and F. Bernardini. A survey on the use of machine learning methods in context-aware middlewares for human activity recognition. *Artificial Intelligence Review*, 55(4):3369–3400, 2022.

[229] R. Kanjilal, M. F. Kucuk, and I. Uysal. Sub-transfer learning in human activity recognition: Boosting the outlier user accuracy. *IEEE Sensors Journal*, 2023.

[230] E. Kim. Interpretable and accurate convolutional neural networks for human activity recognition. *IEEE Transactions on Industrial Informatics*, 16(11):7190–7198, 2020.

[231] X. Cheng, L. Zhang, Y. Tang, Y. Liu, H. Wu, and J. He. Real-time human activity recognition using conditionally parametrized convolutions on mobile and wearable devices. *IEEE Sensors Journal*, 22(6):5889–5901, 2022.

[232] I. E. Jaramillo, C. Chola, J.-G. Jeong, J.-H. Oh, H. Jung, J.-H. Lee, W. H. Lee, and T.-S. Kim. Human activity prediction based on forecasted imu activity signals by sequence-to-sequence deep neural networks. *Sensors*, 23(14):6491, 2023.