A Hybrid Context-aware Framework to Detect Abnormal Human Daily Living Behavior

Roghayeh Mojarad*, Ferhat Attal†, Abdelghani Chibani‡, Yacine Amirat§
Univ Paris Est Creteil, LISSI, F-94400 Vitry, Vitry-sur-Seine, France
*roghayeh.mojarad@u-pec.fr, †ferhat.attal@u-pec.fr, ‡abdelghani.chibani@u-pec.fr, §amirat@u-pec.fr

***Abstract*—In Ambient Assisted Living (AAL) systems, one of the main objectives is to provide intelligent services to enhance the quality of people's lives in terms of safety, well-being, and autonomy. One of the challenges in designing these systems is abnormal human behavior detection, which is critically important to prevent users, especially elderlies, from dangerous situations. Abnormality detection has been widely explored in various fields; however, challenges remain in developing effective approaches that take into account the limitations of data-driven and knowledge-driven approaches in detecting abnormal human behaviors in AAL systems. In this paper, a hybrid context-aware framework combining a machine-learning model and probabilistic reasoning is proposed to detect abnormal human behavior. An LSTM model is firstly used to classify input data into a set of labels describing human activities. Different human activity contexts, including the duration, frequency, time of the day, locations, used objects, and sequences of the frequent activities, are then extracted to analyze human behaviors. The obtained human activities and behaviors are mapped to the proposed ontology called Human AcTivity (HAT) ontology, which conceptualizes human behavior contexts. Afterward, the abnormal human behaviors are detected using Markov Logic Network (MLN), which combines logic and probability. The concepts and relationships defined in HAT ontology are exploited in defining the FOL rules used in MLN. The proposed framework is evaluated on the *Orange4Home dataset* and *HAR dataset using smartphones*. The obtained results demonstrate the ability of the proposed framework to detect abnormal human daily living behavior with high accuracy.**

I. Introduction

In Ambient Assisted Living (AAL) systems, one of the main objectives is to improve the quality of people's lives in terms of safety, well-being, and autonomy using intelligent services [1], [2], [3]. Several challenges are raised in providing assistance services in these systems. One of them is abnormal human behavior detection, which is critically important to prevent users, especially elderlies, from dangerous situations. One of the challenges to detect abnormal human behavior in assistive systems [4], [1] is to propose a comprehensive definition of human behavior considering different contexts of human activities and behaviors. Dey [5] defined context as "any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves". The proposed human behavior definition should be standard and also machine-understandable.

In the state-of-the-art, researchers usually do not distinguish between human activity and human behavior [6]. However, in a few studies, these two terms are defined differently. Human behavior is defined as frequent activities that the user performs in different circumstances [7]. This definition is not comprehensive since it does not consider human behavior contexts such as locations, objects, duration, time of the day, and etc. In this study, a comprehensive definition of human behavior is proposed; human behavior is defined as a structure with six components: (i) frequent activities in specific locations, such as *napping in the bedroom*, (ii) frequent activities with specific objects, such as *napping with the pillow*, (iii) frequent activities in particular times of the day, such as *napping in the afternoon*, (iv) frequent activities within particular ranges of duration, such as *napping takes between d_{min} to d_{max} minutes, where d_{min} and d_{max} represent the minimum and maximum duration of napping activity, respectively*, (v) recurrent activities with particular frequencies per day, such as *frequency of napping per day is between f_{min} and f_{max}, where f_{min} and f_{max} represent the minimum and maximum frequency of napping, respectively*, and (vi) frequent sequences of activities, such as the activity sequence *reading a book-napping*. Abnormal human behavior is defined as an unexpected human behavior since it differs from typical or usual behaviors [8]. Based on the human behavior definition proposed in this study, abnormal human behavior can be categorized into six types of abnormalities: (i) recurrent unexpected activities in specific locations, (ii) recurrent unexpected activities with specific objects, (iii) recurrent unexpected activities in particular times of the day, (iv) recurrent unexpected activities within particular ranges of duration, (v) recurrent unexpected activities with particular frequencies per day, and (vi) recurrent unexpected sequences of activities.

In this paper, a hybrid context-aware framework is proposed to detect abnormal human daily living behavior. This approach aims to overcome the limitations of data-driven and knowledge-driven approaches while exploiting their advantages. It combines probabilistic reasoning and a machine-learning model to consider different human behavior contexts, handle a huge amount of data, and manage uncertain information. This framework consists of four main modules: (i) human activity recognition, (ii) human behavior analysis, (iii) mapping to an ontology, and (iv) abnormal human behavior detection. In the first module, a machine-learning model of Long Short-Term Memory (LSTM) type is used to classify input data into a set of labels describing ongoing human activity. In

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

the second module, the obtained labels are analyzed in terms of locations, used objects, times of the day, duration, frequencies, and activity sequences to provide the six complements used in the proposed human behavior definition. In the third module, the proposed Human AcTivity (HAT) ontology, inspired by the ConceptNet semantic network [9], is used to conceptualize human activities, human behaviors, and their contexts. The obtained human behaviors are mapped to the proposed HAT ontology to conceptualize shared concepts of human activities and human behaviors. Human activity predictions and also data obtained from sensors are generally uncertain; mapping uncertainty over ontology will not yield good performance in the context of activity or behavior recognition [10]. Therefore in the last module, a Markov Logic Network (MLN), combining logic and probability to handle data uncertainty by assigning weights to FOL rules, is used to detect abnormal human daily living behaviors. The use of MLN is motivated by the fact that it allows integrating probabilistic reasoning and inductive logic programming in a unified framework. The proposed framework is evaluated on the *Orange4Home* dataset [11] and *HAR dataset using smartphones* [12].

This paper is organized as follows: section II is dedicated to related works. The details of the proposed framework are presented in section III. The experimental results are provided and discussed in section IV. Finally, section V provides a summary of the proposed framework and research perspectives to enhance the proposed framework.

II. RELATED WORKS

In several studies, researchers do not make any difference between human behavior and human activity; i.e., these two terms are usually used interchangeably [13], [14], [15], [16], [6]. However, in some studies, these two terms are considered as two different abstraction levels of human activity; human behavior is considered as a superior abstraction level of human activity [17], [7], [18]. In these studies, human behavior is usually considered as frequent activities performed in different situations. In [17], low-rank matrix decomposition and time-warping techniques are integrated in order to analyze human activities. Then, the routines and deviations of activities are distributed into different clusters using Dynamic Time Warping (DTW). The well-known Silhouette index is used to find the optimal number of clusters [19]. To find the final memberships of clusters for each day, a cross-product is performed between clusters of activity routines and clusters of routine deviations. In [7], an approach is proposed to extract the human routines from human behavior logs automatically. A Markov Decision Processes (MDP) framework is applied to capture human routines, and the MaxCausalEnt algorithm is then used to predict human behavior.

One of the most challenging research topics in the field of human behavior analysis is abnormal human behavior detection. The latter has received attention from researchers working in several application domains, including ambient assisted living [18] and healthcare [8]. The objective of abnormal human behavior detection is to detect human behaviors that are unexpected since they are different from usual behaviors [8]; i.e., it “refers to the problem of finding patterns in data that do not conform to expected behavior” [20]. The most proposed approaches in this domain are vision-based [21], [22] that show several limitations such as visual occlusions and privacy loss. These limitations were the motivation of this study to focus on sensor-based abnormal human behavior detection. One of the conventional approaches to recognize human behavior and detect abnormal human behavior is machine-learning techniques, such as Hidden Markov Model (HMM) [23] and Support Vector Machine (SVM) [24], [8]. These approaches do not consider the context of human behaviors; hence, they may miss some behaviors that are frequent only under certain conditions. Also, some additional information about a behavior may be missed when the context of human behavior is not considered. However, few studies focus on the human behavior contexts. In [18], the notion of contextualized behavior is introduced where a context can be a specific day (e.g., Tuesday), a specific time (e.g., at 9:00 am) or a specific activity (e.g., sleeping) or the combination of any of them. An algorithm is proposed to find these patterns in a data stream. The algorithm consists of two main steps: segmentation of the stream and the extraction of frequent sequences. This approach cannot handle uncertain information. In order to handle uncertainty besides taking into account the context, in [25], Fine-grained Abnormal BEhavior Recognition (FABER) hybrid approach is proposed to recognize abnormalities in human activities. The authors exploit a semantic integration layer to recognize simple actions or events. Then, MLN reasoner is employed to recognize the activity boundaries, starting and ending points of activities. A knowledge-based inference engine is then used to detect abnormalities based on the recognized activity boundaries. The abnormalities are defined based on only the starting and ending points of activities. In other words, this study does not enough consider human activity and behavior contexts.

In this paper, a hybrid context-aware framework is proposed to detect abnormal human behavior. The main human behavior contexts considered in the proposed framework are the locations, objects, times of the day, duration, frequencies, and sequences of frequent activities. To the best of our knowledge, the proposed framework is the first hybrid one that considers different human behavior contexts while handles the uncertainty of human daily living behaviors to detect abnormal human behaviors.

III. THE PROPOSED FRAMEWORK

The architecture of the proposed framework is shown in Fig. 1. It consists of four main modules: (i) human activity recognition, (ii) human behavior analysis, (iii) mapping to an ontology, and (iv) abnormal human behavior detection. In the first module, an LSTM model is used for the activity, location, and object recognition. This model is able to solve sequential information modeling in the short term and also the long term, which is essential in human behavior analysis. In the second module, the recognized activities, locations, and

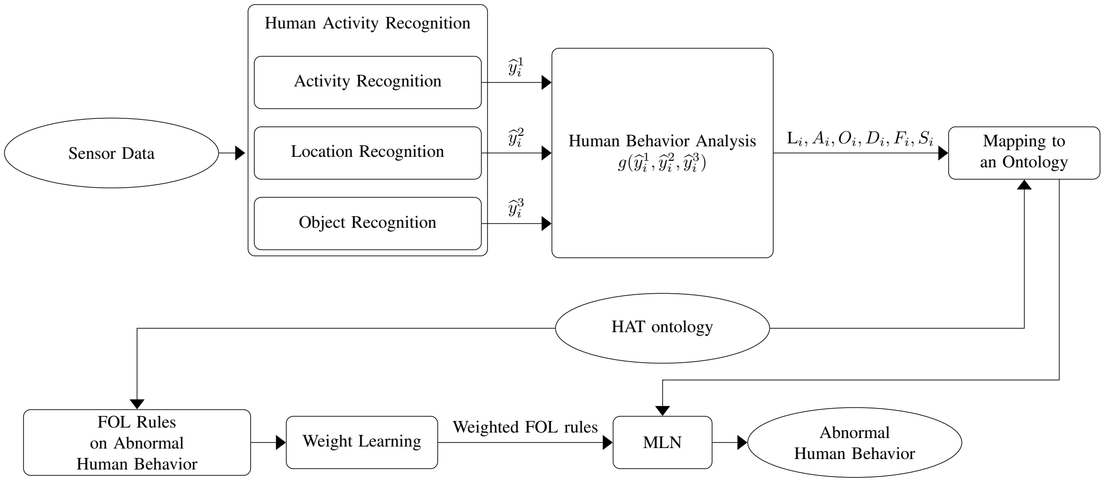

Fig. 1. Architecture of the proposed framework.

objects are analyzed to provide the different contexts of human behaviors. In the third module, the obtained human activities and behaviors are mapped to the proposed HAT ontology to provide shared conceptualized information. In the fourth module, an abnormal human behavior is detected using MLN, which enables probabilistic inferences. This module uses FOL rules about abnormal human behaviors defined by an expert. The HAT ontology is exploited in defining the FOL rules such as providing shared concepts of human behaviors and their contexts as predicates used in FOL rules and discovering automatically inconsistency among knowledge. Then, a weight corresponding to the truth degree of rule is added to each defined FOL rule. Weights are efficiently learned from data by optimizing iteratively a pseudo-likelihood measure.

A. Human activity recognition

In the human activity recognition module, input data are classified into a set of labels describing the ongoing activity. In this study, labels are activity, location, and object used in the activity. Therefore, the input data are represented as pairs composed of the data X_i and labels Y_i:

$$D = \{(X_i, Y_i) \mid 1 \leq i \leq N\} \tag{1}$$

where X_i represents the i^{th} sample data. Y_i represents the vector of labels assigned to the i^{th} sample data. N represents the maximum number of sample data. Each X_i includes d attributes while each Y_i includes three labels:

$$X_i = \{x_i^1, x_i^2, ..., x_i^d\} \tag{2}$$

$$Y_i = \{y_i^1, y_i^2, y_i^3\} \tag{3}$$

where y_i^1, y_i^2, and y_i^3 represent respectively the activity, location, and object labels assigned to the i^{th} sample data. Each label has a specific number of classes; for example, the number of classes for the activity label is 17 while it is 8 for the location label. These labels can be formalized as follows:

$$\begin{aligned} y_i^1 &\in \{c_1^1, c_2^1, ..., c_q^1\} \\ y_i^2 &\in \{c_1^2, c_2^2, ..., c_w^2\} \\ y_i^3 &\in \{c_1^3, c_2^3, ..., c_z^3\} \end{aligned} \tag{4}$$

where q, w, and z represent the number of classes for activity, location, and object labels. To classify input data into these three labels, the human activity recognition module uses three models: activity recognition, location recognition, and object recognition. These models are trained independently to allow the proposed framework to be used even without the information of location or object. These models can be formalized as functions f_A, f_L, and f_O such as:

$$\begin{aligned} \widehat{y}_i^1 &= f_A(X_i) \\ \widehat{y}_i^2 &= f_L(X_i) \\ \widehat{y}_i^3 &= f_O(X_i) \end{aligned} \tag{5}$$

where $\widehat{y}_i^1$, $\widehat{y}_i^2$, and $\widehat{y}_i^3$ represent the predicted labels for human activity, location, and object, respectively. f_A, f_L, and f_O are prediction functions of activity, location, and object models, respectively. In this study, an LSTM model, a type of Recurrent Neural Networks (RNN) that includes special units beside standard units, is used. Each LSTM unit includes a memory cell that can keep information for a long period. Three gates called *forget gate*, *input gate*, and *output gate* are used to control information in this memory cell. This model is appropriate to model human daily living activities that are characterized by time-series data. The used LSTM model consists of 4 layers: (i) LSTM layer with 100 neurons, (ii) LSTM layer with 50 neurons, (iii) Droupout layer with fraction rate 0.5, and (iv) Dense layer with the number of neurons equals with the number of classes. The dropout layer is used to avoid the overfitting problem by randomly and temporarily deleting neurons in the hidden layer of the network at each update of the training phase. In this model, the used optimization function is *Adam* while the lost function model is *categorical-crossentropy*.

B. Human behavior analysis

In the human behavior analysis module, six components used in the proposed human behavior definition are extracted using an algorithm formalized as a function g:

$$L_i, A_i, O_i, D_i, F_i, S_i = g(\widehat{y}_i^1, \widehat{y}_i^2, \widehat{y}_i^3) \tag{6}$$

where L_i represents the list of frequent activities in specific locations; A_i represents the list of frequent activities in specific times of the day. O_i represents the list of frequent activities performed using specific objects. D_i represents the list of frequent activities within specific duration. F_i represents the list of recurrent activities with specific frequencies. S_i represents the list of frequent sequence of activities. In this algorithm, eight lists of hash maps are generated for each activity; these lists are associated with locations, objects, times of the day, minimum duration, maximum duration, minimum frequency, maximum frequency, and previous activity.

C. Mapping to an ontology

The HAT ontology provides a formal specification of a shared conceptualization to describe human activities, human behaviors, and their contexts. This ontology is inspired by the *ConceptNet* semantic network, which is a knowledge graph that makes links between words and phrases in natural language using labeled edges [9]; e.g., the word *earth* is linked with the phrase *grow people* using the *is used for* labeled edge. Figure 2 shows an overview of the HAT ontology modeled using the Semantic Web Ontology Language (OWL) [26]. It consists of two upper-level concepts: *Event* and *Object*. Six other concepts, namely: *Activity*, *Location*, *Time*, *Physical Object*, *Duration*, and *Frequency*, are derived from the mentioned two upper-level concepts. Six different relationships, namely: *has place*, *has frequency*, *has duration*, *has time*, *is used for*, and *is a*, are defined to connect the concepts defined in HAT ontology. Table I represents the formalized relationships between the main concepts in the HAT ontology. The concepts and relationships among these concepts defined in the HAT ontology are exploited in defining FOL rules and predicates used in MLN.

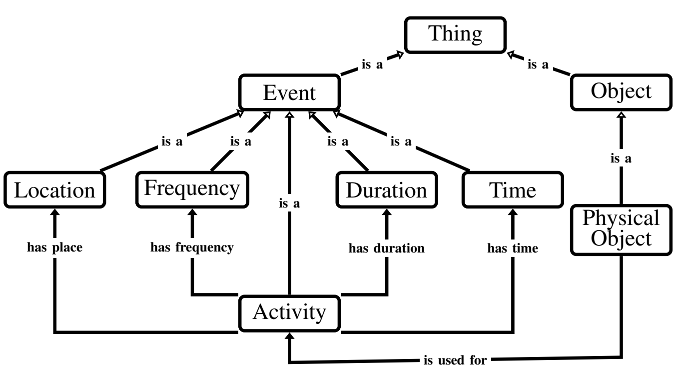

Fig. 2. Overview of the HAT ontology.

TABLE I
FORMALIZED RELATIONSHIPS BETWEEN THE MAIN CONCEPTS IN THE HAT ONTOLOGY.

$\mathsf{Object}(x)$	$\vee\, \mathsf{Event}(x)$	$\rightarrow\ \mathsf{Thing}(x)$.
$\mathsf{Location}(x)$	$\vee\, \mathsf{Activity}(x)$	$\rightarrow\ \mathsf{Event}(x)$.
$\mathsf{Activity}(a)$	$\rightarrow\ \exists o\ \mathsf{Object}(o)\ \wedge\ \mathsf{isusedfor}(o,a)$.	
$\mathsf{Activity}(a)$	$\rightarrow\ \exists l\ \mathsf{Location}(l)\ \wedge\ \mathsf{hasplace}(a,l)$.	
$\mathsf{Activity}(a)$	$\rightarrow\ \exists t\ \mathsf{time}(t)\ \wedge\ \mathsf{hastime}(a,t)$.	
$\mathsf{Activity}(a)$	$\rightarrow\ \exists d\ \mathsf{duration}(d)\ \wedge\ \mathsf{hasduration}(a,d)$.	
$\mathsf{Activity}(a)$	$\rightarrow\ \exists f\ \mathsf{frequency}(f)\ \wedge\ \mathsf{hasfrequency}(a,f)$.	

D. Abnormal human behavior detection

In the proposed framework, MLN is used to detect abnormal human behavior detection module. MLN is constructed by learning weighted FOL rules that enable probabilistic reasoning within a knowledge representation framework. It is a probabilistic logic that applies the foundations of a Markov network to FOL to allow probabilistic inferences. From the perspective of probability, MLN provides a compact language to define extensive Markov networks with the ability to integrate a wide range of knowledge into Markov networks. From the perspective of FOL, MLN allows handling uncertain, imperfect, and even contradictory knowledge.

MLN is defined as weighted FOL rules, where weights are real values representing the truth degree of the rules. In MLN, based on the weight assignment [27], the FOL rules are classified into two categories: soft rules and hard rules. MLN is formalized as follows:

Definition 1: MLN L is a set of pairs (F_j, w_j), where F_j is a FOL rule, and w_j is a real number. A Markov network M_{LC} where C represents a finite set of constants $C = \{c_1, c_2, ..., c_{|C|}\}$ is defined as follows [27]:

1) Each possible grounding of each predicate appearing in L has one binary node. If the ground atom is true, the value of the node is 1; otherwise, it is 0.
2) Each possible grounding of rule F_j in L has one feature. If the ground rule is true, the value of this feature is 1; otherwise, it is 0. The weight of the feature is w_j, associated with F_j in L.

where the groundings of a rule are formed simply by replacing its variables with constants in all possible ways [28]. There is an edge between two nodes of M_{LC} if and only if the ground predicates corresponding to them appear together at least in one grounding of rules in L. MLN can be considered as a *template* for constructing Markov networks. It generates different Markov networks with different sizes using given different sets of constants. Each Markov network composed of an undirected graph G and a set of potential functions ϕ_j. Each variable is represented as a node in the graph. Each clique in the graph has a potential function, which is a non-negative real-valued function of the state of the corresponding clique [27]. All generated Markov networks have specific regularities in structure and parameters given by the MLN. A set of ground atoms makes a possible world in MLN. All MLN possible worlds are true with a probability according to the number of rules they satisfy and the weights of these rules. An MLN can represent hard constraints by assigning an infinite weight to some rules; if a possible world violates these rules, the probability of that possible world will be zero. The probability distribution over a set of possible worlds $Z = \{Z_1, Z_2, ..., Z_r\}$ is calculated as follows [27]:

$$P(Z = z) = \frac{1}{M} \prod_j \phi_j(z_{\{j\}})^{n_j(z)} \tag{7}$$

where M is normalization factor. $n_j(z)$ represents the number of true grounding of F_j in the possible world z. The state, also called truth values, of the ground atoms appearing in F_j is represented using $z_{\{j\}}$. It is worth mentioning that MLN is represented as the products of potential functions in Eq. 7 however it could be represented as log-linear models, which each potential function replaced by an exponentiated weighted sum of features of the state, i.e., $\phi_j(z_{\{j\}}) = e^{w_j}$, as follows:

$$P(Z = z) = \frac{1}{M} exp(\Sigma_j w_j n_j(z)) \tag{8}$$

In domains with a combination of hard and soft constraints such as human activity recognition and abnormal human behavior detection, MLN can be used as the most appropriate approach; some rules with certainty are considered as hard constraints, and others are considered as soft constraints. The learning task of the MLN can be categorized into two subtasks: (i) structure learning and (ii) weight learning. The structure can be provided by rules written by human experts while weight learning is an optimization problem that requires learning data. In this paper, an expert defines the FOL rules based on HAT ontology, and weights are learned from data by optimizing iteratively a pseudo-likelihood measure. The total number of defined rules to detect abnormal human behaviors is 433. Table II reports the main predicates used in these FOL rules. If one of the predicates represented abnormal human behaviors is inferred from reasoning on weighted FOL rules used in MLN, the framework detects an abnormal behavior. Since the different types of abnormal human behaviors are represented using different predicates, the type of abnormal human behavior is detected when an abnormality is identified.

TABLE II
LIST OF THE MAIN PREDICATES USED IN THE DEFINED FOL RULES.

$Act(activity, time)$
$Loc(location, time)$
$Obj(object, time)$
$AbnormalActLoc(activity, location, time)$
$AbnormalActObj(activity, object, time)$
$AbnormalActTime(activity, timeoftheday, time)$
$AbnormalActDur(activity, duration, time)$
$AbnormalActFreq(activity, frequency, time)$
$AbnormalSeqAct(activity1, activity2)$

Equation 9 shows an example of a standard FOL rule in MLN, which is a weighted FOL rule:

$$\begin{aligned} w \quad & Act(Eating, time) \wedge Loc(Bedroom, time) \\ \Rightarrow & AbnormalActLoc(Eating, Bedroom, time) \end{aligned} \tag{9}$$

where w represents the weight of the rule, which is obtained by weight learning from data. The mentioned rule consists of three predicates: $Act(Eating, time)$, $Loc(Bedroom, time)$, and $AbnormalActLoc(Eating, Bedroom, time)$. It is also composed of two constants: $Eating$ and $Bedroom$ and one variable, which is $time$. Merging this type of rules makes a Knowledge Base (KB) on abnormal human behaviors used in the proposed framework; i.e, a list of rules modeling the conditions to discriminate an abnormal behavior are provided in this framework. The number of conditions may differ for each type of abnormal human behavior.

The hybrid nature of MLN enables it to overcome the limitation of data-driven approaches as well as rule-based ones. In contrast to data-driven approaches, MLN can consider the activity sequence and temporal relationships among activities. Additionally, the contexts of human activity and behaviors are not considered in the data-driven approaches while MLN considers them. In terms of performance, MLN can perform better in comparison with rule-based approaches as the unused rules are removed in the weight learning process. Therefore MLN can make inferences more efficiently than rule-based approaches. The time consuming of reasoning depends on the number of rules, in general, if the number of rules is reduced, the time consuming will also be decreased. Moreover, as opposed to pure rule-based approaches, it can deal with some of the unreliability produced in the classification since MLN can manage that during the weight-learning process. In the proposed framework, when the LSTM model can hardly predict one label describing ongoing activity, it might create false abnormality conditions. However, in the training phase of MLN, its effects are mitigated using a lower weight. It is worth mentioning that, in most cases, the misclassifications are rare, and the weight of rules depends on the rate of misclassification. Hence MLN can fully cover the classification errors only if they are common.

Although MLN provides high performance in specific contexts, it has some limitations such as limitations of knowledge representation using FOL. MLN also has limitations to model high-level activities. In addition, the inability to automatically discover inconsistencies among represented knowledge, lack of domain knowledge, and hierarchical association of domain-related concepts are other limitations of MLN [10]. Due to these limitations and the ability of ontology to deal with them, in the proposed framework, HAT ontology is used before MLN to allows offering consistent knowledge, temporal modeling, contextual modeling, high-level activity modeling, and enable probabilistic inferences in a unified framework.

IV. EXPERIMENTS AND RESULTS

In this section, the performance of the proposed framework are evaluated in terms of precision, recall, F-measure and accuracy on the *Orange4Home* dataset [11] and the *HAR dataset using smartphones* [12], benchmarks for human activity recognition. To implement and evaluate the framework, a computer equipped with an Intel i7-8650U 2.11GHz CPU with 32GB RAM is used.

A. Description of the datasets

Orange4Home dataset [11] includes data collected from 236 sensors that capture information about the use of electrical equipment, water consumption, operation of doors, etc. The sensors are placed in the different locations of an instrumented home. One occupant was involved in this dataset to do seventeen daily living activities during four consecutive weeks of working days. The total data recording is around 180 hours. The *Orange4Home* dataset consists of four main contexts of human activities, namely: identity, time-of-day, place, and activity. Identity considers the literal sense of an occupant and also his social role in the home. Time-of-day takes into account temporal information such as date and time. Place considers a geographical location in the home such as *kitchen*. Activity takes into account a set of actions that the occupant performs. *HAR dataset using smartphones* [12] includes data collected from a waist-mounted smartphone with embedded inertial sensors such as accelerometers and gyroscopes. The data capture rate is fifty Hertz. Thirty participants were involved in this dataset to do six activities, namely: (i) Walking, (ii) Walking-upstairs, (iii) Walking-downstairs, (iv) Sitting, (v) Standing, and (vi) Laying.

B. Performance of human activity recognition module

The LSTM model used in the human activity recognition module is evaluated in terms of precision, recall, F-measure, and accuracy. To evaluate the model, the batch size is set to 50 instances, and the epoch number is set to 300 iterations. The internal architecture of the LSTM model and the time step of the sequences are heuristically set.

Since the *Orange4Home* dataset includes the activity and location labels, two LSTM models are independently used to classify input data into activity and location labels. The performance obtained using the LSTM model on the *Orange4Home* dataset are shown in Table III. The latter shows that the performance results obtained in the case of activity recognition and location recognition are more than 95% in terms of precision, recall, F-measure, and accuracy. One can observe that the performance results in the case of location recognition are better than those in the case of activity recognition. This can be explained by the fact that the number of samples in the case of location is higher than the number of samples in the case of activity; therefore, the LSTM model can be trained better in the case of location. Moreover, the classes in the case of location are more distinguishable compared with those in the case of activities. LSTM model is compared with two baseline models [29], namely: MultiLayer Perceptron (MLP) and the SVM model, on the *Orange4Home* dataset, see Table IV. The results show that the proposed LSTM model obtains better results in terms of *F-measure*. This is explained by the fact that the LSTM model is an appropriate model for time-series data while MLP and SVM models do not consider the activity sequences.

The *HAR dataset using smartphones* [12] includes only the activity label; therefore, one LSTM model is used for human activity recognition. The performance obtained using the LSTM model are shown in Table V. The results demonstrate that the model achieves more than 94% in terms of precision, recall, F-measure, and accuracy, which demonstrates its effectiveness. Moreover, the LSTM model is compared with two baseline models [30], namely: K-Nearest Neighbors (KNN) and SVM, on the *HAR dataset*, see Table VI. The results show that the LSTM model performs better compared with the KNN and SVM models and obtains the highest average F-measure with 94.05%.

TABLE III
PERFORMANCE ACHIEVED USING THE LSTM MODEL ON THE *Orange4Home* DATASET.

	Precision	Recall	F-Measure	Accuracy
Activity recognition	96.00	95.71	95.85	95.71
Location recognition	97.98	97.89	97.93	97.90

TABLE IV
PERFORMANCE COMPARISON OF THE LSTM MODEL WITH BASELINES IN THE CASE OF ACTIVITY RECOGNITION ON THE *Orange4Home* DATASET.

	Baselines [29]		Proposed model
Evaluation Metric	MLP	SVM	LSTM
F-measure	77.85	89.60	95.85

C. Performance of abnormal human behavior detection module

The MLN used in the abnormal human behavior detection module is implemented using Tuffy [31], an open-source MLN inference engine. The abnormal human behavior detection module is evaluated in terms of precision, recall, F-measure, and accuracy.

The *Orange4Home* dataset does not include any information about objects. Therefore, abnormal behaviors related to objects are not evaluated in the case of this dataset. The performance evaluation on this dataset is based on five types of abnormal human behaviors: (i) recurrent unexpected activities in specific locations, $AbnormalActLoc$; (ii) recurrent unexpected activities in particular times of the day, $AbnormalActTime$; (iii) recurrent unexpected activities within particular ranges of duration, $AbnormalActDur$; (iv) recurrent unexpected activities with particular frequencies per

TABLE V
PERFORMANCE ACHIEVED USING LSTM MODEL ON THE *HAR dataset*.

	Precision	Recall	F-Measure	Accuracy
Activity recognition	94.08	94.03	94.05	97.98

TABLE VI
PERFORMANCE COMPARISON OF THE LSTM MODEL WITH BASELINES ON THE *HAR dataset*.

	Baselines [30]		Proposed model
Evaluation Metric	K-NN	SVM	LSTM
F-measure	90.16	93.79	94.05

TABLE VII
ABNORMAL HUMAN BEHAVIOR DETECTION PERFORMANCE OBTAINED USING SVM AND THE PROPOSED FRAMEWORK ON THE *Orange4Home* DATASET.

	SVM				Proposed Framework			
Abnormality type	Precision	Recall	F-measure	Accuracy	Precision	Recall	F-measure	Accuracy
$AbnormalActLoc$	95.22	95.11	95.14	95.11	89.40	94.63	91.94	95.03
$AbnormalActTime$	98.66	98.66	98.66	98.66	94.98	99.53	97.20	98.28
$AbnormalActDur$	81.47	79.59	74.53	79.59	72.00	97.29	82.75	94.86
$AbnormalActFreq$	66.45	81.52	73.22	81.52	81.25	76.47	78.78	93.06
$AbnormalSeqAct$	76.61	72.85	72.71	72.85	87.76	70.37	78.11	86.12
Average	83.68	85.54	82.85	85.54	85.08	87.66	85.76	93.47

TABLE VIII
ABNORMAL HUMAN BEHAVIOR DETECTION PERFORMANCE OBTAINED USING SVM AND THE PROPOSED FRAMEWORK ON THE *HAR dataset*.

	SVM				Proposed Framework			
Abnormality type	Precision	Recall	F-measure	Accuracy	Precision	Recall	F-measure	Accuracy
$AbnormalActDur$	80.98	84.89	84.17	84.17	96.15	96.15	96.15	99.15
$AbnormalSeqAct$	62.04	77.31	70.24	70.24	90.44	94.53	92.44	94.69
Average	71.51	81.10	77.20	77.20	93.29	95.34	94.29	96.92

day, $AbnormalActFreq$; and (v) recurrent unexpected sequences of activities, $AbnormalSeqAct$, see Table VII. The *HAR dataset* includes only information about activities; therefore, only two types of abnormal human behaviors, namely: (i) recurrent unexpected activities within a particular ranges of duration, $AbnormalActDu$r; and (ii) recurrent unexpected sequences of activities, $AbnormalSeqAct$, are evaluated on this dataset, see Table VIII.

Since the used datasets do not consist of abnormal activities, an algorithm is implemented to inject different abnormalities to simulate the presence of these abnormal human behavior types. This algorithm randomly selects 30% of sample data and injects abnormal human behavior randomly. For instance, to inject $AbnormalActLoc$ abnormal human behavior type, the algorithm randomly selects a value between 0 and 1 for each sample data; if the value is less than 0.3, this data sample will be selected to change. Afterward, the algorithm randomly chooses another value between 0 and 1; if the value is less than 0.5, the location of that sample data will be changed to an unexpected location according to the activity. Otherwise, the activity of that sample data will be changed to an unexpected activity according to the location. Unexpected location and activity are obtained using the list of the normal locations for each activity.

The rules used for detecting abnormal behavior are intrinsically deterministic. However, in the proposed framework, abnormal behavior detection is coupled with a machine-learning model, LSTM model, that produces probabilistic results; hence, it might cause non-determinacy in abnormal behavior detection. In other words, if the LSTM model falsely predicts a label for a sample data; the abnormal behavior detection may produce a false detection; the LSTM model can hardly predict *leaving* activity, which creates false abnormality conditions and MLN can eliminate false positives created by the classification errors only if they are common. This can be observed through the results of Table VII. The abnormal human behavior detection module performs the best in the case of $AbnormalActTime$ abnormality type, because, it only relies on one output of the LSTM model. Moreover, as this model provides better prediction in the case of location recognition than the activity recognition; consequently, abnormal behavior detection performs better in the case of $AbnormalActLoc$ in comparison with $AbnormalActDur$, $AbnormalActFreq$, and $AbnormalSeqAct$. This can be explained by the fact that detecting $AbnormalActLoc$ abnormality type relies on two outputs from LSTM model, one from the activity recognition and another one from the location recognition while detecting the other three abnormality types rely on multiple activity recognition. It is similar for the *HAR dataset*, see Table VIII. For the *HAR dataset*, the recall rates of abnormal behaviors are similar for both $AbnormalActDur$ and $AbnormalSeqAct$ types of abnormality, however, the precision rate is better in the case of $AbnormalActDur$ abnormality type in comparison with $AbnormalSeqAct$ abnormality type. In the former, the LSTM model only affects the starting and ending points of an activity period; however, in the latter, the LSTM model is involved in the prediction of all points of the activity period. Hence it might produce more prediction errors, which will result in more false positives in abnormal behavior detection.

The proposed framework is compared with an baseline data-driven approach based on the SVM model, the most common model used for abnormality detection [32], [33]. Table VII shows the comparison results on the *Orange4Home* dataset while Table VIII presents those results on the *HAR dataset*. One can observe similar evaluation results in the cases of abnormalities related to the location and time of the day. However, the proposed framework has better performance compared to the SVM-based approach in the case of abnormality related to duration, frequency, and activity sequence, called respectively $AbnormalActDur$, $AbnormalActFreq$, and $AbnormalSeqAct$. These results can be explained by the fact that the SVM model fails to consider the sequence of activities, and also fails to consider the contexts of human behaviors, such as the frequency and duration of human activities whereas using the proposed framework allows taking into account human activity sequences and also human behavior contexts.

V. Conclusion

In this paper, a hybrid context-aware framework is proposed to detect abnormal human daily living behavior. An LSTM model is used to classify input data into appropriate human activities by predicting a set of labels describing the ongoing activity. The predicted labels are then analyzed to capture six components used in the proposed definition of human behavior. The obtained human activities and behaviors are conceptualized using the HAT ontology, which is proposed to provide a formal specification of a shared conceptualization to describe abnormal human activities. Afterward, MLN is used to detect abnormal human behaviors. The proposed framework has been evaluated on two datasets and compared with an baseline data-driven approach based on the SVM model. The obtained results demonstrate the ability of the proposed framework to detect abnormal human behaviors with high performance. These results also illustrate the superiority of the proposed framework to the baseline approach. In terms of research perspectives to this study, an interesting topic is increasing the capability of this framework using probabilistic Answer Set Programming (ASP). Another research direction to explore is to provide a recommendation system to enhance the quality of people's lives.

References

[1] R. Mojarad, F. Attal, A. Chibani, S. R. Fiorini, and Y. Amirat, "Hybrid approach for human activity recognition by ubiquitous robots," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct 2018, pp. 5660–5665.

[2] H. Kordestani, R. Mojarad, A. Chibani, A. Osmani, Y. Amirat, K. Barkaoui, and W. Zahran, "Hapicare: A healthcare monitoring system with self-adaptive coaching using probabilistic reasoning," in *2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)*. IEEE, 2019, pp. 1–8.

[3] R. Mojarad, F. Attal, A. Chibani, and Y. Amirat, "Automatic classification error detection and correction for robust human activity recognition," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2208–2215, 2020.

[4] A. A. Salah, J. Ruiz-del Solar, Ç. Meriçli, and P.-Y. Oudeyer, "Human behavior understanding for robotics," in *Human Behavior Understanding*, A. A. Salah, J. Ruiz-del Solar, Ç. Meriçli, and P.-Y. Oudeyer, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 1–16.

[5] A. K. Dey, "Understanding and using context," *Personal and Ubiquitous Computing*, vol. 5, no. 1, pp. 4–7, Feb 2001.

[6] N. D. Rodríguez, M. P. Cuéllar, J. Lilius, and M. D. Calvo-Flores, "A survey on ontologies for human behavior recognition," *ACM Comput. Surv.*, vol. 46, no. 4, pp. 43:1–43:33, Mar. 2014.

[7] N. Banovic, T. Buzali, F. Chevalier, J. Mankoff, and A. K. Dey, "Modeling and understanding human routine behavior," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: ACM, 2016, pp. 248–260.

[8] A. Lentzas and D. Vrakas, "Non-intrusive human activity recognition and abnormal behavior detection on elderly people: a review," *Artificial Intelligence Review*, pp. 1–47, 2019.

[9] R. Speer, J. Chin, and C. Havasi, "Conceptnet 5.5: An open multilingual graph of general knowledge," in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

[10] K. Gayathri, K. Easwarakumar, and S. Elias, "Probabilistic ontology based activity recognition in smart homes using markov logic network," *Knowledge-Based Systems*, vol. 121, pp. 173 – 184, 2017.

[11] J. Cumin, G. Lefebvre, F. Ramparany, and J. L. Crowley, "A dataset of routine daily activities in an instrumented home," in *Ubiquitous Computing and Ambient Intelligence*, S. F. Ochoa, P. Singh, and J. Bravo, Eds. Cham: Springer International Publishing, 2017, pp. 413–425.

[12] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A Public Domain Dataset for Human Activity Recognition Using Smartphones," *Computational Intelligence*, p. 6, 2013.

[13] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, "Analysis of human behavior recognition algorithms based on acceleration data," in *IEEE International Conference on Robotics and Automation*, May 2013, pp. 1602–1607.

[14] K. Makantasis, A. Doulamis, N. Doulamis, and K. Psychas, "Deep learning based human behavior recognition in industrial workflows," in *IEEE International Conference on Image Processing (ICIP)*, Sep. 2016, pp. 1609–1613.

[15] R. H. Baxter, N. M. Robertson, and D. M. Lane, "Human behaviour recognition in data-scarce domains," *Pattern Recognition*, vol. 48, no. 8, pp. 2377 – 2393, 2015.

[16] G. Batchuluun, J. H. Kim, H. G. Hong, J. K. Kang, and K. R. Park, "Fuzzy system based human behavior recognition by combining behavior prediction and recognition," *Expert Systems with Applications*, vol. 81, pp. 108–133, Sep. 2017.

[17] O. Yürüten, J. Zhang, and P. Pu, "Decomposing activities of daily living to discover routine clusters," in *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, ser. AAAI'14. AAAI Press, 2014, pp. 1348–1354.

[18] P. Lago, C. Jiménez-Guarín, and C. Roncancio, "Contextualized Behavior Patterns for Ambient Assisted Living," in *Human Behavior Understanding*, A. A. Salah, B. J. Kröse, and D. J. Cook, Eds. Cham: Springer International Publishing, 2015, vol. 9277, pp. 132–145.

[19] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis.* John Wiley, 1990.

[20] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1–15:58, Jul. 2009.

[21] J. Sun, J. Shao, and C. He, "Abnormal event detection for video surveillance using deep one-class learning," *Multimedia Tools and Applications*, vol. 78, no. 3, pp. 3633–3647, 2019.

[22] T. Wang, M. Qiao, Y. Deng, Y. Zhou, H. Wang, Q. Lyu, and H. Snoussi, "Abnormal event detection based on analysis of movement information of video sequence," *Optik*, vol. 152, pp. 50–60, 2018.

[23] C. Zhu and W. Sheng, "Motion- and location-based online human daily activity recognition," *Pervasive and Mobile Computing*, vol. 7, no. 2, pp. 256 – 269, 2011.

[24] J. H. Shin, B. Lee, and K. S. Park, "Detection of abnormal living patterns for elderly living alone using support vector data description," *Trans. Info. Tech. Biomed.*, vol. 15, no. 3, pp. 438–448, May 2011.

[25] D. Riboni, C. Bettini, G. Civitarese, Z. H. Janjua, and R. Helaoui, "Fine-grained recognition of abnormal behaviors for early detection of mild cognitive impairment," in *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2015, pp. 149–154.

[26] D. L. McGuinness, F. Van Harmelen *et al.*, "Owl web ontology language overview," *W3C recommendation*, vol. 10, no. 10, p. 2004, 2004.

[27] M. Richardson and P. Domingos, "Markov logic networks," *Machine learning*, vol. 62, no. 1-2, pp. 107–136, 2006.

[28] P. Singla and P. Domingos, "Discriminative training of markov logic networks," in *AAAI*, vol. 5, 2005, pp. 868–873.

[29] J. Cumin, F. Ramparany, J. L. Crowley *et al.*, "Inferring availability for communication in smart homes using context," in *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 2018, pp. 1–6.

[30] B. Romera-Paredes, M. S. Aung, and N. Bianchi-Berthouze, "A one-vs-one classifier ensemble with majority voting for activity recognition." in *ESANN*, 2013.

[31] F. Niu, C. Ré, A. Doan, and J. Shavlik, "Tuffy: Scaling up statistical inference in markov logic networks using an rdbms," *Proceedings of the VLDB Endowment*, vol. 4, no. 6, pp. 373–384, 2011.

[32] G. Diraco, A. Leone, and P. Siciliano, "Ai-based early change detection in smart living environments," *Sensors*, vol. 19, no. 16, p. 3549, 2019.

[33] A. Palaniappan, R. Bhargavi, and V. Vaidehi, "Abnormal human activity recognition using svm based approach," in *International Conference on Recent Trends in Information Technology*, April 2012, pp. 97–102.