Journal of Information Systems Engineering and Management
2025, 10(2)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Review Article

A Survey on use of Ontology for Complex Human Activity Recognition in Video Sequences

Mrs. Rashmi S R[1*], Mrs. Sameia Suha[2], Krishnan Rangarajan[3]
[1*]*Assistant Professor, Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, India*
[1*]*Email: Rashmi-cs@dayanandasagar.edu*
[2]*Assistant Professor, Department of Computer Science and Engineering, Dayananda Sagar college of Engineering, Bangalore, India*
[2]*Email: Sameia-cs@dayanandasagar.edu*
[3]*Professor, Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India*
[3]*Email: krishnan-cs@dayanandasagar.edu*

ARTICLE INFO

Received: 18 Oct 2024

Revised: 22 Dec 2024

Accepted: 30 Dec 2024

ABSTRACT

Video Surveillance systems are ubiquitous in today's world. Traditional Video Surveillance Systems are being replaced by intelligent systems and this process has been happening rapidly since decades. This work lists the Activity Recognition approaches to recognize complex human activities. Knowledge based approaches with video sensors are brought out to be the vital approach in comparison with the data-based approaches, because data driven models have to train on huge data. We investigate the significance of both the approaches in handling complex activities with a long sequence of primitive activities. Knowledge driven approaches with Ontology models the domain expertise to support the long sequenced activity recognition. Hence, the approaches supported by the Ontology for handling complex activities are found to be the most suited model for the recognition of complex human activities

Keywords: Human Activity Recognition, Ontology, Sensors, Video Sensors, Data Driven Approach, Knowledge Driven Approach.

INTRODUCTION

Recognizing complex human activities with long sequences, from a surveillance video, is the need of the hour. Video surveillance systems have been deployed in many public places worldwide since decades. Public areas like Malls, Retail stores, hospitals, Sports stadiums, Schools premises, Highways and traffic junctions, and the list goes on (Kim et al., 2021). Law imposes the usage of surveillance in all these public places to monitor the public safety. They are also in use in many of the private premises like smart homes and elderly care set ups to provide care and assistance to the needy.

With Traditional video surveillance systems, human operators have to continuously monitor and gauge the situations. This is prone to human errors and hence not totally dependable and preferable anymore. With the upheaval of study in the zone of video Processing (starting from traditional image processing techniques to current deep learning techniques combined with Bayesian networks/techniques and/or semantic knowledge) has helped the video surveillance to reach the current state of instinctive human activity recognition in videos (Hussain et al., 2022). Automatic visual surveillance systems use computer software programs to detect objects and events from the audio and video inputs from the camera, just as human eyes do. Automatic Intelligent Video surveillance systems use artificial intelligence in computer software programs to analyze and recognize objects and events from the audio and video inputs of the camera (Mahajan et al.,2020)(Nurnoby and Helmy, 2023). There are continuous attempts to advance the research to match human's scene inferring capabilities (Gull et al., 2022). These systems are expected to monitor the scene to recognize events happening and raise the alerts when required. There is a lot of research going on in this direction for decades. It is easy to find a lot of good review papers on the topic ranging from basic actions to complex activity. These studies helped us to widen our perspective and understand the topic deeply. Through these, we were able to learn facts and narrow down to our focus area and answer the research questions we framed. To mention a few here, the reviews in (Beddiar et al., 2020)(Zhang et al., 2017)(Gupta et al., 2022), highlight the importance of human action recognition with visual sensors and its approaches, spanning from good old vision processing techniques to current deep learning techniques. The reviews in (Gull et al., 2022)(Dhiman and Vishwakarma, 2019) concentrate and summarize human activity recognition approaches. The review in (Manzoor et

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

al., 2021) emphasizes the recent trends in Ontology based intelligent systems helping Robotics to reach the next level. We are going to list a few more important reviews in our work at the points of discussion where it is suitable. Out of this study, we found minimal amount of study regarding the comparison between data driven approaches and Knowledge driven approaches for video activity recognition. This work is an attempt to take the study further in this direction. Also, the existing surveys mostly address Human activity recognition with the sensors other than video. Hence, we have tried bringing out the significance of the knowledge driven approaches with video sequences. To facilitate the study, it is necessary to apprehend the notion of the terms, basic, primitive human activity and complex human activity. Section II addresses the same and gives more insight in to the terminologies.

HUMAN ACTIVITY RECOGNITION

Human involved Activity Recognition basics

To automatically identify activities performed by persons in the video is a tough task because all human activities are complex and highly uncertain (Chen Nugent et al., 2019). Activity recognition means to identify the actions and intent of various actors from a chain of video scenes. The actions may also be under the influence of the environmental conditions. As activity recognition is a necessity for security applications in any domain, since a couple of decades, the research has captured the attention of various related research communities. Activities can be branded according to their length of action into basic actions, sub activity and complex activity. If you take the sports domain as an example, short action such as a kick, can be considered a basic action. Longer periodic actions such as running can be considered a sub activity. More prolonged activities such as playing tennis can be termed as complex activities, constituting a sequence of basic actions and sub activities. Complex activities may further consist of hour-long activities such as Tennis matches and so on. Sub activities can be efficiently represented as sequences of their constituent basic actions. For these activities, it is observed that the temporal order of the constituting actions becomes more absolute. For example, the basic actions comprising the sub activities such as running, jumping and falling cannot be executed in a different order. On the other hand, complex activities, which are long-term, it is obvious that the activities become more flexible in nature. For instance, in the activity of preparing meals, one human might choose to cut bread first and then put vegetables in the vessel; another might choose to perform these activities in the opposite order. A third person might choose to completely skip the first step or the second. This inconsistency makes the activities of this type more difficult to capture automatically (Chen et al., 2019). This is why best of the state of the art activity analysis frameworks are focused on these sub activity levels or primitive activity recognition and there is research scope for complex activities (Huan et al., 2021).To give more examples on complex activities, if we consider activities for household tasks, office tasks, eating activity, shopping activity, exercising activity and grooming activity, they involve a lot of sub activities and are highly complex. At this modern age, humans do not have any other option but to multitask. Each activity is complex and is itself a long sequence of sub activities. To dig one more example as a complex activity, consider shopping activity. It is a series of sub activities like entering a shop, looking for a cart/basket, going to aisles, picking/viewing items, filling of cart/basket, going to the cashier, paying the bills, going to the exit. In the complex activities there may also be repetition of sub activities, out of order sequence of sub activities and basic actions, etc. filling of cart activity can also be followed by going to another aisle and viewing other items. Hence, Automatic recognition of such complex activities can help video surveillance become intelligent. To explain the terms, Basic actions, Primitive Activities and Complex Activities more formally- Numerous basic actions and primitive activities in certain sequences form the complex activity.

Basic actions are established by merely relating a human with his/her close proximity object in a video frame.
Primitive activities are established by a definite temporal sequence of basic actions and they are not variable.
Complex activities are formed by a long variable sequence consisting of a combination of both basic actions as well as primitive activities, which can be out of order, repetitive. For instance, breakfast activity may be formed by a sequence of basic actions - sitting on a chair, holding a spoon, holding a napkin, and so on. The action, "Sitting on chair" is called a primitive activity as this cannot be split further down (Huan et al., 2021). Breakfast activity is considered a complex activity as it involves a sequence of primitive activities.

Human involved Activity Recognition approaches

In this research area, there are different approaches conferred. To broadly categorize, two approaches are identified through the study. First one is the Data Driven approach and the second one is the Knowledge Driven approach (Shi et al., 2022)(Zhang et al., 2023)(Chen and Nugent ,2011). Data oriented activity recognition method uses existing huge datasets of human activities to create human activity models and then uses these learned models to infer on human activities (Bijalwan et al., 2022). This approach uses mining of data and deep/ machine learning practices to learn activity models and to apply them to deduce the activities from the new set of input sensor data (Wang et al., 2020). However, these learned activity models cannot be generally applied to all categories of human activities. A knowledge focused method for identification of activities makes use of rich former knowledge to build models of these activities. These are used to infer human based activities (Chen et al., 2019). This approach uses logic based or semantic technology-based methods to model prior knowledge of human activities as a knowledge base (Chen et al., 2019). These models are further used to infer on new real time data.

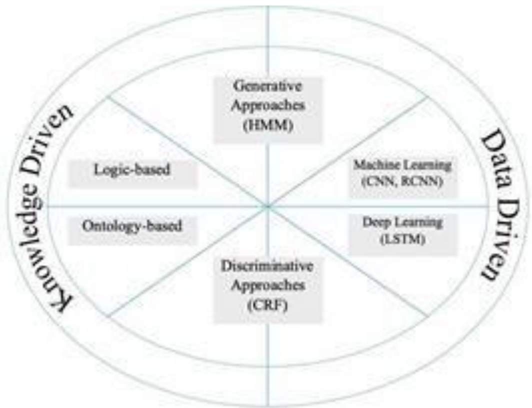

Fig.1: Human Activity Recognition Approaches

Fig.1 shows the human activity recognition approaches. To dig more into the approaches based on both data and prior knowledge, and explain the different methods/models, adopted by them, we will brief on activity recognition through vision based techniques, machine learning techniques, hybrid techniques, logic and knowledge based techniques. To start with, techniques based on vision were the first of the methods to achieve activity recognition as derived from the research. This is a very essential and stimulating challenge to track and know the actions of humans by videos through cameras. The major practice engaged is computer vision. Visual-based activity recognition has caused abundant uses for human computer interface, surveillance etc., you can find a great deal of work in this direction. Researchers have explored techniques like Optical flow, kalman filtering and HMM etc., under diverse means such as video, audio, etc (Kushwaha and Khare , 2023). Adding to it, several facets in this domain, counting pedestrian tracing, group tracing, and identifying abandoned things have been studied. These techniques, if augmented with common sense and contextual perspectives, can be a great accomplishment in the right direction. Machine learning is to make system learn by building mathematical models based on existing data, in order to make decisions on new set of input data without explicitly programming for it. Well known techniques here are Artificial Neural Network (ANN), Convolution Neural Network (CNN), Recurrent CNN, Two-stream CNN, Deep nets, Support Vector Machine (SVM), Decision Trees, and Bayesian Networks (Gull et al.,2022)(https://en.wikipedia.org/wiki/Machine_learnin g#Artificial_neural_networks). ANN models the system through a collection of connection nodes called neurons, imitating our brain neurons. These neurons are aggregated into layers (Srikanth Sagar Bangaru et al., 2021). ANN possesses multiple layers; each layer performs different transformations on its input and provides output to the next layer. Signals travel from input to output layer after traversing middle layers multiple times until error between the expected and actual output is acceptable. Deep nets possess multiple hidden layers compared to ANN. These models make the system learn activity classes and recognize the activities through processing. SVM is supervised method-based learning appropriate for classification problems. Models are built to classify activities to appropriate activity classes (Chathuramali and Rodrigo, 2012). Bayesian Networks model systems using probabilistic relations between activities to predict Dynamic Bayesian Networks (DBN) relates nodes to each other over adjacent time steps to capture temporal relationships or sequence for recognizing complex activities (Kumar et al.,2022). Convolutional Neural Network (CNN) is a simple neural network. It uses convolution instead of common matrix multiplication for at least one layer (Ian Goodfellow et al., 2016). All these models need huge data to train the models to make the models robust. Hence, they face cold start problems as mentioned in table 2 of section II. To recognize complex activities, it is learning more than the relationship between the objects to be identified, which is just enough to identify actions like person standing, person with a bag etc. It is needed to trace the sequence of frames in order to recognize primitive activities like walking, sitting, waving etc. This requires extracting spatial, temporal dependency between the objects existing in the frames. Further, to identify complex activities with longer durations, sequences of the identified primitive activities are to be tracked through long sequences. This requires hybrid models with the stated abilities to be established. Knowledge based techniques range from logic based, grammar based to Ontology based approaches. A logic-based technique keeps record of all the observed actions and monitors their logical consistency. Thus, all probable and reliable activity plans or goals must be considered. This task can be considered as a logical interpretation method of restriction (Chen et al., 2013). This approach is basically a goal-oriented method and knowledge on activities is symbolized by a number of first order statements coded in first order logic. Unpredictable end goals are fine-tuned when new steps are learned. Biggest challenge for logic-based techniques is handling uncertainty. Dynamic learning is difficult with logic-based methods. Domain knowledge-based techniques for activity recognition is the established technique through the decades. As it is difficult to imitate the human brain and its capabilities, working towards that, domain expertise can be modeled as a knowledge base to be used to infer on the activities. Modeling domain knowledge is best done by Ontologies (Rashmi and Krishnan, 2017)(https://www.ontotext.com).

Human Activity Recognition phases

Human Activity Recognition has evolved from recognition of basic actions to complex activities (Maurya et al., 2022). Though the current state of the research focus is mainly on Complex Human Activity Recognition (CHAR), research in all these phases is equally significant. Unlike the human eye, machines have to start with the basic step of identifying objects in the frame, recognize basic actions and primitive activities in order to arrive at complex activities. Fig 2 represents the stages of activity recognition in terms of how Activity recognition research has evolved. Almost equal amounts of research have occurred in each of these phases causing the real time analysis of videos to recognize normal or abnormal activities. As robust as the basic step, that accurate is the final activity recognition step. Table 1 lists the phases of activity recognition along with the models and approaches applicable to each phase, and certain examples depicting each phase of the above-mentioned process.

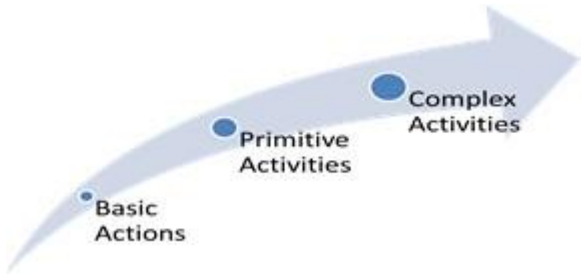

Fig 2: Evolution of Activity Recognition phases

Human involved Activity Recognition through visual sensors

Video sensors involved in Human Activity Recognition have to go through basic video processing steps first. Video sequences are split into frames first in order to begin any kind of processing. Then many of the state-of-the-art models such as YOLO may be used to identify the objects in the frame (https://pjreddie.com/media/files/papers/YOLO v3.pdf). Once objects are identified, basic action has to be established from the frames. Work presented in (Kushwaha and Khare, 2023)(Li Choo and Chuah Sbgar, 2017)(Gao et al., 2013)(Tanget al., 2018)(Kongand Fu, 2022) discusses the optical flow method for action recognition. Works (Chathuramali and Rodrigo ,2012)(Latahet al., 2017) discuss the SVM method for action classifications. References listed at (Latahet al., 2017)(Gedat et al., 2017)(Tranet al., 2015) (Ji et al., 2012)(Simonyan and Zisserman, 2014)(Karpathy et al., 2014) and more discuss the NN and CNN methods for action recognition. Generative models such as HMM for action recognition are listed at (Gedat et al., 2017)(Ahmad et al., 2006). Decades of research gone into this phase has ensured almost 100% accuracy with all the listed promising models. Recognizing primitive activities is a different approach than recognizing the basic actions. This is because a single frame cannot recognize the primitive activities like running, jumping, waving and sitting. Number of continuous frames has to be tracked and learned. Hence approaches like Machine learning, Deep learning, Generative and Discriminative approaches, which can handle sequences of frames, are best suited for primitive activities. These models are realized through the papers listed in the reference section from (Pienaar and Malekian, 2019) to (Castro et al., 2015) and more.

Complex activity recognition is still the challenge to the research community in matching the perceiving accuracy of a human vision. Lot of factors contribute to it. Human activities are uncertain to a very high degree. A human can perform one activity in diverse ways. Sequences of primitive activities involved may happen in shuffled order for the same complex activity. Primitive activities in a sequence can be repeated or can form a loop. Concurrent and interleaving primitive activities are common amongst complex activities. Hence, any one of the above models is not sufficientto handle the complex activity recognition. Combination of the models such as Deep learning with probability-based approaches or knowledge driven approaches with probability-based approaches are good options. Research contributions (Ranasinghe et al., 2016)(Ramasamy Ramamurthyet al., 2018) (Vrigkas and Nikou Kakadiaris, 2015)(Sreenu and Durai,2019)(Sakr et al., 2018)(Kaleet al., 2019)(Srivastavaet al., 2014) Showcase the current state of the approaches required for complex activity recognition.

Research questions to be addressed in the section III and IV

RQ1. What are the well-recognized Complex Human Activity Recognition methods based on learning through data?

RQ2. What are the benefits of having video sensors for Complex Human Activity Recognition?

RQ3. List the knowledge focused approaches for

Complex Human Activity Recognition.

RQ4. What role Ontology plays in Complex Human Activity Recognition?

RQ5. Discuss the research contributions for Ontology based Complex Human Activity Recognition

Section III responds to the first two of the identified research questions and gives the details on Human Activity Recognition with data-based approaches, knowledge-based approaches, and section IV explains Ontology based approaches with its representation aspects and associated reasoning techniques. Section IV is followed by the conclusion of the survey with open research challenges to be addressed further.

Summary on Approaches based on DATA AND knowledge

Human Activity Recognition Steps

The phases that occur while recognizing human activities are described in table 1, as identified through the in-depth research survey. This table highlights the appropriate methods for each phase, possible identifiable actions/activities along with the research work involved. This clearly brings out the difference between the notions of human activities as identified.

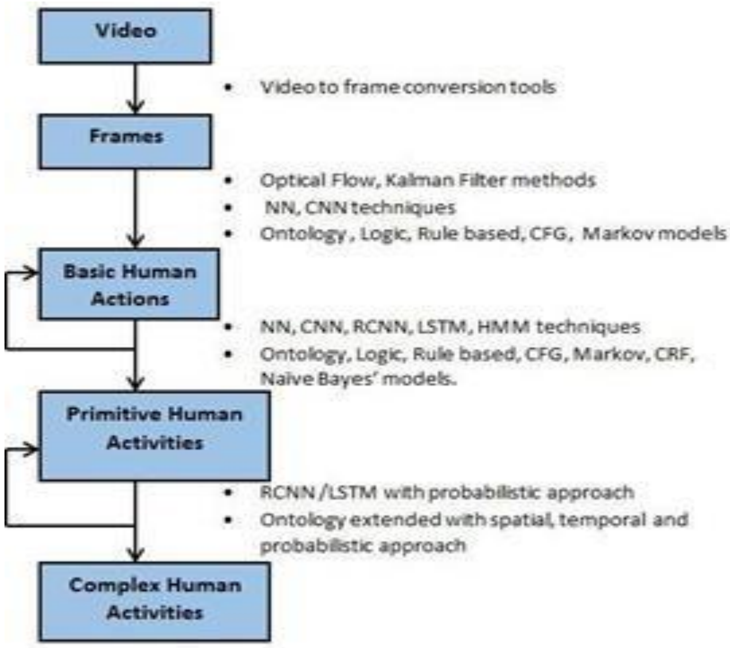

Fig3: Flowchart: Human activity recognition

Fig. 3 provides the steps for the complex human based activity recognition, starting from videos to the final complex activities. This also lists the techniques/approaches applicable for each step.It clearly brings out the repetitive sequence of basic actions and primitive activities showcasing the uncertain behavior of activities when humans are involved. Fig. 3 provides the steps for the complex human based activity recognition, starting from videos to the final complex activities. This also lists the techniques/approaches applicable for each step. It clearly brings out the repetitive sequence of basic actions and primitive activities showcasing the uncertain behavior of activities when humans are involved.

Human Activity Recognition Characteristics to evaluate the approaches

Table 2 lists the characteristics of both the approaches listed, showcasing their capabilities towards complex human activity recognition. The characteristics range from model building, activity representation, and system start issues, dynamic learning to different variations of complex human activities. This brings out the fact that when complex activities are to be recognized, models training on huge data are not preferable. This will be a better idea to exploit existing domain expertise through knowledge driven ontology kind of a model to begin the system with. This knowledge is extendible with the learning aspect incorporated.

Table 1: Human Activity Recognition Phases

HAR Phases	**Methods used**		**Activities identified**	
	Data Driven	**Knowledge Driven**	**Actions/Activities**	
Basic actions	Optical Flow, Kalman Filter, NN, CNN, RCNN, LSTM, HMM, CRF, Naïve Bayes’	Ontology, Logic, Rule based, CFG, Markow models	• Person standing • Person sitting • Etc.	(Zhang et al.,2017)(Sun et al.,2010)(Florence Simon et al.,2019)(Wanyan et al.,2024)(Latah ,2017)(Gedat et al.,2017)(Tran et al.,2015)(Ji et al.,2012)(Simonyan and Zisserman ,2014)(Karpathy et al.,2014)(Ahmad and Lee ,2006)(Schuldt et al.,2004)(Ullah et al.,2017) (Wanyan et al.,2024)(Tran et al.,2015)(Ji et al.,2012)(Pienaar and Malekian ,2019)(Deng et al.,2015)
Primitive activities	NN, CNN, RCNN, LSTM, HMM, CRF, Naïve Bayes’	Ontology, Logic, Rule based, CFG, Markow models	• Person running • Person lifting • Etc.	(Baradel et al.,2018)(Schuldt et al.,2004)(Ryoo ,2011)(Duong et al.,2005)(Wang et al.,2006) (Choi et al.,2011)(Varol et al.,2017)(Yang and Shah ,2012)(Ryoo and Aggarwal ,2006)(Luo et al.,2014)(Wei and Shah,2017)(Castro et al.,2015)
Complex activities	RCNN /LSTM with probabilistic approach extended to track sequence of primitive activities	Ontology, Logic, Rule based, CFG, Markow models	• Person cooking • Person eating • Person playing • Etc.	(Li ,2017)(Swears et al.,2014)(Hussein et al.,2019)(Ramasamy Ramamurthy and Roy ,2018)(Alexiou et al.,2015)(Liuet al.,2018)

Table 2: Characteristics of the activity recognition approaches influencing the choosing of methods.

Characteristics	Data Driven approach	Knowledge Driven approach
Model building	Dynamic, Uses existing data to build the model	Static, knowledge driven approach builds model which replicate the existing domain expertise
Action/Activity representation	Difficult as long temporal sequences pose challenges	Easy using domain expertise
System start	System start has a cold start problem, as huge data is needed to train the model in case of unsupervised learning.	Does not have a cold start problem. Domain knowledge base is built which is used to reason on input data
Dynamic Learning	Yes, new identified activity outputs can be learned.	Models need to be extended with learning capability
Handling Uncertainty	Yes, with fuzzy and probability-based approaches' support	Models to be extended with probabilistic methods to handle uncertainty
Handling Shuffled sequence of primitive activities	Difficult to learn on all sequences	Domain expertise can handle this with the rule sequences
Handling concurrent and interleaving activities	Yes, with fuzzy and probability-based approaches' support	Yes, with fuzzy and probability-based approaches' support

Data Driven Approaches summarized

Data driven approaches have evolved from basic video processing to, making machines learn to recognize the activities. Table 3 has summarized the data driven approaches and their pros, cons. Approaches like SVM to Machine Learning to Deep Learning, all the methods are studied.

Knowledge Driven Approaches

Knowledge based approach, as established to be an important option, offers a great range of opportunities through its various methods for complex human action recognition. The table 4 gives an insight in to the existing research on these approaches, along with their pros and cons.

Table 3: Data Driven Approaches

Data Driven Approaches				
Models	**Description**	**Advantages**	**Disadvantages**	**Relevant References**
Image/Video processing SVM/BOW /k-NN	Use classification, sequential, space time approaches	Action classification is quick and promising with a high level of accuracy Robust to noise	Complex action representation will be difficult, not expressive enough	(Zhang et al.,2017)(Sun et al.,2010)(Florence Simon ,2019)(Wanyan et al.,2024)(Latah ,2017)(Schuldtet al.,2004)(Ryoo ,2011)(Duong et al.,2005)(Riboni et al.,2011)

Machine Learning based on ANN, CNN/RCNN	Use huge data to make the model learn and infer on the new set of data to recognize actions	Dynamic in representation and reasoning	Huge data to train the model and failing to capture long temporal information of video	(Wanyan et al.,2024)(Gedat et al.,2014)(Tran et al.,2015)(Ji et al.,2012)(Simonyan ,2014)(Karpathyet al.,2014)(Pienaar ,2019)(Baradel et al.,2018)(Kong et al.,2012)(Luo et al.,2014)(Castroet al.,2015)(Ullah et al.,2017)
Generative and Discriminative	These methods map the input states and produce description to activity labels principally by considering video frames as temporal sequences.	Suitable for sequential activities. Less computational burden than DBN Good in managing uncertainty	Major dependency on human silhouettes making it is challenging for real world applications. Sensitive to noise	(Li ,2017)(Swears et al.,2014)(Wanyan et al.,2024)(Gedat et al.,2017)(Ahmad and Lee ,2006)(Alexiouet al.,2015)(Wang et al.,2006)(Liu et al.,2018)(Ramasamy Ramamurthy ,2018)
Deep Learningbased on LSTM,DBN	Uses multiple learning layer to extract higher level features from data	Network can be easily scaled up forthe increasing data, can learn dependencies between variables	To work on temporal modelling difficult to train, more computational burden	(Tran ,2015)(Ji ,2012)(Varol et al.,2017)(Yang ,2012)(Hussein et al.,2019)(Wei,2017)
Hybrid approach es	Uses Graphical models like HMM with DBN or ML with HMM	Handling temporal relations	Need to work on concurrent activities	(Karpathy et al.,2014)(Hussein et al.,2019)(Zhang et al.,2013)

Table 4: Knowledge Driven Approaches

Knowledge Driven Approaches				
Models	Description	Advantages	Disadvantage s	Relevant References
Semantic models	Semantic models describe inherent characteristics of activities very well even in case of unclear flow of activities performed by humans.	These based models are suitable to represent rich temporal relations amongst actions, along with spatial relations well captured	Lack of expressivity to Symbolize uncertainties within temporal relations. Framing the semantic formula with the associated weights from scratch is cumbersome	(Kushwaha and Khare ,2023) (https://pjreddie.com/media/files/papers/YOLO v3.pdf)(Swears et al.,2014)(Sun et al.,2010)(Florence Simon ,2019)

			in real world scenarios.	
Context Free Grammar -based model	The CFG-based representation facilitates formal description of complex human activities based on basic actions.	Representation and recognition of actions like hug, punch, push, hand-shake with high accuracy	Uncertainty in complex activities recognition is lacking. To frame CFG grammar by scratch is tedious. Not good for concurrent activities	(https://en.wikipedia.org/wiki/Machine_learning#Artificial_neural_networks)(Rashmi ,2017)(Maurya et al.,2022)
Markov logic network	MLN Knowledge base is explored	Activities can be easily inferred along with the time points indicating the whole duration of activity.	Formulae used may affect recognition rate.	(Bangaru et al.,2021)(Chathuramali ,2012)(Kumar ,2022)(Goodfellow ,2016)(Chen et al.,2013)
Fuzzy, spatial, temporal logic based	Temporal, spatial and Fuzzy action semantics in the knowledge base are captured.	Handling uncertainty to some extent	Dynamic learning to update knowledge base to be added	(Babangida et al.,2022)(Bijalwan et al.,2022)(Wang et al.,2020)(Chen ,2019)
For Ontology based approach refer table 6				

List of Application domains for Activity Recognition with the appropriate input sensors.

Based on the requirements, the activity recognition process has to adopt the input sensor type. If the application domains are of indoor type, such as assisted living, smart homes, and hospitals, need sensors other than video for taking the input. One reason may be because of privacy issues. Another reason is to avoid the video processing challenges and overhead. And, if a smaller number of humans and objects are involved in the act, then sensors other than video can be preferred. But in some of the public domains, with a greater number of actors and objects, having a lot of sensors on all the objects is not a preferred choice. This will be the situation in most of the domains such as traffic, retail malls, bus and railway stations, airports, and security sensitive domains etc. Table 5 lists application domains with preferred sensors.

Table 5: Application domains with preferred sensors

Application domain	Preferred sensors types
Health care	• Wearable sensors
Ambient assisted living	• Sensors attached to surrounding objects/RFID or beacon sensors • Wearable sensors

Smart homes	• Sensors attached to surrounding objects/RFID or beacon sensors • Wearable sensors • Ambient sensors • Temperature sensors
Natural disaster	• Visual sensors
Retail/Mall	• Visual sensors
Defense	• Visual sensors
Classroom	• Visual sensors
Other Public places	• Visual sensors
Traffic monitoring	• Visual sensors

In this section, sub section A discussed the phases of the activity recognition, sub section B discussed the characteristics of activity recognition approaches, sub section C answered the research question RQ1, subsection D answered research question RQ2 and sub section E answered RQ5.

HUMAN ACTIVITY RECOGNITION USING ONTOLOGY

This section brings out the significance of using ontology as a knowledge driven model for activity representations and reasoning for important activities recognition. This highlights the salient and features of ontology as an activity-modelling tool. Also presents a survey of ontology-based research.

What role an Ontology plays in Complex Human Activity Recognition

Ontology is a modelling means for knowledge representation. It offers to represent worldly concepts and relations amongst them, enables sharing and reusing of knowledge. Knowledge can be updated with new concepts easily. Ontology provides an official description of knowledge. To create ontology, we need to formally put down constituents such as objects, classes, characteristics, and associations as well as constraints, rubrics and axioms. Ontologies try to replicate brain functions. So, they help systems to perceive the world like humans do.

[1] Some of the major features of ontologies are that they guarantee a shared understanding of knowledge and that they make clear domain norms. With the essential relationships between concepts, the automated reasoning is possible. Ontologies afford a more intelligible and easy steering for users to link concept to concept. Extendibility is the additional valued characteristic of ontologies. This is because new relationships and concepts are easy to be added to current ontologies. Last decade has witnessed an acceptance of OWL as the major Ontology language amongst all. OWL is a semantics and logic driven computational language, intended to signify opulent and compound knowledge. Owl supports Ontology models with constant and logical distinctions between concepts, characteristics and relationships between concepts. (https://www.ontotext.com)

Ontology concepts and relationships clearly capture the activity relationships. These are expressed using Ontology language, OWL that can be machine processed further into PROLOG or JESS environments. These environments use the rule engine to infer on the activities. Rules can also be mapped

onto states of the state machine to process input videos to recognize activities (https://www.ontotext.com).

Survey of Ontology based HAR

The below table 6, lists the Ontology based works for activity recognition. There is a substantial amount of study found to prove usage of ontology for capturing domain concepts and their relationships as given by experts. It is observed to be less amount of work on video-based recognition and more on other sensor based. This is due to the initial hurdle of processing videos to get the basic actions identified. But as today's robust video processing techniques provide almost 100% accuracy for basic human actions from a single or few numbers of frames, initial video processing is no more a challenge. This survey brings out the salient features ofontology, which are flexibility, expressing ability, information sharing capability, extendibility, and reusability.

Table 6: Ontology based Activity Recognition survey

SL. No.	Relevant reference	Year	Description	Dataset
1	(Schuldt et al.,2004)	2011	Ontological approach is assessed in smart home dataset. It shows that ontology can be extended for temporal reasoning, with efficiency comparable to the methods that support temporal based reasoning. Also showcased ontological reasoning to identify activities with the help of context information.	State changing Sensors
2	(Ryoo ,2011)	2014	Focus of the paper is a complete analysis of ontologies for flexibility, reasoning, information sharing, and knowledge representation, which showed ontologies as well capable techniques for activity recognition.	Sensors
3	(Duong et al.,2005)	2012	Presented Activity Recognition based on Ontology (OBAR) system which recognized human doings with great accuracy, and facilitated the searching of the semantic information in the system.	Sensors
4	(Wang et al.,2006)	2017	The focus of the paper is, representing the concepts using RDF and ontologies to investigate the behavior of the crowd.	Visual sensors
5	(Choi et al.,2011)	2014	Ontology is combined with a temporal approach to provide amodel capable of composite activities modelling and reasoning.	Sensors
6	(Liu et al.,2018)	2014	Ontological and data driven learning models are combined for smart home activities recognition.	Sensors
7	(Riboni et al.,2011)	2012	This paper makes use of ontology to infer a continuous stream of activities through different means in a smart home setup. It uses semantic reasoning on the ontologies to infer on the activities.	Sensors
8	(Varol et al.,2017)	2007	A hybrid model consisting of ontology and probability to perceive human activity based on humans present whereabouts. RFID is used to monitor the human actions, which are then modeled using ontology, representing the things, and actions and naive Bayesian method is used for helping with real location semantics from the terms.	Sensors
9	(Fenz ,2012)	2012	The developed method enables the efficient construction and modification of Bayesian networks based on existing ontologies.	Visual Sensor
10	(Klein ,2007)	2007	With an aim to provide assistance to the elderly to a context ontology is created. OWL-Lite version of Web Ontology	Sensors

			Language was used to construct the ontology, which was then used as the reference to deduce the actions.	
11	(Riboni and Bettini ,2011)	2011	The work showcases features of OWL 2 over the main shortcomings of OWL 1 for the activities	Sensor
12	(Magherini et al.,2013)	2013	The paper concentrates recognizing daily happening activities automatically with the help of ADL ontology clubbed with propositional Temporal Logic to satisfy the expressive wants of Ambient assisted living.	Home automation sensors
13	(Rafferty et al.,2017)	2017	The paper introduces intention recognition mechanism combined into intelligent agent design. Activities are modelled as goal ontology and sensors are used to recognize atomic actions are modelled in belief ontology.	Bluetooth low energy estimote sticker beacons
14	(Crispim-Junior et al.,2016)	2016	A framework is presented using Ontology Language based on Web (OWL) for activity modeling. To handle uncertainty, probability based inference methods are made use of.	Visual Sensor
15	(Schlenoff et al.,2012)	2012	Robotics ontology is developed and further extended to include state information. This paper recognizes the intentions by identifying the state of multiple people working together.	Robots sensor system
16	(Rafferty et al.,2015)	2015	Work proposes a goal driven approach wherein a goal repository is built using ontology and goals of inhabitants residing in smart home are inferred by recognizing their atomic actions	NFC tags
17	(Mayer et al.,2016)	2007	This paper recognizes human activities with the help of location semantics; the locations are modelled in ontology and inferred through Naive Bayes method.	RFID
18	(Latfi et al.,2007)	2007	Ontology is used to describe the domain and to load the Bayesian network to infer the activities of elderly suffering from cognitive impairment.	Sensors and other remote monitoring material
19	(Al-Wattar et al.,2016)	2016	The approach presented in this paper combines spatial and temporal contexts of a person and recognizes the activity to initiate actions as required.	Conventional Cameras
20	(Greco et al.,2016)	2016	Semantic web technology is used to annotate the output of people tracing algorithms. Ontology is used to model the domain and semantic reasoned with rules is used to recognize the simple and complex events.	IP Cameras
21	(PRIYA ,2015)	2015	This paper proposes Ontology building for video frames with spatial, temporal and topological relations explored between them.	Surveillance cameras

22	(McKeever et al.,2010)	2010	Activities such as sleeping, leaving home, sleeping, preparing breakfast, preparing dinner, and preparing a drink were recognized in a scenario of smart home using sensors. This was done on the fact that certain activities takes place at certain time intervals	Smart home sensors
23	(Okeyo et al.,2012)	2012	The paper concentrates on activity modeling. The activities are divided into three categories: simplex activity, composite activity and actions. Owl and temporal is combined and the activity modeling graph is shown.	Sensors
24	(Meditskos et al.,2013)	2013	Describes a hybrid framework called SP-ACT, which combines the owl ontology and the SPQRL rules for high-level activity recognition. The framework consists of the interpretation layer and the representation layer. The activities take place inside a temporal boundary, which have HasStartTime and HasEndTime as the properties.	Atomic Activities (Cameras, microphones, etc.)
25	(Le Yaouanc et al.,2012)	2012	Combine fuzzy and spatial temporal approach for activity recognition. Various relations such as going away, moving along etc. are defined. The basic principles of fuzzy logic are explained and activities are modeled using situation graphs. Spatial temporal relations along with uncertainty are modeled using boundaries. This fuzzy based expert system that considers vagueness in spatial and temporal relations.	Sensors
26	(Zhu et al.,2012)	2013	Activity recognition with a mathematical model along with context and Spatial-temporal relationships is described for anomaly detection in videos.	Video
27	(Onofri et al.,2016)	2016	In this work, emerging trends of knowledge-centered methods for human activity recognition have been surveyed.	Video
28	(Amirjavid et al.,2011)	2011	The work concentrates on uncertainty involved while recognizing activities of daily living. Domain Knowledge is represented using fuzzy concept ontology and reasoning is done using fuzzy rules. Temporal and spatial extensions are also added to the fuzzy concept ontology.	RFID
29	(Manzoor et al.,2021)	2021	In this article, ontology-based semantic representation unified into the current state of robotic knowledge base systems is reviewed to explore mainly the recent developments in ontology-based knowledge representation systems that have led to the effective solutions of real-world robotic applications	Video
30	(Chen et al.,2019)	2019	In the vision-based activity recognition community, researchers have realized that symbolic activity definitions based on manual specification of a set of rules suffer from limitations in their applicability, i.e., the definitions are only deployable to the scenarios for which they have been designed.	Sensors

			There is a need for an explicit commonly agreed representation of activity definitions, i.e., ontologies, for activities that are independent of algorithmic choices, thus facilitating portability, interoperability and reuse and sharing of both underlying technologies and systems.	
31	(Abishek and Nair ,2023)	2024	The paper focuses on ontology-based action recognition in sports videos, achieving 96% accuracy. The semantic verification model proposed in this work uses ontology and relies on rules and logic that can be easily interpreted enhancing accuracy.	Videos
32	(Abishek and Nair ,2023)	2023	The contributions of this work lie in its advancements in Humar activity recognition, sensor technology, and maintaining non-intrusiveness to develop personalized support systems; this improves real-time decision-making, and the potential for diverse applications in enhancing daily living and safety for users.	State of the art sensors
33	(Foudeh and Salim ,2023)	2023	The paper proposes a fully probabilistic ontology model for human activity recognition, enhancing flexibility with multiple candidates and improved recognition rates, addressing scalability and efficiency challenges in current research.	Hybrid sensors
34	(Jabla et al.,2022)	2022	The paper proposes a knowledge-driven activity recognition framework for learning unknown activities, utilizing ontology.	Sensors
35	(Hooda and Rani ,2020)	2020	The paper introduces an ontology-driven model for human activity recognition using sensor data, enhancing reusability and interoperability while addressing sensor heterogeneity in activity recognition systems.	Heterogeneous Sensors
36	(Noor et al.,2020)	2020	The paper investigates ontology-based sensor fusion for activity recognition, enhancing accuracy by combining wearable and ambient sensors to infer activities more effectively, achieving 91.5% recognition accuracy.	Sensors

RESULTS AND DISCUSSION

This section indicates the related work reviewed on CHAR approaches for video sensor applications as well as other applications like home automation and healthcare. Fig 4 shows the most relevant work on these approaches, also tabulated below in table 7.

Table 7: Existing Work on CHAR approaches

Approaches for CHAR	No. of most relevant papers
Data driven methods for visual sensors	54
Data driven methods for other sensors	10
Knowledge driven approaches for visual sensor	34
Knowledge driven approaches for other sensor	17

Video surveillance is need of the hour for the security in public domains. This study brings out, especially Table 4 and Table 6, there is a scope for Ontology driven knowledge-based approaches for Complex Human Activity Recognition in video sequences.

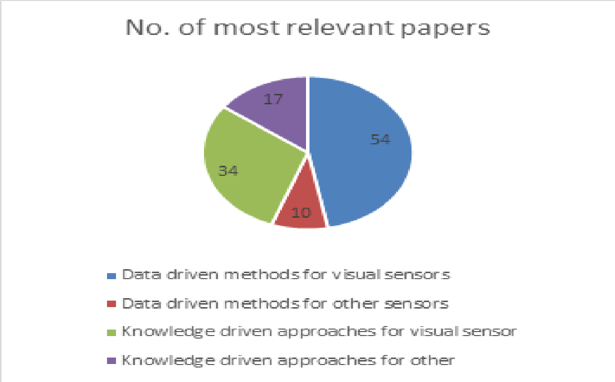

Fig 4: Graph of Existing Work on CHAR approaches

Data Availability Statement

Data sharing does not apply to this article as no new data has been created or analyzed in this study.

Funding Information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

REFERENCES

[1] A,Hussain T,Hussain Ullah SW.W, Baik Vision transformer and deep sequence learning for human activity recognition in surveillance videos. Computational Intelligence and Neuroscience. 2022;2022(1):3454167.. https://doi.org/10.1155/2022/3454167

[2] Abhilash, C. B., & Mahesh, K. (2022, November). Ontology-based method for semantic association rules. In *2022 IEEE 19th India Council International Conference (INDICON)* (pp. 1-7). IEEE.

[3] Abishek, V. S., & Nair, A. (2023, October). Human Activity Recognition using Ontology. In *2023 4th IEEE Global Conference for Advancement in Technology (GCAT)* (pp. 1-7). IEEE.

[4] Ahmad, M., & Lee, S. W. (2006, August). HMM-based human action recognition using multiview image sequences. In *18th international conference on pattern recognition (ICPR'06)* (Vol. 1, pp. 263-266). IEEE.

[5] Alexiou, I., Xiang, T., & Gong, S. (2015, September). Learning a joint discriminative-generative model for action recognition. In *2015 International Conference on Systems, Signals and Image Processing (IWSSIP)* (pp. 1-4). IEEE.

[6] Al-Wattar, M., Khusainov, R., Azzi, D., & Chiverton, J. (2016, September). Activity recognition from video data using spatial and temporal features. In *2016 12th International Conference on Intelligent Environments (IE)* (pp. 250-253). IEEE.

[7] Amirjavid, F., Bouchard, K., Bouzouane, A., & Bouchard, B. (2011). Spatiotemporal knowledge representation and reasoning under uncertainty for action recognition in smart homes.

[8] Arrotta, L., Civitarese, G., & Bettini, C. (2022, June). Knowledge infusion for context-aware sensor-based human activity recognition. In *2022 IEEE International Conference on Smart Computing (SMARTCOMP)* (pp. 1-8). IEEE.

[9] Arrotta, L., Civitarese, G., & Bettini, C. (2023). Probabilistic knowledge infusion through symbolic features for context-aware activity recognition. *Pervasive and Mobile Computing*, *91*, 101780.

[10] Arrotta, L., Civitarese, G., Presotto, R., & Bettini, C. (2023, July). DOMINO: A Dataset for Context-Aware Human Activity Recognition using Mobile Devices. In *2023 24th IEEE International Conference on Mobile Data Management (MDM)* (pp. 346-351). IEEE.

[11] Babangida, L., Perumal, T., Mustapha, N., & Yaakob, R. (2022). Internet of things (IoT) based activity recognition strategies in smart homes: a review. *IEEE sensors journal*, *22*(9), 8327-8336..

[12] Babangida, L., Perumal, T., Mustapha, N., & Yaakob, R. (2022). Internet of things (IoT) based activity recognition strategies in smart homes: a review. *IEEE sensors journal*, *22*(9), 8327-8336.

[13] Bangaru, S. S., Wang, C., Busam, S. A., & Aghazadeh, F. (2021). ANN-based automated scaffold builder activity recognition through wearable EMG and IMU sensors. *Automation in Construction*, *126*, 103653.

[14] Baradel, F., Neverova, N., Wolf, C., Mille, J., & Mori, G. (2018). Object level visual reasoning in videos. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 105-121).

[15] Beddiar, D. R., Nini, B., Sabokrou, M., & Hadid, A. (2020). Vision-based human activity recognition: a survey. *Multimedia Tools and Applications*, *79*(41), 30509-30555.

[16] Bijalwan, V., Semwal, V. B., & Gupta, V. (2022). Wearable sensor-based pattern mining for human activity recognition: Deep learning approach. *Industrial Robot: the international journal of robotics research and application*, *49*(1), 21-33.

[17] Buoncompagni, L., Kareem, S. Y., & Mastrogiovanni, F. (2021). Human activity recognition models in ontology networks. *IEEE Transactions on Cybernetics*, *52*(6), 5587-5606.

[18] Castro, D., Hickson, S., Bettadapura, V., Thomaz, E., Abowd, G., Christensen, H., & Essa, I. (2015, September). Predicting daily activities from egocentric images using deep learning. In *proceedings of the 2015 ACM International symposium on Wearable Computers* (pp. 75-82).

[19] Chathuramali, K. M., & Rodrigo, R. (2012, December). Faster human activity recognition with SVM. In *International conference on advances in ICT for emerging regions (ICTer2012)* (pp. 197-203). IEEE.

[20] Chen, L., Nugent, C. D., Chen, L., & Nugent, C. D. (2019). An Ontology-Based Approach to Activity Recognition. *Human Activity Recognition and Behaviour Analysis: For Cyber-Physical Systems in Smart Environments*, 49-75.

[21] Chen, L., Nugent, C., & Okeyo, G. (2013). An ontology-based hybrid approach to activity modeling for smart homes. *IEEE Transactions on human-machine systems*, *44*(1), 92-105..

[22] Chen, S., Liu, J., Wang, H., & Augusto, J. C. (2013, October). A hierarchical human activity recognition framework based on automated reasoning. In *2013 IEEE International Conference on Systems, Man, and Cybernetics* (pp. 3495-3499). IEEE.

[23] Choi, W., Shahid, K., & Savarese, S. (2011, June). Learning context for collective activity recognition. In *CVPR 2011* (pp. 3273-3280). IEEE.

[24] Crispim-Junior, C. F., Buso, V., Avgerinakis, K., Meditskos, G., Briassouli, A., Benois-Pineau, J., ... & Bremond, F. (2016). Semantic event fusion of different visual modality concepts for activity recognition. *IEEE transactions on pattern analysis and machine intelligence*, *38*(8), 1598-1611.

[25] Deng, Z., Zhai, M., Chen, L., Liu, Y., Muralidharan, S., Roshtkhari, M. J., & Mori, G. (2015). Deep structured models for group activity recognition. *arXiv preprint arXiv:1506.04191*.

[26] Dhiman, C., & Vishwakarma, D. K. (2019). A review of state-of-the-art techniques for abnormal human activity recognition. *Engineering Applications of Artificial Intelligence*, *77*, 21-45.

[27] Duong, T. V., Bui, H. H., Phung, D. Q., & Venkatesh, S. (2005, June). Activity recognition and abnormality detection with the switching hidden semi-markov model. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (Vol. 1, pp. 838-845). IEEE.

[28] Fenz, S. (2012). An ontology-based approach for constructing Bayesian networks. *Data & Knowledge Engineering*, *73*, 73-88.

[29] Florence Simon, Prof. Ujwal Harode, “Human Action Recognition using Optical Flow Estimation”, International Research Journal of Advanced Engineering and Science, 2019.

[30] Foudeh, P., & Salim, N. (2023). ONTOLOGICAL, FULLY PROBABILISTIC KNOWLEDGE MODEL FOR HUMAN ACTIVITY RECOGNITION. *Jurnal Teknologi*, *85*(2), 183-199.

[31] Gao, L., Campbell, H. A., Bidder, O. R., & Hunter, J. (2013). A Web-based semantic tagging and activity recognition system for species' accelerometry data. *Ecological Informatics*, *13*, 47-56..

[32] Gedat, E., Fechner, P., Fiebelkorn, R., & Vandenhouten, R. (2017, September). Human action recognition with hidden Markov models and neural network derived poses. In *2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY)* (pp. 000157-000162). IEEE.

[33] Goodfellow, I. (2016). Deep learning.

[34] Greco, L., Ritrovato, P., Saggese, A., & Vento, M. (2016). Abnormal event recognition: a hybrid approach using semantic web. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 58-65).

[35] Gull, I., Selwal, A., & Sabha, A. (2022, April). An Analysis of Video-based Human Activity Detection Approaches. In *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)* (pp. 1554-1561). IEEE.

[36] Gupta, N., Gupta, S. K., Pathak, R. K., Jain, V., Rashidi, P., & Suri, J. S. (2022). Human activity recognition in artificial intelligence framework: a narrative review. *Artificial intelligence review*, *55*(6), 4755-4808.

[37] Herath, S., Harandi, M., & Porikli, F. (2017). Going deeper into action recognition: A survey. *Image and vision computing*, *60*, 4-21.

[38] Hooda, D., & Rani, R. (2020). Ontology driven human activity recognition in heterogeneous sensor measurements. *Journal of Ambient Intelligence and Humanized Computing*, *11*(12), 5947-5960.

[39] https://en.wikipedia.org/wiki/Machine_learnin g#Artificial_neural_networks

[40] https://pjreddie.com/media/files/papers/YOLO v3.pdf.

[41] https://www.ontotext.com

[42] Huan, R., Jiang, C., Ge, L., Shu, J., Zhan, Z., Chen, P., ... & Liang, R. (2021). Human complex activity recognition with sensor data using multiple features. *IEEE Sensors Journal*, *22*(1), 757-775.

[43] Hussein, N., Gavves, E., & Smeulders, A. W. (2019). Timeception for complex action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 254-263).

[44] I,Gull A,Selwal A.Sabha An Analysis of Video-based Human Activity Detection Approaches. In2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) 2022 Apr 7 (pp. 1554-1561). IEEE. doi: 10.47992/ijmts.2581.6012.0318

[45] Jabla, R., Khemaja, M., Buendia, F., & Faiz, S. (2022). A knowledge-driven activity recognition framework for learning unknown activities. *Procedia Computer Science*, *207*, 1871-1880.

[46] Ji, S., Xu, W., Yang, M., & Yu, K. (2012). 3D convolutional neural networks for human action recognition. *IEEE transactions on pattern analysis and machine intelligence*, *35*(1), 221-231.

[47] Kale, G. V. (2019). Human activity recognition on real time and offline dataset. *Int. J. Intell. Syst. Appl. Eng*, *7*(1), 60-65.

[48] Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 1725-1732).

[49] Kim, J. S., Kim, M. G., & Pan, S. B. (2021). A study on implementation of real-time intelligent video surveillance system based on embedded module. *EURASIP Journal on Image and Video Processing*, *2021*(1), 35.

[50] Klein, M., Schmidt, A., & Lauer, R. (2007, September). Ontology-centred design of an ambient middleware for assisted living: The case of soprano. In *Towards Ambient Intelligence: Methods for Cooperating Ensembles in Ubiquitous Environments (AIM-CU), 30th Annual German Conference on Artificial Intelligence (KI 2007), Osnabrück* (Vol. 15)..

[51] Kong, Y., Jia, Y., & Fu, Y. (2012). Learning human interaction by interactive phrases. In Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part I 12 (pp. 300-313). Springer Berlin Heidelberg.

[52] Kumar, K. V., Harikiran, J., & Chandana, B. S. (2022, February). Human Activity Recognition with Privacy Preserving using Deep Learning Algorithms. In *2022 2nd International Conference on Artificial Intelligence and Signal Processing (AISP)* (pp. 1-8). IEEE.

[53] Kushwaha, A., & Khare, A. (2023). Human Activity Recognition in Video Sequences Based on the Integration of Optical Flow and Appearance of Human Objects. In *Robotics, Control and Computer Vision: Select Proceedings of ICRCCV 2022* (pp. 117-125). Singapore: Springer Nature Singapore.

[54] L, Chen CD,Nugent L, Chen Nugent CD. Sensor-based activity recognition review. Human Activity Recognition and Behaviour Analysis: For Cyber-Physical Systems in Smart Environments. 2019:23-47.. doi: 10.1007/978-3-030-19408-6_2

[55] L,Chen CD,Nugent L,Chen Nugent CD. Composite activity recognition. Human Activity Recognition and Behaviour Analysis: For Cyber-Physical Systems in Smart Environments. 2019:151-81.. doi: 10.1007/978-3-030-19408-6_7

[56] Latah, M. (2017). Human action recognition using support vector machines and 3D convolutional neural networks. *Int. J. Adv. Intell. Informatics*, *3*(1), 47-55.

[57] Latfi, F., Lefebvre, B., & Descheneaux, C. (2007, June). Ontology-Based Management of the Telehealth Smart Home, Dedicated to Elderly in Loss of Cognitive Autonomy. In *OWLED* (Vol. 258).

[58] Le Yaouanc, J. M., & Poli, J. P. (2012). A fuzzy spatio-temporal-based approach for activity recognition. In Advances in Conceptual Modeling: ER 2012 Workshops CMS, ECDM-NoCoDA, MoDIC, MORE-BI, RIGiM, SeCoGIS, WISM, Florence, Italy, October 15-18, 2012. Proceedings 31 (pp. 314-323). Springer Berlin Heidelberg.

[59] Li, X., & Choo Chuah, M. (2017). Sbgar: Semantics based group activity recognition. In *Proceedings of the IEEE international conference on computer vision* (pp. 2876-2885).

[60] Liu, L., Wang, S., Hu, B., Qiong, Q., Wen, J., & Rosenblum, D. S. (2018). Learning structures of interval-based Bayesian networks in probabilistic generative model for human complex activity recognition. *Pattern Recognition*, *81*, 545-561.

[61] Luo, G., Yang, S., Tian, G., Yuan, C., Hu, W., & Maybank, S. J. (2014). Learning human actions by combining global dynamics and local appearance. *IEEE transactions on pattern analysis and machine intelligence*, *36*(12), 2466-2482.

[62] Magherini, T., Fantechi, A., Nugent, C. D., & Vicario, E. (2013). Using temporal logic and model checking in automated recognition of human activities for ambient-assisted living. *IEEE Transactions on Human-Machine Systems*, *43*(6), 509-521.

[63] Mahajan, M., RajeshwariBajre, R., Sharma, L., & Varma, S. L. (2020, December). Computer Vision Based Campus Surveillance. In *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)* (pp. 1013-1018). IEEE.

[64] Manzoor, S., Rocha, Y. G., Joo, S. H., Bae, S. H., Kim, E. J., Joo, K. J., & Kuc, T. Y. (2021). Ontology-based knowledge representation in robotic systems: A survey oriented toward applications. *Applied Sciences*, *11*(10), 4324.

[65] Maurya, R., Teo, T. H., Chua, S. H., Chow, H. C., & Wey, I. C. (2022, December). Complex human activities recognition based on high performance 1D CNN model. In *2022 IEEE 15th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoC)* (pp. 330-336). IEEE.

[66] Mayer, S., Verborgh, R., Kovatsch, M., & Mattern, F. (2016). Smart configuration of smart environments. *IEEE Transactions on Automation Science and Engineering*, *13*(3), 1247-1255.

[67] McKeever, S., Ye, J., Coyle, L., Bleakley, C., & Dobson, S. (2010). Activity recognition using temporal evidence theory. *Journal of Ambient Intelligence and Smart Environments*, *2*(3), 253-269.

[68] Meditskos, G., Dasiopoulou, S., Efstathiou, V., & Kompatsiaris, I. (2013, March). 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)* (pp. 25-30). IEEE.

[69] Noor, M. H. M., Salcic, Z., & Wang, K. I. K. (2020). Ontology-based sensor fusion activity recognition. *Journal of Ambient Intelligence and Humanized Computing*, *11*, 3073-3087.

[70] Nurnoby, M. F., & Helmy, T. (2023). A real-time deep learning-based smart surveillance using fog computing: a complete architecture. *Procedia Computer Science*, *218*, 1102-1111.

[71] Okeyo, G., Chen, L., & Wang, H. (2014). Combining ontological and temporal formalisms for composite activity modelling and recognition in smart homes. *Future Generation Computer Systems*, *39*, 29-43.

[72] Okeyo, G., Chen, L., Wang, H., & Sterritt, R. (2012, June). A hybrid ontological and temporal approach for composite activity modelling. In *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications* (pp. 1763-1770). IEEE.

[73] Onofri, L., Soda, P., Pechenizkiy, M., & Iannello, G. (2016). A survey on using domain and contextual knowledge for human activity recognition in video streams. *Expert Systems with Applications*, *63*, 97-111.

[74] PRIYA, K. (2015). EFFECTIVE EXTRACTION OF SEMANTIC CONTENT IN VIDEOS USING META-ONTOLOGY. *Journal of Theoretical & Applied Information Technology*, *73*(2).

[75] Rafferty, J., Chen, L., Nugent, C., & Liu, J. (2015). Goal lifecycles and ontological models for intention based assistive living within smart environments. *Comput. Syst. Sci. Eng*, *30*(1), 7-18.

[76] Rafferty, J., Nugent, C. D., Liu, J., & Chen, L. (2017). From activity recognition to intention recognition for assisted living within smart homes. *IEEE Transactions on Human-Machine Systems*, *47*(3), 368-379.

[77] Ramasamy Ramamurthy, S., & Roy, N. (2018). Recent trends in machine learning for human activity recognition—A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(4), e1254.

[78] Ranasinghe, S., Al Machot, F., & Mayr, H. C. (2016). A review on applications of activity recognition systems with regard to performance and evaluation. *International Journal of Distributed Sensor Networks*, *12*(8), 1550147716665520.

[79] Rashmi, S. R., & Krishnan, R. (2017, February). Domain ontologies and their use in building intelligent systems: A comprehensive survey. In *2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 611-616). IEEE.

[80] Riboni, D., & Bettini, C. (2011). OWL 2 modeling and reasoning with complex human activities. *Pervasive and Mobile Computing*, *7*(3), 379-395.

[81] Riboni, D., Pareschi, L., Radaelli, L., & Bettini, C. (2011, March). Is ontology-based activity recognition really effective?. In *2011 IEEE international conference on pervasive computing and communications workshops (PERCOM workshops)* (pp. 427-431). IEEE.

[82] Rodríguez, N. D., Cuéllar, M. P., Lilius, J., & Calvo-Flores, M. D. (2014). A survey on ontologies for human behavior recognition. *ACM Computing Surveys (CSUR)*, *46*(4), 1-33.

[83] Rodríguez, N. D., Cuéllar, M. P., Lilius, J., & Calvo-Flores, M. D. (2014). A fuzzy ontology for semantic modelling and recognition of human behaviour. *Knowledge-Based Systems*, *66*, 46-60.

[84] Ryoo, M. S. (2011, November). Human activity prediction: Early recognition of ongoing activities from streaming videos. In *2011 international conference on computer vision* (pp. 1036-1043). IEEE.

[85] Ryoo, M. S., & Aggarwal, J. K. (2006, June). Recognition of composite human activities through context-free grammar based representation. In *2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06)* (Vol. 2, pp. 1709-1718). IEEE.

[86] Sah, M., & Direkoglu, C. (2017, August). Semantic annotation of surveillance videos for abnormal crowd behaviour search and analysis. In *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (pp. 1-6). IEEE..

[87] Sakr, N. A., Abu-Elkheir, M., Atwan, A., & Soliman, H. H. (2018). Current trends in complex human activity recognition. *Journal of Theoretical and Applied Information Technology*, *96*(14), 4564-4583.

[88] Schlenoff, C., Foufou, S., & Balakirsky, S. (2012, October). An approach to ontology-based intention recognition using state representations. In *International Conference on Knowledge Engineering and Ontology Development* (Vol. 2, pp. 178-183). SCITEPRESS.

[89] Schuldt, C., Laptev, I., & Caputo, B. (2004, August). Recognizing human actions: a local SVM approach. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.* (Vol. 3, pp. 32-36). IEEE..

[90] Shi, F., Zhou, F., Liu, H., Chen, L., & Ning, H. (2022). Survey and Tutorial on Hybrid Human-Artificial Intelligence. *Tsinghua Science and Technology*, *28*(3), 486-499.

[91] Simonyan, K., & Zisserman, A. (2014). Two-stream convolutional networks for action recognition in videos. *Advances in neural information processing systems*, *27*.

[92] Sreenu, G. S. D. M. A., & Durai, S. (2019). Intelligent video surveillance: a review through deep learning techniques for crowd analysis. *Journal of Big Data*, *6*(1), 1-27.

[93] Srivastava, Rajeev Kushwaha, Alok. (2014). "Human Activity Recognition Using Object Silhouettes for Automatic Video Surveillance System"

[94] Sun, D., Roth, S., & Black, M. J. (2010, June). Secrets of optical flow estimation and their principles. In *2010 IEEE computer society conference on computer vision and pattern recognition* (pp. 2432-2439). IEEE..

[95] SW,Pienaar R. Malekian Human activity recognition using visual object detection. In2019 IEEE 2nd Wireless Africa Conference (WAC) 2019 Aug 18 (pp. 1-5). IEEE. https://ieeexplore.ieee.org/xpl/conhome/8834888/proceeding

[96] Swears, E., Hoogs, A., Ji, Q., & Boyer, K. (2014). Complex activity recognition using granger constrained dbn (gcdbn) in sports and surveillance video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 788-795).

[97] Taha, A., Zayed, H. H., Khalifa, M. E., & El-Horbaty, E. S. M. (2015, May). Human activity recognition for surveillance applications. In *Proceedings of the 7th International Conference on Information Technology* (pp. 577-586).

[98] Tang, Y., Wang, Z., Li, P., Lu, J., Yang, M., & Zhou, J. (2018, October). Mining semantics-preserving attention for group activity recognition. In *Proceedings of the 26th ACM international conference on Multimedia* (pp. 1283-1291).

[99] Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 4489-4497).

[100] Ullah, M., Ullah, H., & Alseadonn, I. M. (2017). Human action recognition in videos using stable features.

[101] Varol, G., Laptev, I., & Schmid, C. (2017). Long-term temporal convolutions for action recognition. *IEEE transactions on pattern analysis and machine intelligence*, *40*(6), 1510-1517.

[102] Vrigkas, M., Nikou, C., & Kakadiaris, I. A. (2015). A review of human activity recognition methods. *Frontiers in Robotics and AI*, *2*, 28.

[103] Wang, A., Zhao, S., Zheng, C., Yang, J., Chen, G., & Chang, C. Y. (2020). Activities of daily living recognition with binary environment sensors using deep learning: A comparative study. *IEEE Sensors Journal*, *21*(4), 5423-5433.

[104] Wang, S. B., Quattoni, A., Morency, L. P., Demirdjian, D., & Darrell, T. (2006, June). Hidden conditional random fields for gesture recognition. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)* (Vol. 2, pp. 1521-1527). IEEE.

[105] Wanyan, Y., Yang, X., Dong, W., & Xu, C. (2024). A Comprehensive Review of Few-shot Action Recognition. *arXiv preprint arXiv:2407.14744*.

[106] Wei, L., & Shah, S. K. (2017, February). Human Activity Recognition using Deep Neural Network with Contextual Information. In *VISIGRAPP (5: VISAPP)* (pp. 34-43).

[107] Wongpatikaseree, K., Ikeda, M., Buranarach, M., Supnithi, T., Lim, A. O., & Tan, Y. (2012, November). Activity recognition using context-aware infrastructure ontology in smart home domain. In *2012 Seventh International Conference on Knowledge, Information and Creativity Support Systems* (pp. 50-57). IEEE.

[108] WX,Li N. Vasconcelos Complex activity recognition via attribute dynamics. International Journal of Computer Vision. 2017 Apr;122:334-70.. https://doi.org/10.1007/s11263-016-0918-1

[109] Yamada, N., Sakamoto, K., Kunito, G., Isoda, Y., Yamazaki, K., & Tanaka, S. (2007). Applying ontology and probabilistic model to human activity recognition from surrounding things. *IPSJ Digital Courier*, *3*, 506-517.

[110] Yang, Y., & Shah, M. (2012). Complex events detection using data-driven concepts. In Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part III 12 (pp. 722-735). Springer Berlin Heidelberg.

[111] Zhang, S., Wei, Z., Nie, J., Huang, L., Wang, S., & Li, Z. (2017). A review on human activity recognition using vision-based method. *Journal of healthcare engineering*, *2017*(1), 3090343.

[112] Zhang, X., Cui, S., Zhu, T., Chen, L., Zhou, F., & Ning, H. (2023). CASL: Capturing Activity Semantics through Location Information for enhanced activity recognition. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*.

[113] Zhang, Y., Zhang, Y., Swears, E., Larios, N., Wang, Z., & Ji, Q. (2013). Modeling temporal interactions with interval temporal bayesian networks for complex activity recognition. *IEEE transactions on pattern analysis and machine intelligence*, *35*(10), 2468-2483.

[114] Zhu, Y., Nayak, N. M., & Roy-Chowdhury, A. K. (2012). Context-aware activity recognition and anomaly detection in video. *IEEE Journal of Selected Topics in Signal Processing*, *7*(1), 91-101.