ORIGINAL RESEARCH

Activity recognition in a smart home using local feature weighting and variants of nearest-neighbors classifiers

Labiba Gillani Fahad[1] · Syed Fahad Tahir[2]

Abstract

Recognition of activities, such as preparing meal or watching TV, performed by a smart home resident, can promote the independent living of elderly in a safe and comfortable environment of their own homes, for an extended period of time. Different activities performed at the same location have commonalities resulting in less inter-class variations; while the same activity performed multiple times, or by multiple residents, varies in its execution resulting in high intra-class variations. We propose a Local Feature Weighting approach (LFW) that assigns weights based on both inter-class and intra-class importance of a feature in an activity. Multiple sensors are deployed at different locations in a smart home to gather information. We exploit the obtained information, such as frequency and duration of activation of sensors, and the total sensors in an activity for feature weighting. The weights for the same features vary among activities, since a feature may have more importance for one activity but less for the other. For the classification, we exploit the two variants of K-Nearest Neighbors (KNN): Evidence Theoretic KNN (ETKNN) and Fuzzy KNN (FKNN). The evaluation of the proposed approach on three datasets, from CASAS smart home project, demonstrates its ability in the correct recognition of activities compared to the existing approaches.

Keywords Smart home · Activity recognition · Ambient assisted living · Machine learning · Nearest neighbors · Feature weighting

1 Introduction

Recognition of daily activities executed by a smart home resident supports the remote monitoring of elderly and enable them to live independently in their own homes (Patel and Shah 2019; Cook 2012; Chen et al. 2012). Activity recognition improves the quality of life and maintains the well being of a smart home occupant through the analysis of the performed activities and identification of changes in the daily routine (Fahad et al. 2013; Stikic et al. 2011). One of the important applications of sensor based activity recognition could be the monitoring of patients kept in isolation in anticipation of spread of disease, such as in the recent scenarios of Covid-19 (Kannan et al. 2020). A smart home is equipped with sensors to gather observations about the resident and its context. The obtained data takes the form a stream of sensor activations occurred over a period of time, which is partitioned into multiple segments (sequence of sensor events). Each segment represents an activity instance (Fahad et al. 2014, 2015a). These instances are utilized in training a classifier to recognize the newly detected instances.

The major challenges involved in activity recognition include high intra-class and less inter-class variations, which can decline the performance of a learning classifier. Intra-class variations can occur in the same activity performed by different residents because of their personal preferences and the individual human nature. For example, let us consider the activity of "making tea", which comprises of a sequence of events such as taking a cup from cupboard, poring the hot water from the kettle, taking milk form the fridge and adding sugar from the sugar pot. Some people like tea without milk or sugar, while others may use either milk or sugar or both.

✉ Labiba Gillani Fahad
labiba.fahad@nu.edu.pk

Syed Fahad Tahir
fahad.tahir@mail.au.edu.pk

1 Intelligent Knowledge Mining and Analytics Lab,
Department of Computer Science, National University
of Computer and Emerging Sciences, Islamabad, Pakistan

2 Department of Computer Science, Faculty of Computing &
Artificial Intelligence, Air University, Islamabad, Pakistan

Such intra-class variations result in less discriminative information. In the case of the making tea activity, a high weight should be assigned to the objects such as tea, cup and kettle while milk and sugar should be assigned less weight (Tahir et al. 2019). However, in the case of two different activities, such as "making tea" and "making coffee" less inter-class variations exist. The notable events include the use of cup, hot water, sugar, coffee and tea. In such a scenario, tea and coffee would be the most discriminative features and thus, should have the higher weights among the identified features. Therefore, it is important to weight the features by considering both the intra-class and inter-class variations.

In this paper, we propose a Local Feature Weighting approach (LFW) for recognizing the smart home activities with high intra-class and less inter-class variations. Unlike the existing Key Feature Selection (KFS) approach based on the intra-class variations only (Tahir et al. 2019), LFW exploits both the intra-class and the inter-class feature importance in assigning the weights. Further, we weight the same feature differently in each activity class, in accordance with the feature's contribution in the correct classification of that class. For activity classification, in contrast to an adhoc classification method specifically designed for KFS, we use simpler and computationally less intensive variants of KNN classifiers: FKNN and ETKNN. The proposed approach is evaluated and compared with the existing approaches using three challenging datasets with inter-class and intra-class variations from CASAS smart home project: Kyoto7, Kasteren and Kyoto1. The obtained results of the proposed models show better performance in comparison to the existing activity recognition approaches.

The rest of the paper is organized as follows: Sect. 2 describes the related work on activity recognition. The proposed activity recognition approach is discussed in Sect. 3. The datasets for evaluation and experimental analysis are discussed in Sect. 4. Finally, Sect. 5 draws the conclusions.

2 Related work

We recognize the activities of daily living such as eating, sleeping and mediation executed by a smart home resident, using the data obtained from the ambient sensors (Rashidi et al. 2011; Wang et al. 2019; Mckeever et al. 2010; Chen et al. 2012; Fleury et al. 2010).

Activities in a smart home can be recognized using Hidden Markov Model (HMM) and Conditional Random Fields (CRF). The performance of two probabilistic classification models HMM and CRF are compared for recognition of daily activities performed by a smart home resident (Kasteren et al. 2008). Discriminative features of activity classes are identified using Information Gain (IG), while representation of activity classes with fewer instances is

improved through SMOTE and then activities are classified using Evidence Theoretic K-Nearest Neighbors to achieve better recognition performance (Fahad et al. 2015b). A hierarchical classification approach by Fahad et al. (2014), first groups the activities with similar features into clusters using K-means and then evidence based nearest neighbor classifier is used for recognition of activities within each cluster.

A data driven approach improves the precision and sensitivity in recognition of activities by capturing the representation of sensor-activations before and after the time of prediction (Hamad et al. 2019). In the next step, deep learning models namely, Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM), are used for classification of activities. The classification performance of Dempster Shafer Theory (DST) of evidence and Dynamic Bayesian Network in recognition of daily activities shows that both the models are cable of performing better in uncertain situations (Tolstikov et al. 2011). Incorporation of context temporal information such as begin time and duration of an activity can improve the performance of DST of evidence in activity classification (Mckeever et al. 2010). A hybrid approach exploits the generative and discriminative models to exploit the best of both for activity classification (Fahad and Rajarajan 2015), where probability estimation is obtained using curve fitting, and combined with the direct distance minimization for activity recognition.

Naive Bayes Classifier (NBC) is used to recognize the daily activities by exploiting the information from switch state binary sensors (Tapia et al. 2004). Discriminative features within each activity class are identified using inter-class distance, while activities are recognized using Back Propagation Neural Network (BPNN). Changes in the daily behavior of a smart home resident are identified using Probabilistic Neural Network (PNN) and K-means clustering. PNN is used to assign label to the activity instances, while deviated patterns from routine are monitored through clustering (Fahad et al. 2013). Intra-class and inter-class variations in activities are addressed using the neuro-fuzzy classification technique (Ordonez et al. 2013). The approach proposed by Fahad et al. (2015a), learns the correct and incorrect distances to assign labels to the activity instances and then the confidence score of assigned labels is measured using the sub-clustering within activity classes. Another recognition approach, for activities sharing similar events, uses frequent pattern mining to find activity patterns in a particular location and then activity clusters are formed using DBSCAN clustering method (Hoque and Stankovic 2012). An online activity recognition approach recognizes the label from incomplete sensor stream using headmost sensors only rather than all sensor events, which reduces the average time (Liu et al. 2019). Headmost sensors are considered as the key sensors of an activity. A comparison of different deep learning methods for classification of human activities in a

smart home shows the best performance of LSTM (Liciotti et al. 2019).

Emerging patterns and random forest are used to recognize the activities of daily living, where first emerging pattern are mined to extract the meaningful features and then the recognition model is built using random forest (Malazi and Davari 2018). Frequent pattern mining and latent Dirichlet allocation statistical model is applied to group the similar activity patterns into clusters (Chikhaoui et al. 2012). Frequent pattern mining can be used to group similar patterns into clusters, and then ten activities are categorized into different classes using Hidden Markov Model (HMM) (Rashidi et al. 2011). In order to introduce long range dependencies in the sequential labeling algorithms, an activity recognition approach exploits sequential pattern mining along with hidden semi markov model (Avci and Passerini 2014).

A knowledge driven approach using partially observable Markov decision process exploits the information of task description, while the location of its execution is integrated with the events generated by the deployed sensors in a smart home (Liciotti et al. 2019). Activity recognition approaches based on the domain Knowledge uses ontological modeling and semantic reasoning (Chen et al. 2012). An ontology based reasoning framework is discussed by Matassa and Riboni (2019), which recognizes the normal and anomalous behavior of the elderly. A generic hybrid approach to recognize the composite activities occurring in a sequential or parallel order, such as preparing dinner or dish washing, integrates ontology, temporal knowledge and the inference rules (Okeyo et al. 2014). To overcome the limitation of insufficient data for a knowledge driven approach, the knowledge driven and the data driven approaches are combined, while the domain knowledge is also used to improve the learning (Sukor et al. 2019).

Most of the existing approaches (Liciotti et al. 2019; Tolstikov et al. 2011; Rashidi et al. 2011) focus on improving the recognition in the case of high intra-class variations because of different ways of executing the same activity. Some approaches also exist that consider less inter-class variations, where different activities are performed at the same location (Hoque and Stankovic 2012; Tahir et al. 2019; Avci and Passerini 2014). Hardly any approach focuses on both. Also the traditional feature selection and weighting approaches weight a feature based on its overall performance on all activities. However, in the case of activities a feature useful in the case of one activity becomes irrelevant in the case of another activity. Thus features need to be weighted varyingly for different activities, while taking into account the importance of the same feature in the case of high intra-class as well as low inter-class variations.

3 Activity recognition

Let R binary sensors be deployed at different objects and locations in a smart home. We assume that the activity detection part has already been solved. Let $\mathbf{A} = \{A_n\}_{n=1}^{N}$ be a set of N activity classes, where each A_n has M pre-segmented activity instances/samples, $\mathbf{I}_n = \{I_{mn}\}_{m=1}^{M}$. We propose an activity recognition approach that first extracts the features from I_{mn}. The extracted features are weighted differently for each activity class. Finally, we perform the activity recognition on the selected features by applying the KNN, its two variants: ETKNN and FKNN, and their ensemble. Figure 1 shows the block diagram of the proposed approach.

3.1 Feature extraction

Each I_{mn} is represented by a feature vector $\mathbf{F}_{mn} = \{f_{mn}^{r}\}_{r=1}^{R}$, where R is the number of installed sensors. The number of features is the same as the number of sensors installed in a smart home. Each sensor can remain active in multiple intervals during an activity. We compute the time in milliseconds during a sensor remains active in an activity and represent it as f_{mn}^{r}. Since all the sensors are part of a feature set, the features representing inactive sensors are assigned zeros.

3.2 Local feature weighting (LFW)

We weight the features based on their intra-class and inter-class feature importance; hence, in contrast to an

Fig. 1 Block diagram of the proposed activity recognition approach
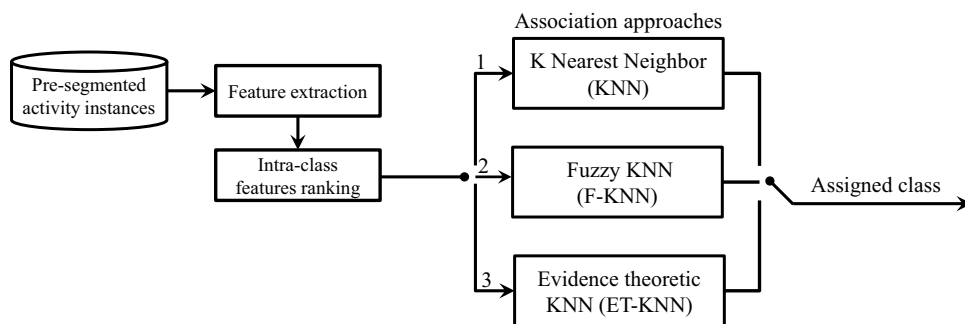
R-dimensional feature vector, we obtain an $R \times N$ dimensional weighted-feature matrix $\mathbf{S}$. Each column in $\mathbf{S}$ represents the features for each activity class A_n, weighted according to the importance. We count the number of instances of each activity A_n where the rth feature f_{mn}^r exists, given as

$$C_n^r = |f_{mn}^r\}_{m=1}^M > 0|, \tag{1}$$

where $| \cdot |$ is the cardinality of a set, and C_n^r counts the number of times f^r appears in all instances of A_n. Thus we can find the intra-class importance of f^r in A_n. Higher the value of C_n^r the more important f^r is for that activity.

In contrast to the fixed weighting based on the number of selected features per activity class (Tahir et al. 2019), we propose a two-level varying weights method to deal with both intra-class and inter-class variations. Let $\hat{\mathbf{C}}$ be an $R \times N$ feature-count matrix obtained by column-wise scaling the values of C_n^r given as

$$\hat{C}_n^r = \frac{C_n^r}{\max_r \hat{C}_n^r}. \tag{2}$$

Each column of $\hat{\mathbf{C}}$ represents the counts of appearances of all feature in an activity A_n, while each row has the counts of appearances of the same feature in N activities. The intra-class feature importance, $\hat{C}_n^r$, is combined with the inter-class feature importance to generate an overall weight W_n^r of the feature f^r for activity A_n given as

$$W_n^r = \hat{C}_n^r + \sum_{n=1}^N \frac{\max_n \hat{C}_n^r - \hat{C}_n^r}{N}, \tag{3}$$

where $\max_n \hat{C}_n^r$ represents the maximum value in $\hat{\mathbf{C}}_r$. Thus a feature with high importance in A_n and less importance in other activity classes is assigned a high weight for A_n. The weights can vary for the same feature in different activity classes. Thus, an $R \times N$ weight matrix $\mathbf{W}$ is obtained.

Finally, the weighted-feature matrix $\mathbf{S}$ for each activity is obtained by element-wise multiplication of each column of $\mathbf{W}$ with the feature vector $\mathbf{F}_{mn}$ given as

$$\mathbf{S} = \mathbf{F}_{mn} \circ \{W_n\}_{n=1}^N, \tag{4}$$

where $\circ$ is the Hadamard product, $\mathbf{F}_{mn}$ is the R dimensional feature vector, and W_n is the R dimensional weight vector of an activity A_n. The obtained feature matrix $\mathbf{S}$ for each activity instance is input to KNN for classification.

3.3 Label assignment using variants of KNN

KNN is a well-known and simple to use method in pattern classification (Lloyd 1982), yet it has good performance. In the case of activity recognition, we have N activity classes.

For every new activity instance $I_{(x)}$, A_n needs to be determined. An unclassified activity instance $I_{(x)}$ is assigned to the activity class represented by the majority of its K-nearest neighbors in the training set known as *voting KNN rule*, where K can be user defined.

ETKNN, a modification of KNN, is based on Basic Belief Assignment (BBA) (Zouhal and Denoeux 1998). The first belief, $Bel(A_n)$, represents the overall probability of a new activity instance $I_{(x)}$ belonging to a particular activity class A_n, given as

$$Bel(A_n) = P(x = n) = \frac{|I_{mn}|}{M * N}. \tag{5}$$

The second belief also known as plausibility, $Pl(A_n)$, is the conditional probability of $I_{(x)}$ belonging to a particular class, given that the class of its neighbors is known. In order to assign the labels the two beliefs, $Bel(A_n)$ and $Pl(A_n)$ are aggregated using the Dempster Shafer Theory (DST) of belief. DST generates a single score s using the orthogonal sum $\oplus$ of the two beliefs (Shafer 1976), given as

$$s = Bel(A_n) \oplus Pl(A_n). \tag{6}$$

FKNN assigns a degree of class membership to an activity instance rather than associating the instance to a particular activity class (Keller et al. 1985), which leads to a membership vector $\mu_{(x)} = \{\mu_{xn}\}_{n=1}^N$ containing N membership values for the new instance $I_{(x)}$, one for each class. The membership value is viewed as a confidence in the assignment of an instance to that activity class. For example, if the new activity instance $I_{(x)}$ is associated with class A_i with membership value $\mu_{xi} = 0.9$ and with another class A_j with $\mu_{xi} = 0.05$, then $I_{(x)}$ would be assigned to A_i with high membership value. Contrary to this, if $I_{(x)}$ is associated with A_i, A_j and A_k with membership values of 0.55, 0.44 and 0.01, respectively, then it can be concluded that $I_{(x)}$ does not belong to A_k, while $I_{(x)}$ has high relevance with both A_i and A_j. This degree of membership can be useful for further analysis in the classification. μ_{xn} of $I_{(x)}$ for an activity class is higher if there are more samples of a particular class in its neighborhood and their distances from $I_{(x)}$ is smaller compared to the other samples in the neighborhood given as

$$\mu_{xn} = \frac{1}{\sum_{j=1}^{\hat{K}} \left\| I_{(x)} - I_{jn} \right\|}, \tag{7}$$

where $\hat{K}$ is the number of samples belonging to the activity A_n in the K nearest neighbors of $I_{(x)}$. The new instance can be assigned to the class with maximum membership value as

$$x = \arg \max_n \mu_{xn}, \tag{8}$$

where x is the ID of the activity class with the highest membership value, assigned to the new activity instance.

4 Evaluation and discussion

In the evaluation we represent the two variants of the proposed approach as $LFW + ETKNN$, and $LFW + FKNN$, where value of K is set equal to the number of classes in each dataset for all variants of KNN. The evaluation is performed using three publicly available smart home datasets; two are from CASAS project (Rashidi et al. 2011) namely: Kyoto1, and Kyoto7, and one is from Kasteren (Kasteren et al. 2008). Leave one day out cross validations is used for the evaluation, where one day data is used for testing and remaining days for training. The evaluation metric comprises of four measures: Precision, Recall, F1score and Accuracy. Moreover, activity level performance is also shown through confusion matrices. The comparison of obtained results with the existing activity recognition approaches (Avci and Passerini 2014; Hoque and Stankovic 2012) shows its superior performance compared to the state-of-the-art approaches.

4.1 Evaluation measures

For the evaluation of the proposed approach and comparison with the existing methods, we use the following four evaluation measures:

$$Precision = \frac{TP}{TP + FP} \times 100, \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \times 100, \tag{10}$$

$$F1score = \frac{2 \times Precision \times Recall}{Precision + Recall}, \tag{11}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100. \tag{12}$$

4.2 Datasets

Table 1 presents the characteristics of datasets from CASAS and Kasteren smart home projects: Kyoto1, Kyoto7 (Rashidi et al. 2011) and (Kasteren et al.2008). Total number of activity instances in Kyoto1, Kyoto7 and Kasteren are 120, 272 and 499, respectively. In Kasteren, single resident performed the 10 activities, while in Kyoto1, 20 participants performed the five activities. In Kyoto7, two occupants living together in a smart home performed the 14 activities. These three datasets have activities with less inter-class and high intra-class variations. Some activities in these datasets have a few instances. We evaluate the performance of the proposed approach on these challenging benchmark datasets.

4.3 Comparison with existing approaches

Table 2 shows the results of using LFW with the variants of KNN compared to using all the features without weighting on the three datasets. In the case of Kyoto1, activity classes are quite separable with less intra-class variations, therefore all the approaches show comparable results in the classification of activity instances .

The Kasteren dataset contains the activities performed at the same location which share similar features, such as 'Use toilet', 'Take shower' or 'Prepare breakfast', 'Prepare dinner', and 'Get drink'. These activities are difficult to discriminate because of less inter-class variations. The significance of LFW can be observed in this dataset. Since LFW learns the fine grained differences between the activities in assigning the weights, the results of both ETKNN and FKNN are improved compared to using all features without weighting. $LFW + FKNN$ achieves the accuracy of 92.03, better than the $LFW + ETKNN$ (91.82), and the rest. The F1scores of $LFW + FKNN$, $LFW + ETKNN$, FKNN, ETKNN and KNN approaches are 0.88 and 0.86, 0.85, 0.83 and 0.81, respectively. These results show that the LFW is able to perform better in the case of less inter-class variations.

In the case of Kyoto7, two residents live together and perform 14 activities without mutual cooperation. The dataset contains similar activities (with different labels),

Table 1 Characteristics of CASAS and Kastern smart homes datasets used in the evaluation

Datasets	Activity classes	Activity instances	Name of activities
Kyoto1	5	120	Clean, Cook, Eat, Phone call, and Wash hands
Kyoto7	14	499	R1 bed to toilet, R1 prepare breakfast, R1 groom, R1 sleep
			R1 work at computer, R1 work at dining room table, R2 bed to toilet
			R2 prepare breakfast, R2 groom, R2 prepare dinner, R2 prepare lunch
			R2 sleep, R2 watch tv, and R2 work at computer
Kasteren	10	272	Leave house, Use toilet, Take shower, Sleep, Prepare breakfast
			Prepare dinner, Get snack, Get drink, Washing machine, and Dish wash

Table 2 Performance evaluation metrics on three smart home datasets for proposed *LFW + ETKNN* and *LFW + FKNN*, and the base classification models: FKNN, ETKNN, and KNN; without feature weighting using leave one day out cross validation. Precision, Recall and Accuracy are in percentages (%), The range of F1score is between [0 ; 1]. The highest values in the performance evaluation metrics are highlighted in bold

Datasets	Folds	Approach	Precision	Recall	F1score	Accuracy
Kyoto1	One day out	*LFW + FKNN*	**98.15**	**97.44**	**0.97**	**97.44**
		LFW + ETKNN	97.12	96.15	0.96	96.85
		FKNN	97.22	96.15	0.96	96.52
		ETKNN	97.01	96.51	0.96	95.92
		KNN	95.71	93.53	0.94	95.61
Kasteren	One day out	*LFW + FKNN*	**88.14**	**89.68**	**0.88**	**92.03**
		LFW + ETKNN	86.03	87.24	0.86	91.82
		FKNN	84.28	86.55	0.85	90.51
		ETKNN	83.62	83.32	0.83	90.93
		KNN	81.18	82.00	0.81	90.23
Kyoto7	One day out	*LFW + FKNN*	**73.26**	**77.38**	**0.75**	**78.16**
		LFW + ETKNN	71.83	76.11	0.73	76.44
		FKNN	70.12	72.74	0.71	76.05
		ETKNN	72.31	73.10	0.72	76.12
		KNN	70.00	71.32	0.70	73.41

Table 3 The accuracy comparison of existing Key feature selection method (Tahir et al. 2019) with the proposed models, *LFW + FKNN* and *LFW + ETKNN*, on Kyoto1 and Kasteren datasets

Datasets	Folds	Approach	Accuracy (%)
Kyoto1	One day out	Key Feature Selection	97.5
		LFW + ETKNN	96.85
		LFW + FKNN	97.44
Kasteren	*One day out	Key Feature Selection	73.33
		LFW + ETKNN	91.82
		LFW + FKNN	92.03

such as'R1 prepare breakfast' and 'R2 prepare breakfast', representing less inter-class as well as high intra-class variations. The inclusion of intra-class feature importance in the weighting process of LFW (3) results in the superior performance of LFW in the case when high intra-class variations also exist along with the less inter-class variations. Both *LFW + FKNN* and *LFW + ETKNN* show better results compared to the approaches without LFW, where *LFW + FKNN* achieves the highest F1score of 0.75 and accuracy of 78.

It can be observed that FKNN shows an overall better performance compared to the other approaches on all three datasets. FKNN counts of the number of neighbors of a particular class as in KNN, moreover, FKNN combines the counts with their distances from the instance of interest to generate a membership score. Thus FKNN becomes less biased towards the classes with more number of instances, which makes it a better choice for the activity recognition problem.

Table 3 shows the comparison of *LFW + ETKNN* and *LFW + FKNN* with the reported results of KFS (Tahir et al.

2019) on Kyoto1 and Kasteren. Since Kyoto1 comprises of well-discriminative activities with high inter-class variations and less intra-class variations, all the three approaches show comparable results with high accuracies of 96.85, 97.44, and 97.5. However, in the case of the more challenging dataset, Kasteren, with less inter-class variations, *LFW + FKNN* and *LFW + ETKNN* outperform KFS (Tahir et al. 2019) with a large margin. The accuracy of *LFW + FKNN* and *LFW + ETKNN* are 92.03 and 91.82, which is much better compared to KFS (73.33). It can be observed that in the case of overlapping activities, the inclusion of inter-class variation in the proposed feature weighting process significantly improves the results.

4.4 Activity level performance analysis

We also analyze the performance of the proposed models for each activity class. For comparisons we apply the same evaluation measures used in the compared methods. Figure 2 shows the activity level accuracy comparison of our best performing model, *LFW + FKNN*, with an existing approach, active learning for overlapping activities (AALO) (Hoque and Stankovic 2012), on Kasteren. It can be noted that *LFW + FKNN* outperforms AALO in the case of most of the activities, while shows comparable results in the two activities: *Prepare dinner* and *Sleep*.

Figure 3 shows F1score comparison of *LFW + FKNN* with another existing approach, activity recognition approach using segmental pattern mining AR-SPM (Avci and Passerini 2014), on Kyoto7. In nine out of the fourteen activities *LFW + FKNN* obtains a higher F1score, while in 'R1 groom' shows the comparable results. In the remaining four activities 'R1 prepare breakfast', 'R1 work at dining room table', 'R2 prepare breakfast' and 'R2

Fig. 2 Activity level accuracy comparison of proposed *LFW + FKNN* with existing activity recognition approach: Active learning in the presence of overlapping activities (AALO) (Hoque and Stankovic 2012), using "leave one day" out cross validation on Kasteren
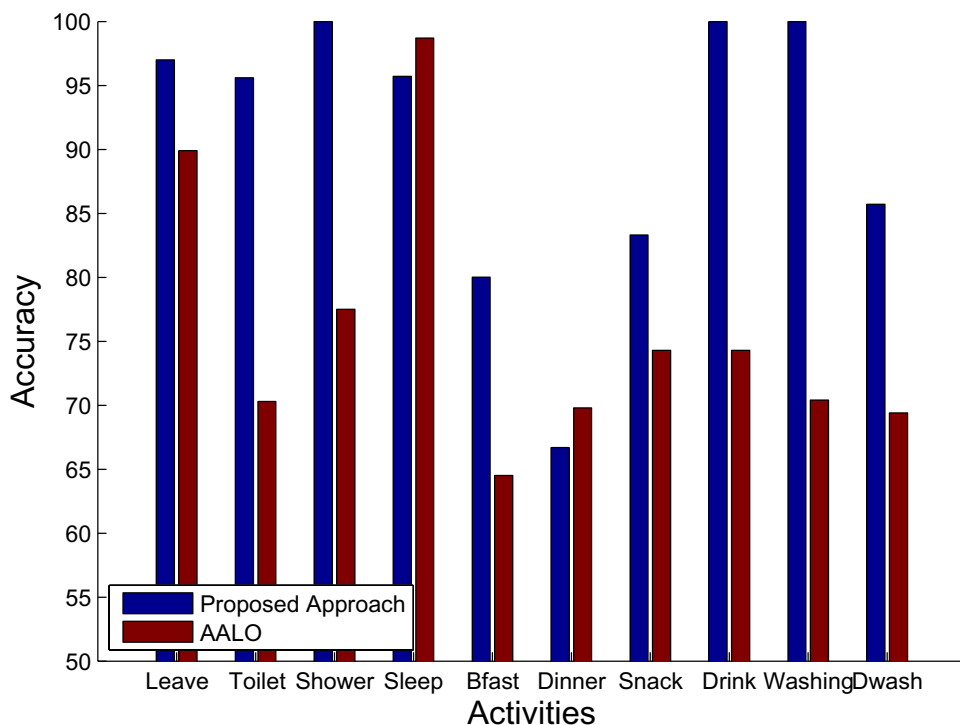
Fig. 3 Activity level performance comparison of proposed *LFW + FKNN* with existing activity recognition approach: Activity recognition using segmental pattern mining (AR-SPM) (Avci and Passerini 2014), through F1score using "leave one day" out cross validation on Kyoto7
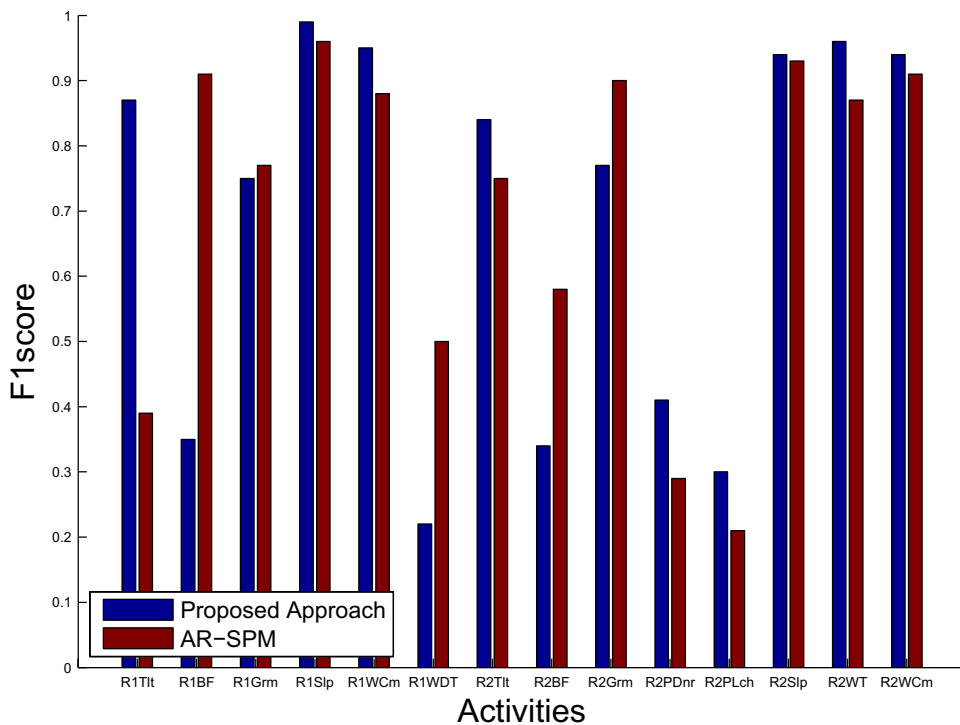
groom', AR-SPM has the better performance. The above comparison shows that the proposed approach correctly recognizes the instances of activity classes in the case of less inter-class and high intra-class variations than the existing approaches.

Table 4 shows the confusion matrix of activities in Kyoto1. In this dataset, the activities are well discriminative with high inter-class and low intra-class variations. The obtained results show that *LFW + FKNN* is able to classify most of the activity instances accurately. The Proposed

Table 4 The accuracy breakdown of the recognized activities by *LFW + FKNN* on Kyoto1. Rows represent the actual activities and columns represent the predicted activities

Activities	Clean	Cook	Eat	Phone call	Wash hands
Clean	89.50	5.30	4.70	0	0
Cook	0	100.00	0	0	0
Eat	4.20	0	92.30	0	3.50
Phone call	0	0	0	100.00	0
Wash hands	0	0	0	0	100.00

approach recognizes the activities of 'Cook', 'Phone Call' and 'Wash hands' with 100% accuracy. Only the 'Eat' and 'Clean' activities share 4% of their instances with each other. It could be due to the reason of overlapping features, since these activities are carried out at the same location such as kitchen.

Table 5 shows the confusion matrix of activities in Kasteren. It can be noted that the the activity instances in almost all the classes are correctly recognized. *LFW + FKNN* obtains more than 90% accuracy in six, and more than 80% in three. It shows comparatively less accuracy of 66% in 'Prepare dinner' (Pdnr) activity, whose 22% of instances are confused with the 'Get snack' (Gsnk) and 11% with the 'Prepare breakfast' (Pbf) activities. The less accuracy in Pdnr is due to its resemblance with Pbf and Gsnk, as these activities are homogeneous and involve the usage of similar features.

Table 6 shows the confusion matrix of activities in Kyoto7. The obtained results show that most of the activities are correctly recognized except activities related to meal preparation such as 'R1 prepare breakfast', 'R2 prepare breakfast', 'R2 prepare dinner' and 'R2 prepare lunch'. Because of the execution of these activities at the same location, kitchen, similar sensors are used such as sensors attached to the water tab, fridge, stove or cupboard. The same case can be observed with the activity of 'R1 work at dinning room table', where 44% of its instances are incorrectly recognized as 'R1 work at computer' and 22% as 'R2 watch TV'. This may be due to the reason that work activity can also be performed at computer table or during the 'watching TV' activity.

5 Conclusion

We proposed a feature weighting approach for improving the activity recognition in smart homes. The features are individually weighted in accordance with their significance in each activity class. The features are weighted differently for different classes. Next, we exploited variants of K-nearest neighbors such as ETKNN and FKNN for activity classification. Experimental evaluation using

Table 5 The accuracy breakdown of the recognized activities by *LFW + FKNN* on Kasteren. Rows represent the actual activities and columns represent the predicted activities

Activities	Leave house	Use toilet	Take shower	Sleep	Prepare b. fast	Prepare dinner	Get snack	Get drink	Wash machine	Dish wash
Leave house	97.00	0	0	0	0	0	0	3.00	0	0
Use toilet	0	95.60	4.40	0	0	0	0	0	0	0
Take shower	0	0	100.00	0	0	0	0	0	0	0
Sleep	0	4.30	0	95.70	0	0	0	0	0	0
Prepare b. fast	0	0	0	0	80.00	5.00	5.00	5.00	5.00	0
Prepare dinner	0	0	0	0	11.10	66.70	22.20	0	0	0
Get snack	0	0	0	0	8.30	8.30	83.30	0	0	0
Get drink	0	0	0	0	0	0	0	100.00	0	0
Wash machine	0	0	0	0	0	0	0	0	100.00	0
Dish wash	0	0	14.30	0	0	0	0	0	0	85.70

Table 6 The accuracy breakdown of the recognized activities by *LFW + FKNN* on Kyoto7. Rows represent the actual activities and columns represent the predicted activities. Key: R1btlt—R1 bed to toilet, R1pbf—R1 prepare breakfast, R1grm—R1 groom, R1slp—R1 sleep, R1wcmp—R1 work at computer, R1wdrt—R1 work at dining room table, R2btlt - R2 bed to toilet, R2pbf—R2 prepare breakfast, R2grm—R2 groom, R2pdnr—R2 prepare dinner, R2plch—R2 prepare lunch, R2slp—R2 sleep, R2wtv—R2 watch tv, R2wcmp—R2 work at computer

Activities	R1btlt	R1pbf	R1grm	R1slp	R1wcmp	R1wdrt	R2btlt	R2pbf	R2grm	R2pdnr	R2plch	R2slp	R2wtv	R2wcmp
R1btlt	94.10	0	5.90	0	0	0	0	0	0	0	0	0	0	0
R1pbf	0	47.10	0	0	5.90	0	0	14.70	0	8.80	23.50	0	0	0
R1grm	22.20	0	73.30	0	0	0	0	0	4.40	0	0	0	0	0
R1slp	0	0	0	100.00	0	0	0	0	0	0	0	0	0	0
R1wcmp	1.70	0	0	0	98.30	0	0	0	0	0	0	0	0	0
R1wdrt	0	11.10	0	0	44.40	22.20	0	0	0	0	0	0	22.20	0
R2btlt	0	0	0	0	0	0	94.90	0	5.1	0	0	0	0	0
R2pbf	0	34.50	0	0	0	0	0	41.40	0	20.70	3.40	0	0	0
R2grm	0	2.50	0	0	0	0	25.00	0	72.50	0	0	0	0	0
R2pdnr	0	27.80	0	0	0	0	0	11.10	0	41.70	19.40	0	0	0
R2plch	0	41.40	0	0	0	0	0	6.90	0	17.20	34.50	0	0	0
R2slp	0	0	0	2.90	2.90	0	0	0	0	0	0	94.30	0	0
R2wtv	0	0	0	0	0	3.20	0	0	0	0	0	0	96.80	0
R2wcmp	0	0	0	0	0	0	4.40	0	0	0	0	0	0	95.60

three smart home datasets from Kastern and CASAS projects demonstrates a superior classification performance of the proposed approach in comparison to the existing approaches. The comparison of nearest neighbors and its variants shows that the best performance is achieved by FKNN and followed by ETKNN with an accuracy of up to 77% in *Kyoto*7, while 93% and 97% in Kasteren and *Kyoto*1 smart home datasets, respectively. Future work includes the recognition of multi-resident activities performed concurrently in a smart home.

References

Avci U, Passerini A (2014) Improving activity recognitionby segmental pattern mining. IEEE Trans Knowl Data Eng 26(4):889–902

Chen L, Nugent C, Wang H (2012) A knowledge-driven approach to activity recognition in smart homes. IEEE Trans Knowl Data Eng 24(6):961–974

Chikhaoui B, Wang S, Pigot H (2012) Adr-splda: activity discovery and recognition by combining sequential patterns and latent dirichlet allocation. Pervasive Mobile Comput 8(6):845–862

Cook DJ (2012) Learning setting-generalized activity models for smart spaces. IEEE Intelligent Syst 27(1):32–38

Fahad LG, Rajarajan M (2015) Integration of discriminative and generative models for activity recognition in smart homes. Appl Soft Comput 37:992–1001

Fahad LG, Ali A, Rajarajan M (2013) Long term analysis of daily activities in smart home. In: Proc. of the European symp. on artificial neural networks, computational intelligence and machine learning, Bruges, Belgium, pp 419–424

Fahad LG, Tahir SF, Rajarajan M (2014) Activity recognition in smart homes using clustering based classification. In: Proc. of IEEE Int. Conf. on pattern recognition, pp 1348–1353

Fahad LG, Khan A, Rajarajan M (2015a) Activity recognition in smart homes with self verification of assignments. Neurocomputing 149(2):1286–1298

Fahad LG, Tahir SF, Rajarajan M (2015b) Feature selection and data balancing for activity recognition in smart homes. In: Proc. of IEEE Int. Conf. on Communications, pp 512–517

Fleury A, Vacher M, Noury N (2010) Svm-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results. IEEE Trans Inform Technol Biomed 14(2):274–283

Hamad RA, Hidalgo AS, Bouguelia MR, Estevez ME, Quero JM (2019) Efficient activity recognition in smart homes using delayed fuzzy temporal windows on binary sensors. IEEE Journal of Biomedical and Health Informatics accepted:2168–2194

Hoque E, Stankovic J (2012) Aalo: Activity recognition in smart homes using active learning in the presence of overlapped activities. In: Proc. of IEEE Int. Conf. on Pervasive Computing Technologies for Healthcare (PervasiveHealth), pp 139–146

Kannan S, Shaik P, Ali S, Sheeza A, Hemalatha K (2020) Covid-19 (novel coronavirus 2019)-recent trends. Euro Rev Med Pharmacol Sci 24(4):2006–2011

Kasteren TV, Noulas A, Englebienne G, Krose B (2008) Accurate activity recognition in a home setting. In: Proc. of Int. conf. on Ubiquitous computing, Seoul, Korea, pp 1–9

Keller J, Gray M, Givens J (1985) A fuzzy k-nearest neighbor algorithm. IEEE Trans Syst Man Cybernet 15(4):580–585

Liciotti D, Bernardini M, Romeo L, Frontoni E (2019) A sequential deep learning application for recognizing human activities in smart homes. Neurocomputing In press

Liu Y, Wang X, Zhai Z, Chen R, Zhang B, Jiang Y (2019) Timely daily activity recognition from headmost sensor events. ISA Transactions In press

Lloyd S (1982) Least squares quantization in pcm. IEEE Trans Information Theory 28(2):129–137

Malazi HT, Davari M (2018) Combining emerging patterns with random forest for complex activity recognition in smart homes. Appl Intelligence 48(2):315–330

Matassa A, Riboni D (2019) Reasoning with smart objects' affordance for personalized behavior monitoring in pervasive information systems. Knowledge and Information Systems accepted, Mar 2019:1–24

Mckeever S, Ye J, Coyle L, Bleakley C, Dobson S (2010) Activity recognition using temporal evidence theory. J Ambient Intell Smart Environ 2(3):253–269

Okeyo G, Chen l, Wang H (2014) Combining ontological and temporal formalisms for composite activity modelling and recognition in smart homes. Future Generation Comput Syst 39:29–43

Ordonez FJ, Iglesias JA, de Toledo P, Ledezma A, Sanchis A (2013) Online activity recognition using evolving classifiers. Expert Syst Appl 40(4):1248–1255

Patel A, Shah J (2019) Sensor-based activity recognition in the context of ambient assisted living systems: a review. J Ambient Intell Smart Environ 11(4):301–322

Rashidi P, Cook DJ, Holder LB, Schmitter-Edgecombe M (2011) Discovering activities to recognize and track in a smart environment. IEEE Trans Knowl Data Eng 23(4):527–539

Shafer G (1976) A mathematical theory of evidence. Princeton University Press, Princeton

Stikic M, Larlus D, Ebert S, Schiele B (2011) Weakly supervised recognition of daily life activities with wearable sensors. IEEE Trans Pattern Anal Mach Intell 33(12):2521–2537

Sukor A, Syafiq A, Zakaria A, Rahim NA, Kamarudin L, Setchi R, Nishizaki H (2019) A hybrid approach of knowledge-driven and data-driven reasoning for activity recognition in smart homes. J Intell Fuzzy Syst 36:4177–4188

Tahir SF, Fahad LG, Kifayat K (2019) Key feature identification for recognition of activities performed by a smart-home resident. J Ambient Intell Human Comput 1–11

Tapia EM, Intille SS, Larson K (2004) Activity recognition in the home using simple and ubiquitous sensors. Pervasive Comput 3001:158–175

Tolstikov A, Hong X, Biswas J, Nugent C, Chen L, Parente G (2011) Comparison of fusion methods based on dst and dbn in human activity recognition. J Control Theory Appl 9(1):18–27

Wang J, Chen Y, Hao S, Peng X, Hu L (2019) Deep learning for sensor-based activity recognition: a survey. Pattern Recognit Lett 119:3–11

Zouhal LM, Denoeux T (1998) An evidence-theoretic k-nn rule with parameter optimization. IEEE Trans Systems Man Cybernet Part C 28(2):263–271