The International Journal of Engineering and Science (IJES)
|| Volume || 10 || Issue || 7 || Series I || Pages || PP 57-63 || 2021 ||
ISSN (e): 2319-1813 ISSN (p): 20-24-1805

SVM-based Learning Model for Abnormal Behaviors in Smart Home

Zhi-Wei Chou, Ying-Kai Lu, Ke-Nung Huang*

Department of Electronic Engineering, I-Shou University, Kaohsiung, Taiwan

** Corresponding Author: Ke-Nung Huang*

ABSTRACT

The ability to recognize abnormal behaviour is important for health monitoring systems in smart home. Taiwan will have at least 20 percent of the population over 65 by 2025. Smart home is one of solution used to assist elderly live independently. We designed a SVM-based learning model to classify activities of daily living and to analyze an individual's daily routines and habits, typically for the elderly who live alone. The temporal information and spatial sequences collected over time are used to generate pattern, which can be fitted to the training data and the fitted model can be used to make a prediction. One of the CASAS smart home datasets was used to train and to retest the algorithm. A non-trained dataset was also used to validate the accuracy of the algorithm. Abnormal behaviors can be detected by compared with individual's daily activity pattern as baseline. Through detecting and distinguishing the abnormal behaviors of elder's daily activities, it can provide assistances on elder's independent living and improvement of aged quality of life.

KEYWORDS;- *activity daily life (ADL); telemonitoring; abnormal behavior; smart home*

Date of Submission: 26-06-2021 Date of Acceptance: 08-07-2021

I. INTRODUCTION

There is great change for elderly to live alone for most of the time, especially during pandemic. In order to provide a safer environment for elderly, abnormal behaviors observed during activities of daily living are a good indicator that the person is more likely to have health and behavioral problems that need intervention and assistance [1]. Almost every country in the world is experiencing a growing and aging population. Societies in which those aged 65 years and older account for 7%, 14%, and 20% are referred to internationally as aging societies, aged societies, and super-aged societies respectively, as illustrated in Fig. 1. Taiwan became an aging society in 1993, became an aged society in 2018, and is projected to become a super-aged society in 2025 according to the National Development Council. Population aging is not only happening in Taiwan, but also the phenomenon for most of the country in the world. In 2020, there are approximately 4.5 people of working age to support one elderly person [2].

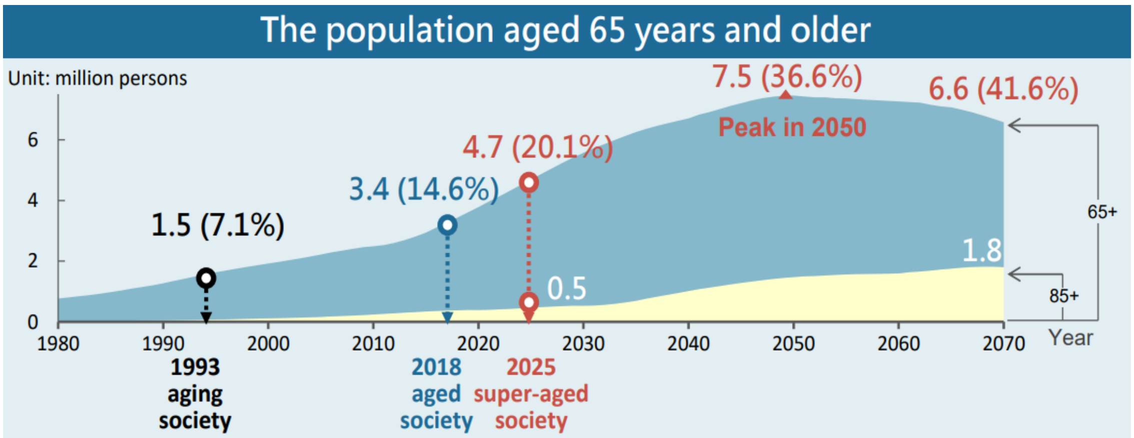

Fig. 1. Speed of population aging in Taiwan. [2]

Population aging raises numbers of public issues. One important issue is an increasing demand for medical services. With universal health coverage, the government have to raise taxes to finance the extra public medical expense. Smart home is one of the solutions to meet the diverse needs of the elderly and building an aged society that is inclusive and self-dependent [2]. It offers the ability to track people's activities without interfering in their daily life; and to track behaviors and monitor their health by using sensors embedded in their living spaces [3-4]. Tracking user behavior for abnormality detection is becoming a primary goal for some researchers [5]. Abnormal behavior detection approaches are based mainly on machine learning algorithms and specifically on supervised learning techniques [6]. Hung et al. [7] proposed a homecare sensory system that combines Support Vector Machine (SVM) and Hidden Markov model (HMM). They used RFID-based sensor networks to collect the elder's daily activities to estimate whether the elder's behavior is abnormal or not. In the work of Chua, Foo and Guesgen [8], they show the important of spatial and temporal information in interpreting human activity and how such information can be represented for activity recognition. The prediction by partial matching method is extended to capture spatio-temporal information by exploiting the repetitions from activity events.

In this study, CASAS (Center of Advance System in Adaptive System) dataset was used. CASAS collected real data through a smart home environment located on the Washington state University Pullman campus. Sensors embedded in the smart home generate readings while residents perform their daily routines. The sensor readings including date, time, sensor number, sensor status and activity are collected by a computer network and stored in a database that an intelligent agent uses to generate useful knowledge such as patterns, predictions, and trends. The purpose of this study is to design a SVM-based learning model for activity recognition, especially for elderly, in Smart Home.

II. MATERIAL AND METHOD

SVMs (Support Vector Machines) are a useful technique for data classification method introduced in 1992 by Boser, Guyon, and Vapnik [9]. Given a training set of instance-label pairs $(x_i, y_i), i = 1,2, \dots, l)$ where $x_i \in R^n$ and $y \in \{1, -1\}^l$, the SVM require the solution of the following optimization problem:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i$$
$$\text{subject to } y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i \qquad (1)$$
$$\xi_i \geq 0$$

the function$\phi(x_i)$ maps the vectors x_i into a higher dimensional space. Where $\xi_i$are slack variables that allow an example to be in the margin (also called a margin tolerance). C > 0 is the penalty parameter of the error term. The C parameter sets the relative importance of maximizing the margin and minimizing the amount of slack. For greater values of C, a smaller margin will be accepted if the decision function is better at classifying all training points correctly, the less support vectors, the closer to the concept of hard-margin SVM, but easy to overfitting. A smaller C will encourage a larger margin, therefore a simpler decision function, at the cost of training accuracy, the more support vectors, the larger margin can be tolerated. In other word, C works as a regularization parameter in the SVM. This formulation is called the soft-margin SVM.

Furthermore, $\mathrm{K}(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is called the kernel function. The widely used kernel is the Radial Basis function (RBF) kernel defined by

$$\mathrm{K}(x_i, x_j) = \exp(-\gamma \parallel x_i - x_j \parallel^2), \gamma > 0 \qquad (2)$$

A small γ means a Gaussian with a large variance so the influence of x_j is more, i.e. if $x_j$is a support vector, a small γ implies the class of this support vector will have influence on deciding the class of the vector x_i even if the distance between them is large. If γ is large, then variance is small implying the support vector does not have wide-spread influence. Proper parameter setting in the kernels increases SVM classification accuracy. In this study, RBF kernel and Soft margin are the core concept of the algorithm written with C language and executed in Linux system. There are two parameters to be determined in the SVM model with the RBF kernel: C and γ . The γ parameter automatically defines the distance which a single training example can reach, with low values meaning 'far' and high values meaning 'close'. Experiments are carried out to evaluate SVM performance through variations of the γ and C parameters.

There are several data sets that can be used to evaluate activity recognition, even synthetic ones [10]. CASAS (Center of Advance System in Adaptive System) dataset was used in this study [11]. The CASAS data sets are the most popular ones, as they not only provide data from smart homes with single occupancy but also provide data from homes with multiple residents. We used the Aruba data set to evaluate our approach in the context of smart homes with one resident. The sensor events are generated from motion sensors (sensor IDs begin with "M"), door closure sensors (sensor IDs begin with "D"). Fig. 2 shows an example of activity events in a smart home. We separated the Aruba dataset into training and testing sets. Each instance in the training set contains one target value (i.e. the class labels). Sensors are often deployed in the smart home to collect information about the inhabitant's daily activities.

Date	Time	Sensor	State	Activity labeled
2010-11-04	11:41:32.570925	M022	OFF	
2010-11-04	11:41:34.029848	D004	OPEN	Leave_Home begin
2010-11-04	11:41:37.192624	M030	ON	
2010-11-04	11:41:43.345957	D004	CLOSE	Leave_Home end
2010-11-04	11:41:44.121	M030	OFF	
2010-11-04	11:43:30.094537	D004	OPEN	Enter_Home begin
2010-11-04	11:43:30.658939	M030	ON	
2010-11-04	11:43:34.541657	M030	OFF	
2010-11-04	11:43:34.683398	D004	CLOSE	Enter_Home end
2010-11-04	11:43:35.454279	M022	ON	

Fig. 2. CASAS Aruba dataset

III. RESULT AND DISCUSSION

In this study, the SVM-based algorithm was modified from the open source algorithm LIBSVM developed by Professor C. J. Lin [12]. LIBSVM, a library for Support Vector Machines (SVMs), has gained wide popularity in machine learning and many other areas. The Aruba data set contains 11 types of activities. We pre-processed the Aruba data set by computing the activity space for each type of activity recorded in the data set. The SVM-based learning model was trained with 325732 data to identity different daily activities which were classified in to 11 categories including Bed_to_Toilet, Eating, Enter_home, Housekeeping, leave_Home, Meal_Preparation, Relax, Respirate, Sleeping, Wash_Dishes and Work. And then, activity tags were removed from dataset (325732data) to test the trained SVM algorithm. In this study, γ was set to 0.01 and C was set to 10 to perform the best classification of the algorithm. Accuracy was around 90% compared with original dataset (with activity tags), as illustrated in Fig. 3. In addition, a non-trained dataset (119216 data) were used to test the accuracy of the algorithm. Fig. 4 indicates the 92% of accuracy with the non-trained dataset. The result confirmed the practicability of our SVM-RBF algorithm.

Activities	Number	Number / All data (%)
Sleeping_Data_Number	34365	10.55%
Bed_to_Toilet_Data_Number	553	0.17%
Meal_Preparation_Data_Number	122565	37.63%
Relax_Data_Number	147047	45.14%
Housekeeping_Data_Number	4172	1.28%
Eating_Data_Number	8021	2.46%
Wash_Dishes_Data_Number	143	0.04%
Leave_Home_Data_Number	1229	0.38%
Enter_Home_Data_Number	613	0.19%
Work_Data_Number	6894	2.12%
Respirate_Data_Number	140	0.04%
All data	325742	

Accuracy 89.33%

Parameter C :10.00

Parameter GAMMA :0.0100

Fig. 3. γ=0.01 and C =10, accuracy of original dataset.

Activities	Number	Number / All data (%)
Sleeping_Data_Number	8650	7.26%
Bed_to_Toilet_Data_Number	103	0.09%
Meal_Preparation_Data_Number	43503	36.49%
Relax_Data_Number	60300	50.58%
Housekeeping_Data_Number	1061	0.89%
Eating_Data_Number	2139	1.79%
Wash_Dishes_Data_Number	66	0.06%
Leave_Home_Data_Number	307	0.26%
Enter_Home_Data_Number	183	0.15%
Work_Data_Number	2759	2.31%
Respirate_Data_Number	145	0.12%
All data	119216	

Accuracy 92.30%
Parameter C :10.00
Parameter GAMMA :0.0100

Fig. 4. Accuracy with the non-trained dataset.

Fig. 5 shows percentage of time spend and number of times motion sensors trigged for each activity in one day from the training dataset (Aruba 2010/11/10). Among 11 daily activities, sleeping (48%) and relax (41%) were two activities accounted for large portion of time whereas leaving home and washing dishes were only accounted small percentage of time in one day. The time frame to perform each activity is different. For example, the sleeping activity was longer to perform whereas wash_dish activity was shorter. Base on the infrared sensor as a motion detector, the number of times motion sensors triggered for each activity were counted. The more the individual moved the more data were collected.

Sleeping activity represented 48% of time in one day, but only 18% of motion data were accounted for sleeping activity. The percentage of time spend on Meal_preparation was only 5%, but 33% of motion data were accounted for this activity. There was a lot of movement involved in Meal_Preparation activity such as food preparation or carry a plate, therefore, more motion data were collected for this activity. On the other contrary, sleeping activity was often at resting mode, therefore, motion sensors were not triggered as often. This information can help to monitor behavior changes over times. If percentage of time on sleeping activity is reduced but the number of motion data is increased, it could be interpreted that the individual has difficulty sleeping or poor sleeping quality due to insomniac or physical condition. Recognising the abnormal behaviour is a challenging task, since they can be complex, irregular, and very substantially between instances. Therefore, we need simultaneous analysis of the time spend and number of sensors trigged.

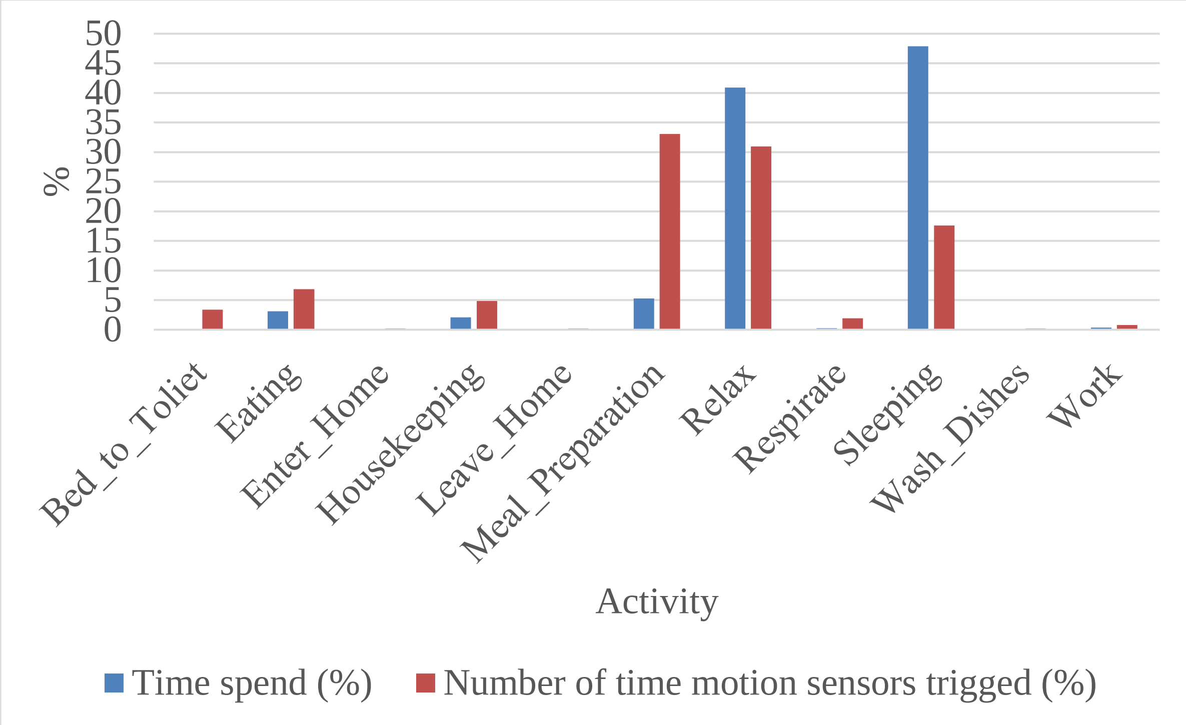

Fig. 5. Time spend (%) and number of times motion sensors trigged (%) for each activity in one day from the training dataset.

We also analyzed the same individual's daily activity for one week from March 1, 2011 (Tuesday) to March 7, 2011 (Monday). Time pattern between 7 days was similar which could interpret the daily habit of the individual was regular and routine (Table I). It can be divided in to two parts (Monday to Saturday vs Sunday). Compared to Sunday, Monday to Saturday had higher percentage on work, higher percentage on sleeping and less percentage on leave_home. On Sunday, individual spend less time on sleeping (32%) but had higher percentage on leave_ home (4%). The reason for less percentage of sleeping was very likely that individual had to get up early in order to go out. Table II indicated the number of times motion sensors trigged for each activity for the same week. The average percentage of time spend on each activity for one week (2011/3/1 - 2011/3/7) in Fig. 6 is similar to the percentage of each activity on 2010/11/10 in Fig. 5. The Daily habit of the resident did not change much after 4 months from 2010/11/10. The characteristic of daily habit can be analyzed. It also indicates the feasibility of our algorithm.

Table 1 Percentage of time for each activity for one week (%)

Date / Activity	3/1 Tue	3/2 Wed	3/3 Thu	3/4 Fri	3/5 Sat	3/6 Sun	3/7 Mon
Bed_to_Toliet	0.00	0.08	0.34	0.19	1.78	0.81	0.71
Eating	0.72	1.76	1.34	2.51	0.77	0.9	5.64
Enter_Home	0.26	0.02	0.06	0.04	0.19	0.12	0.02
Housekeeping	2.44	4.61	2.01	3.9	2.42	3.64	2.77
Leave_Home	0.13	0.07	0.13	0.23	0.41	4.12	0.39
Meal_Preparation	16.57	12.65	5.73	5.26	7.51	8.28	6.78
Relax	41.99	36.23	46.5	39.91	45.37	49.56	40.43
Respirate	0.06	0.19	0.01	0.04	0.05	0.00	0.01
Sleeping	35.26	41.35	37.55	43.14	36.23	32.03	41.5
Wash_Dishes	0.91	0.59	0.99	0.73	0.79	0.20	0.86
Work	1.65	2.45	5.34	4.11	4.49	0.33	0.89

Table 2 The number of times motion sensors trigged for each activity for the same week (%)

Date / Activity	3/1 Tue	3/2 Wed	3/3 Thu	3/4 Fri	3/5 Sat	3/6 Sun	3/7 Mon
Bed_to_Toliet	0.03	0.01	0.26	0.11	0.89	0.65	0.68
Eating	1.30	2.04	2.77	1.72	1.53	1.6	5.75
Enter_Home	0.27	0.14	0.39	0.33	0.80	0.52	0.23
Housekeeping	7.57	10.13	12.05	7.54	11.72	18.06	12.81
Leave_Home	0.50	0.6	0.75	0.40	1.60	0.13	0.45
Meal_Preparation	55.97	41.68	29.67	32.20	37.36	36.89	34.87
Relax	25.37	32.91	44.77	45.34	28.31	35.4	33.23
Respirate	0.14	0.54	0.08	0.18	0.29	0.03	0.09
Sleeping	3.40	2.48	4.21	3.84	6.89	5.25	8.66
Wash_Dishes	1.12	1.19	2.08	1.48	1.71	0.48	0.51
Work	4.33	8.29	2.97	6.85	8.89	0.99	2.72

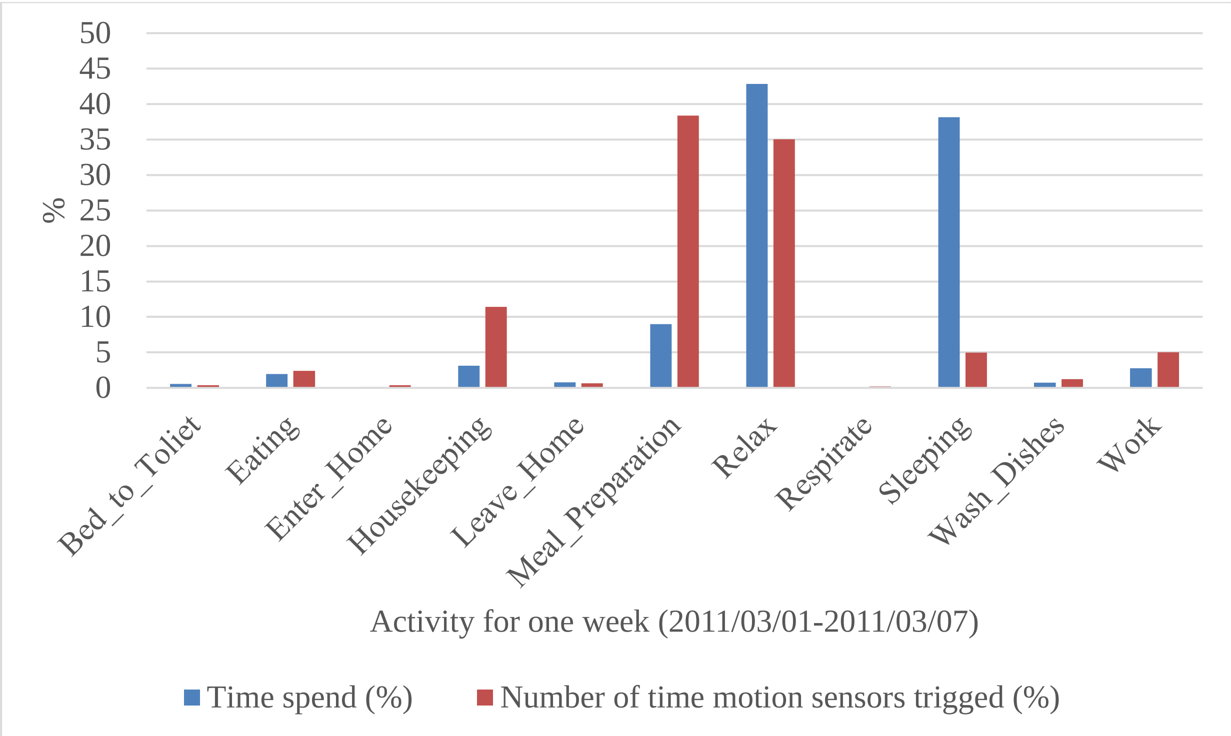

Fig.6. Time spend (%) and number of times motion sensors trigged (%) for each activity for one week.

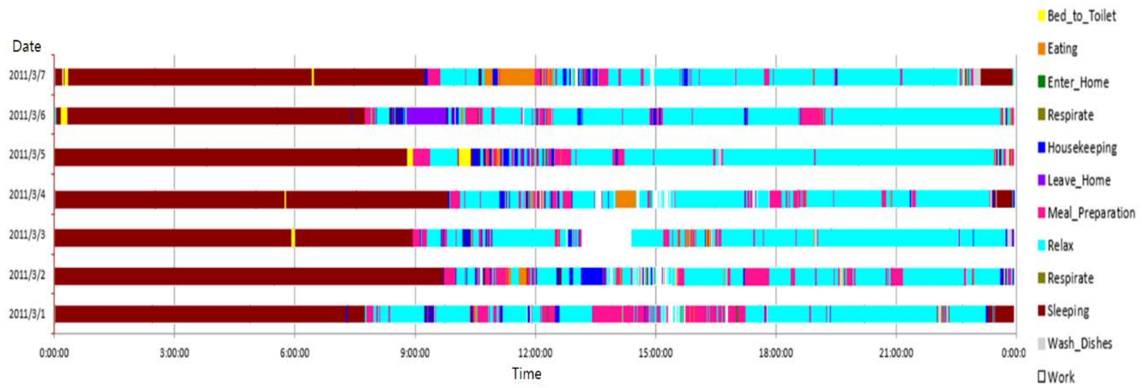

Fig7. Distribution of daily activities for everyday.

Everyone's daily habit is different. Some people are used to go to bed early and some people are used to bed late. Fig 7. Shown when each activity occurred for every day. For example, we can understand this individual wake up around 9, spend most of the time at home, and have the habit going to toilet during sleep. Data collected for a long period of time can provide a baseline for the analysis of abnormal behaviors.

IV. CONCLUSIONS

Based on the analysis from the SVM-based learning model, an individual's habit of daily living activities can be drawn. It can be used to monitor daily activity and to detect a significant change of daily habits which could indicate a problem. We could also detect abnormal behaviors by compared with the individual's daily activity pattern as a baseline. These analyses provide important insights for healthcare provider. As a next step, we will focus on immediately anormal behaviors detection and link the system to a medical alarm system. Healthcare providers will be notified as soon as an anomaly is detected.

REFERENCES

[1]. Zerkouk, M.; Chikhaoui, B. Spatio-Temporal Abnormal Behavior Prediction in Elderly Persons Using Deep Learning Models. *Sensors* **2020**, *20*, 2359. https://doi.org/10.3390/s20082359

[2]. Population Projections for the Republic of China (Taiwan): 2020-2070, Department of Human Resources Development, National Development Council (Taiwan), August 2020.

[3]. Lê, Q.; Nguyen, H.B.; Barnett, T. Smart Homes for Older People: Positive Aging in a Digital World. Futur. Internet 2012, 4, 607–617.

[4]. A. Fleury, M. Vacher and N. Noury, "SVM-Based Multimodal Classification of Activities of Daily Living in Health Smart Homes: Sensors, Algorithms, and First Experimental Results," in *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 274-283, March 2010, https://doi.org/10.1109/TITB.2009.2037317.

[5]. Yahaya, S.W.; Langensiepen, C.; Lotfi, A. Advances in Computational Intelligence Systems; Springer Science and Business Media LLC: Berlin, Germany, 2019; Volume 840, pp. 362–371.

[6]. Mukhopadhyay, S.C. Next Generation Sensors and Systems; Springer Nature: Cham, Switzerland, 2016; Volume 16, pp. 1–330.

[7]. Hung, Y.-X.; Chiang, C.-Y.; Hsu, S.J.; Chan, C.-T. Abnormality Detection for Improving Elder's Daily Life Independent. In Aging Friendly Technology for Health and Independence; Springer Nature: Cham, Switzerland, 2010; Volume 6159, pp. 186–194.

[8]. Chua, S.-L.; Foo, L.K.; Guesgen, H.W. Predicting Activities of Daily Living with Spatio-Temporal Information. *Future Internet* **2020**, *12*, 214. https://doi.org/10.3390/fi12120214

[9]. Boser, B.E., Guyon, I.M., and Vapnik, V.N. (1992) A training algorithm for optimal margin classifiers. In D. Haussler, editor, 5th Annual ACM Workshop on COLT, pp. 144–152, Pittsburgh, PA. ACM Press.

[10]. Synnott, J.; Nugent, C.; Jeffers, P. Simulation of Smart Home Activity Datasets. Sensors 2015, 15, 14162–14179.

[11]. Cook, D. Learning Setting-Generalized Activity Models for Smart Spaces. Intell. Syst. 2012, 27, 32–38.

[12]. C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1-27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Ke-Nung Huang*. "SVM-based Learning Model for Abnormal Behaviors in Smart Home." *The International Journal of Engineering and Science (IJES)*, 10(07), (2021): pp. 57-63.