International Journal of Innovative Computing, Information and Control
Volume **17**, Number **3**, June **2021**
ICIC International ©2021 ISSN 1349-4198
pp. 905–918

INTELLIGENT MONITORING FOR ELDER CARE USING VISION-BASED TECHNOLOGY

Pann Thinzar Seint, Thi Thi Zin and Pyke Tin

Graduate School of Engineering
University of Miyazaki
1 Chome-1 Gakuenkibanadainishi, Miyazaki 889-2192, Japan
{ pannthinzarseint; pyketin11 }@gmail.com; thithi@cc.miyazaki-u.ac.jp

Received January 2021; revised April 2021

Abstract. *Nowadays, smart home care systems are being developed in response to various demands, though challenges remain in realizing various required functionalities. Among many considerations used in developing the proposed system, this paper focuses on ways of recording the consumption of medicine and food by elderly people living alone, as well as ways of communicating information to caregivers. Primarily, we used color coding for objects to facilitate their identification and use. Firstly, we propose useful features, not only between the skin surfaces of hands and mouth, but also the contact between body parts and the objects involved. An Eigen value detector is used to overcome the skin occlusion problem. And then, action detection is performed (such as for picking up or grasping medicine, taking medicine, eating, drinking water, and using a towel) by using a combination of the proposed feature and conditional rule-based learning. Secondly, the proposed system uses context awareness for assessing the subject's actions using statistical analysis. Finally, the entire system is implemented through the user interface of the application platform. Using this system, caregivers can easily see a record of daily activities, provided with contextual information useful in improving the quality of care. Our proposed system is easy to learn and can provide an economical labor-saving solution for caregivers.*
Keywords: Consumption record, Context awareness, Elderly people, Medicine and food intake, Smart home care, Vision-based technology

1. **Introduction.** Lessons learned from behavior analysis of the elderly can be intelligently applied, contributing to longer and healthier lives for our aging population. By the year 2050, the world population of people aged 60 or older is projected to grow 250% compared to figures for 2013 [1]. The aging process sometimes involves a gradual loss of mental and physical abilities. This contributes to increased risk of accidents in older people's homes. In providing long-term care, this places greater demand on formal health care services, as well as on informal caregivers, such as family, friends, and neighbors. However, as the population ages, a trend toward increasing numbers of people with disabilities places an additional burden on caregivers. Another trend toward fewer caregivers motivates researchers to develop smart home-care monitoring systems to overcome these problems. In recent years, an awareness has emerged of requirements for developing an efficient approach to monitor the eating and drinking habits of elderly people. In addition, the elderly does not always take medicine as prescribed. Addressing the challenges of aging has become the focus of extensive scientific research during the last decade, though gaps in such research must be addressed in the near future [2].

DOI: 10.24507/ijicic.17.03.905

Recognizing and analyzing behavior is a continuing challenge, due to the breadth of the domain in the real world, and the wide variety of possible behaviors. Depending on the purpose, monitoring elderly behavior is sometimes quite straightforward. Providing home care monitoring for the elderly is more helpful than monitoring the daily activities of normal people. It is often a matter of observing whether they are correctly adhering to proper routines for taking meals and medications. For this purpose, we proposed intelligent monitoring system for elderly care. By using image processing method, we extract the features to know the interaction between the detected skin parts and identified color objects. Depending on the extracted features, action detection process is performed by using the proposed conditional rules. After analyzing the action sequences, the proposed system will inform the user history such as the quantity of food intake and medication consumed, and whether which medicine is skipped, date, time taken.

2. **Related Works.** In developing a smart home-care environment, a vision-based monitoring system for elder care is proposed in [3], in which ALMOND (Assisted Living MONitoring Dataset) was used as the experimental data. This dataset includes actions such as standing, lying, vomiting, falling, eating, drinking, and reading. For disabled person, the author in [4] proposed an intelligent electric wheelchair which can be used in manual and electric mode. The ultrasonic sensor is used to avoid the obstacles in front of user. In [5], artificial intelligence is used to assess nutrient intake for hospitalized patients using a method of food segmentation. A method of automatically monitoring meal intake is introduced to analyze eating behavior in [6]. In this work, Microsoft Kinect is used to analyze skeletal motion by tracking joint positions, and the Hidden Markov Model is used to classify gestures. In [7], Chidananda and Reddy used hand movements in a rules-based classification method of recognizing behavior associated with eating and drinking, though they did not discuss object information. The method of detecting eating and drinking behavior described in [8] addresses the challenge of dealing with the great variety of individual eating styles. In [9], the proposed system involves an analysis of food intake actions, using the distance between the joints of one or both of hands. The detection of eating behavior is made when the joints move toward the head. An Android-based application is developed by Ameta et al. in [10] as a medical reminder system. In this system, alarms can be set for various kinds of medicine. Using a vision-based technique in [11], medication intake monitoring is proposed, using body part detection, face and hand differentiation, and bottle detection based on color. This technique incorporates color-based object tracking, Hu image moments, and edges. Medication intake recognition is performed by Petri network, and the proposed techniques achieved an accuracy of 75%.

In [12], a method of identifying medicine bottles is presented featuring a combination of video camera surveillance and radio frequency identification (RFID), for detecting bottles as moving objects. A method of monitoring medication intake through a state transition system is introduced in [13], and its contribution lies in differentiating medication intake from other behavior. Our previously proposed system described in [14] analyzed the sub-actions involved in medication intake, such as normal action, grasping the medicine (bottle 1, bottle 2, and bottle 3), grasping a water cup, taking medicine, and drinking water by using the color object information described in [11]. However, state-of-the-art techniques still lack parameter learning for interpreting activity details. No statistical analysis was performed on the duration of each sub-action. In [15], recognizing medication intake in real time, using a three-level hierarchical approach is proposed. The levels in this hierarchy consist of the duration of each action, the time interval between each action, and the point at which medication intake occurs. This system includes a differentiation of normal and abnormal behavior for medication intake. The referenced research includes numerous ways of recognizing the behavior involved in medication and meal intake.

Our previously proposed system in [16] introduces a new approach for monitoring medication and meal intake. This system provides information derived from an observation history of each subject being monitored by knowing the sequences of each action. However, this system incorrectly identified some medication actions by using the pattern recognition neural network (PRNN) classifier. It also had some conflict for "touching towel to mouth" and "eating actions" because of rule weaknesses and skin detection problems. Our currently proposed monitoring system reduces the errors of our previous work, and also focuses on details that are much more useful for attaining a reliable healthcare management solution. This paper presents modifications of our previous system, along with detailed explanations. These modifications and fixes include the use of extracted feature parameters, a new classification system for medication sub-actions, and meal intake model. They also include modified conditional rule-based learning, as well as the addition of a newly defined action for towel use. Moreover, the steps for learning classified action sequences are covered in great detail. After learning the sequences of sub-actions, the main activity is identified, with a status notification sent to caregivers. The proposed work is implemented in the user interface of the application, which can also be combined with information and communication technology (ICT) to make a smart environment in the future.

3. **Materials and Methods.** The experimental videos are collected using an RGB camera which can capture ultra HD video images at 30 frames per second. The proposed system has been developed for home health-care monitoring, especially for recognizing behavior associated with medicine and nutrition intake. The system has two environmental constraints: only one person can live in the room, and the person being monitored is instructed not to wear clothing of the same color as objects in the room. The proposed system began monitoring after obtaining skin images for three body parts (the two hands, as well as face). The system has three main functions: (i) detection, labelling and tracking of body parts and color-coded objects, (ii) feature extraction, and then (iii) classification of simple actions and identifying activities using context. An overview of the proposed system is shown in Figure 1.

3.1. **Detection, labelling and tracking of body parts and color-coded object.** Color-coded objects are used to indicate the kind of interaction to be performed. Figure 2(a) and Figure 2(b) show the experimental environment, as well as the models used for medication and meal intake. The medication intake model is provided with red, blue, and green medicine bottles, along with a yellow cup for water. For the nutrition or meal intake model, a white plate and a yellow cup are used. Figure 2(c) illustrates the color used for each item. Detecting these color-coded objects is performed in the RGB color space. Threshold values for the color-coded objects are chosen by observations of our video data. As seen in Figure 3, skin color detection is performed to obtain skin regions, including mouth and hand regions. Skin regions are detected using the YC_bC_r color space [17,18].

After detecting body parts and objects, the labeling process is performed. As shown in Figure 3(a), face and hand regions are individually extracted. Firstly, the face region is differentiated from the other body parts using measurements for major and minor axes, as well as location parameters. After obtaining the face region, the mouth region M is differentiated using the geometric properties of the extracted face. The other two body parts are then labeled as right hand H_1 and left hand H_2, respectively. In addition,

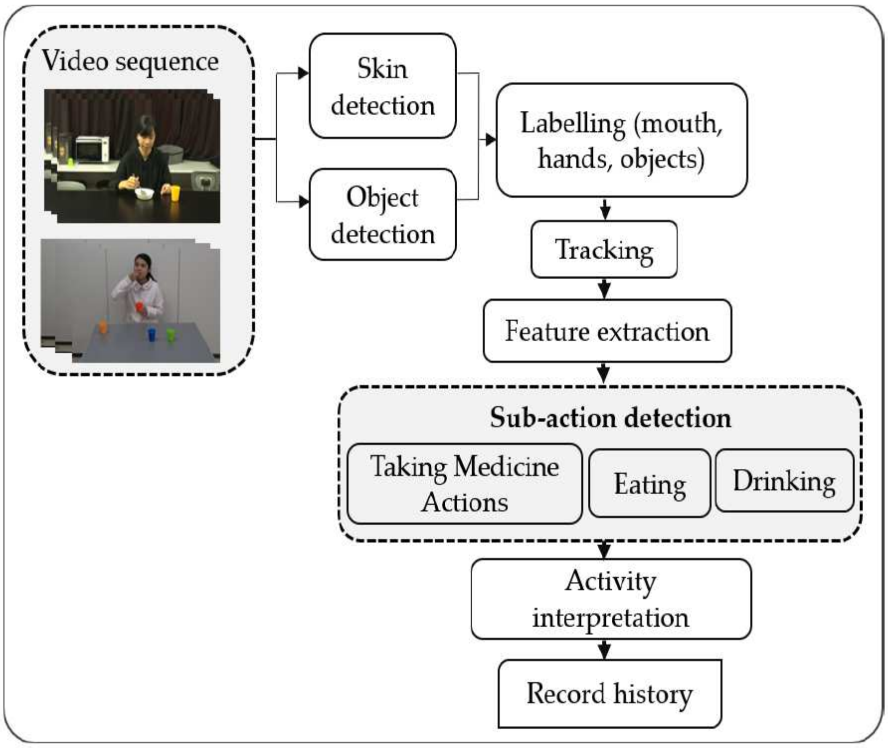


Figure 1. Overview of the proposed system

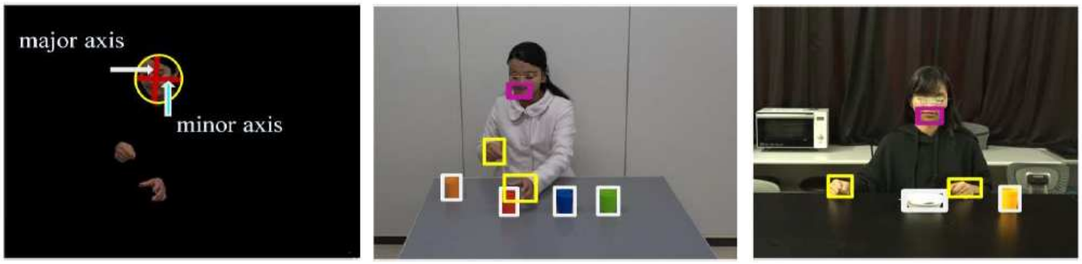
(a) (b) (c)

Figure 2. (color online) (a) Medication intake model, (b) meal intake model, and (c) color coding for labeling objects

objects that are color-coded using red, blue, green, yellow, and white are labeled as medicine bottle M_1, M_2, M_3, water cup W, and food plate F. Bounding boxes (BB) for all regions involved in the medicine and meal intake process are shown in Figure 3(b) and Figure 3(c).

The three body parts are sometimes occluded in some actions, such as touching two skins together, taking medicine, eating and drinking water. When the person is taking medicine or drinking water, the face and hand regions are combined to become one object. In these conditions, they are sometimes obtained as a single body part, or possibly two. Examples of such combinations in specific actions are shown in Figure 4(a) and Figure 4(b). We have developed some tracking considerations for overcoming the occlusion problem. Details of the tracking process are shown in Figure 5.

The first step involves checking the number of detected body parts, whether it equals TH or not (where $TH = 3$). When all three body parts are detected, individual regions are labeled, and then Eigen values are computed to trace each region separately. By



(a) (b) (c)

FIGURE 3. (color online) (a) Skin extraction, (b) labeling for medication intake model, and (c) labeling for meal intake model

(a) (b)

FIGURE 4. Label combination for (a) taking medicine, and (b) drinking water

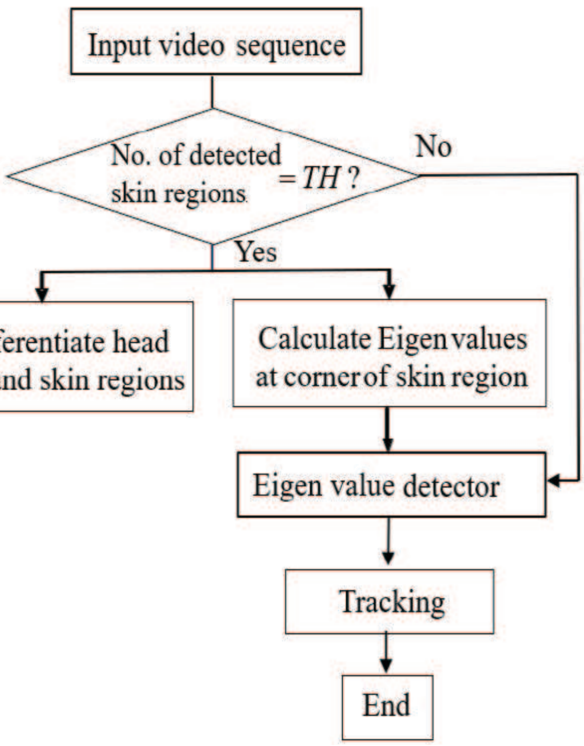
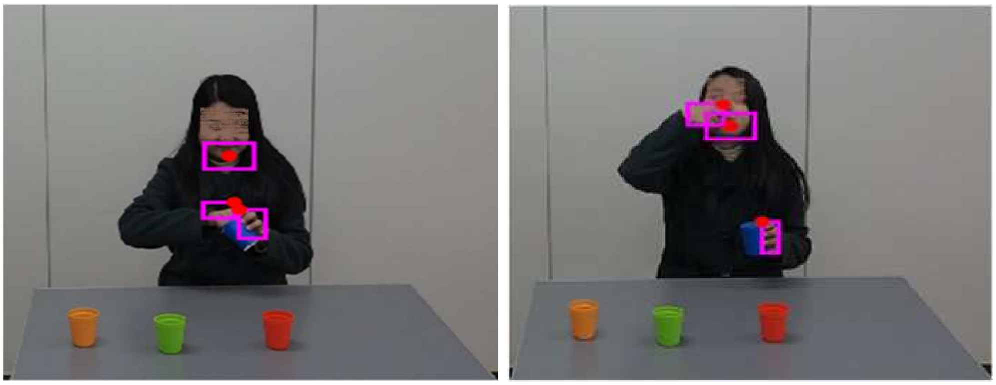


FIGURE 5. Flowchart for tracking process



(a) Hand over-hand occlusion (b) Hand over-mouth occlusion

Figure 6. Tracking with Eigen detector

using minimum eigenvalue algorithm developed by Shi and Tomasi [19], corner's interest points are detected on the input image and these points are recorded. To perform tracking process, the similarity point pairs are estimated by using geometric transformation which map the previous and current images. Otherwise, when the body parts are occluded or combined, the previous Eigen values are used to find the similar points on the current image as before. In this way, the occlusion problem can be solved using the proposed tracking method [13], and the results for handling skin occlusion are described in Figure 6(a) and Figure 6(b). The next section describes feature extraction for detecting the actions of human-object interaction.

3.2. **Feature extraction.** Feature extraction is a principal component of computer vision, and also a key component of the observation process for obtaining outputs for classification. Therefore, distinct features are provided to learn the proposed two models. The proposed feature extraction methods are explained in the following subsections.

3.2.1. *Feature extraction for medication intake actions.* For the medication intake model, 11 features are extracted to establish the interaction between four color-coded objects and body parts. By using these features, seven sub-actions are detected, as described in the next section 3.3.1 for action classification.

The interaction feature described in Equation (1) is the ratio of the intersecting area between the two regions defined by bounding boxes (BB) to the union area between them, i.e., no overlap value exists if no action has occurred.

$$OverlapRatio = \frac{BB_A \cap BB_B}{BB_A \cup BB_B} \tag{1}$$

where BB_A and BB_B are the bounding boxes for regions A and B. In this medication intake system, the 11 interaction features are extracted in the following:

H_iM_j where $i = 1, 2$ (right hand, left hand), $j = 1, 2, 3$ (medicine 1, 2 and 3)
H_iM where $i = 1, 2$ (right hand, left hand), M (mouth)
H_iW where $i = 1, 2$ (right hand, left hand), W (water cup)
WM where W is water cup and M is mouth

3.2.2. *Feature extraction for meal intake actions.* For the meal intake model, we propose pixel distance as the most useful feature. The three sub-actions are defined using pixel distance information upon the reference point r. It can be calculated by the following equation:

$$d_{ij} = \sqrt{(r_i(x) - r_j(x))^2 + (r_i(y) - r_j(y))^2} \tag{2}$$

where d_{ij} is the distance between reference points r_i and r_j; $r_i(x, y)$ and $r_j(x, y)$ are the x and y coordinates of reference points r_i and r_j.

Here, in this system, five reference points are used: r_i $(i = 1, 2, 3, 4, 5)$.

r_1 = centroid of water of water cup

r_2 = centroid of food plate

r_3 = centroid of mouth

r_4 = upper left corner of right hand

r_5 = upper right corner of left hand

By using these reference points as illustrated in Figure 7, the five distance features between two points (r_3r_4, r_3r_5, r_2r_4, r_2r_5 and r_1r_3) are extracted for meal intake actions.

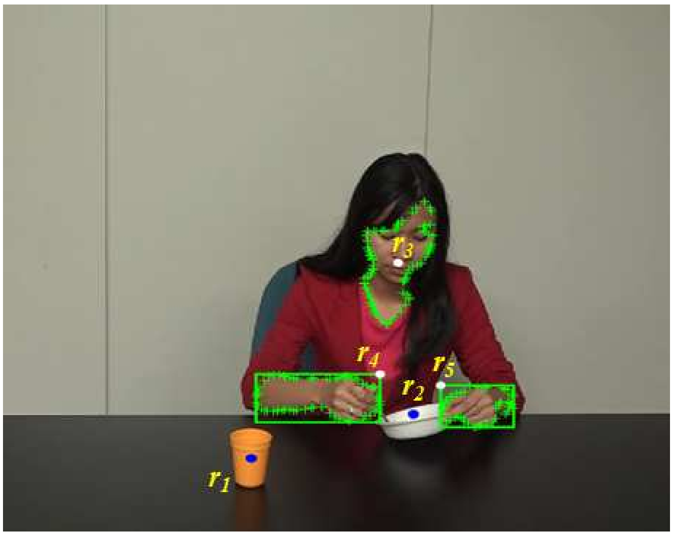


FIGURE 7. Five reference points for meal intake action

3.3. **Classifying sub-actions and interpreting activities using context.** In this section, we propose an action classification method using the two models: (1) medication intake model, and (2) meal intake model. The first model has seven action classifications, and the second model has three. The proposed actions and their related classes are described in Table 1, and subsequent subsections present details of these action classifications.

TABLE 1. Proposed actions for the two models

Class	Medication intake action	Meal intake action
1	Non-action	Preparing to eat
2	Grasp medicine 1	Eating
3	Grasp medicine 2	Drinking
4	Grasp medicine 3	
5	Grasp water cup	
6	Take medicine	
7	Drink water	

3.3.1. *Activity interpretation for medication intake model.* In this model, we extract 11 features in classifying seven actions. Depending on the active feature values, we can classify the seven actions using the following target output classes: '1', '2', '3', '4', '5', '6' and '7' by proposing conditional rule-based learning. The extracted feature values are checked to meet the step-by-step conditional rules. The proposed rules for medication actions are as follows:

Transition: All feature values are zeros

Drinking: $(H_1M > 0$ and $H_1W > 0)$ or $(H_2M > 0$ and $H_2W > 0)$ or $(H_1W > 0$ and $WM > 0)$ or $(H_2W > 0$ and $WM > 0)$ or $WM > 0$

Take medicine: $H_1M > 0$ or $H_2M > 0$

Grasp medicine 1: $H_1M_1 > 0$ or $H_2M_1 > 0$

Grasp medicine 2: $H_1M_2 > 0$ or $H_2M_2 > 0$

Grasp medicine 3: $H_1M_3 > 0$ or $H_2M_3 > 0$

Grasp water cup: $H_1W > 0$ or $H_2W > 0$

The seven actions obtained from the proposed method are shown in Figure 8. The extracted feature patterns and their related actions can be seen in Table 2.

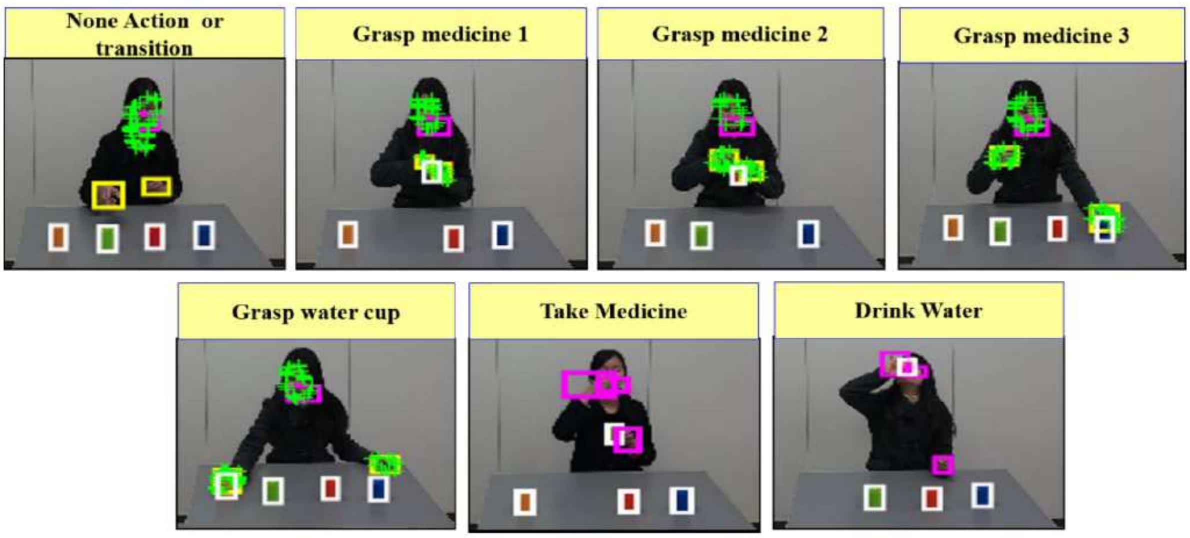


FIGURE 8. Seven medication intake action classifications obtained using the proposed method

After observing the action sequences in the input video, frames containing Action Class '1' none-actions are removed, and remaining actions are then compressed to obtain shorter action sequences, as described in Figure 9.

Using the occurrence threshold value, some actions are discarded as unrelated to any identified class. The compressed action sequences are fed into the support vector machine (SVM) classifier for providing context, which contains '9' statements, such as whether the medication intake event is completed or not. The various possible action patterns or compressed action sequences are trained in a multi-class SVM [20]. The complete workflow for activity interpretation can be seen in Figure 10. A user history can be seen in the application GUI, as shown in Figure 11. In this GUI figure, the caregiver can see the action sequences as a user history resulted from the video. '363575262575464575' is the

TABLE 2. Feature patterns and their related medication intake actions

Features for medication intake model (overlap area ratios)											Actions	Action class
H_1_M_1	H_1_M_2	H_1_M_3	H_1_W	H_2_M_1	H_2_M_2	H_2_M_3	H_2_W	H_1_M	H_2_M	W_M		
0	0	0	0	0	0	0	0	0	0	0	None action	1
0.09	0	0	0	0.13	0	0	0	0	0	0	Grasp medicine 1	2
0	0	0	0	0.24	0	0	0	0	0	0	Grasp medicine 1	2
0	0.16	0	0	0	0	0	0	0	0	0	Grasp medicine 2	3
0	0.18	0	0	0	0.11	0	0	0	0	0	Grasp medicine 2	3
0	0	0.12	0	0	0	0	0	0	0	0	Grasp medicine 3	4
0	0	0	0.06	0	0	0	0	0.05	0	0.34	Grasp water cup	5
0	0	0	0	0.12	0	0	0	0.07	0	0	Take medicine	6
0	0	0.15	0	0	0	0	0	0	0.06	0	Take medicine	6
0	0	0	0.08	0	0	0	0	0.05	0	0.32	Drink water	7
0	0	0	0	0.07	0	0	0.03	0	0	0.35	Drink water	7

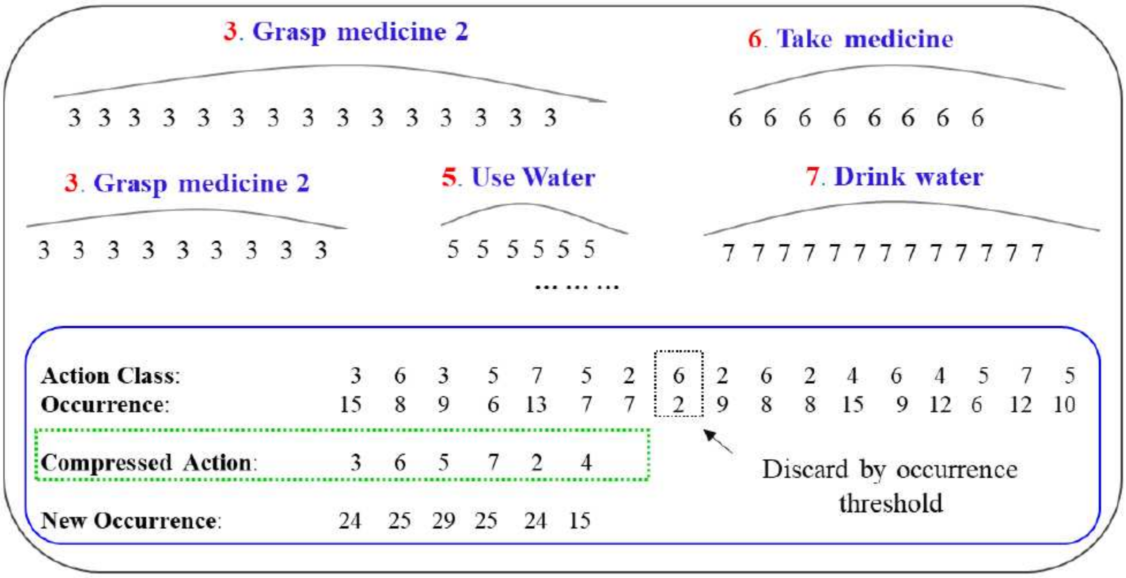


FIGURE 9. Medication intake action sequences after compression

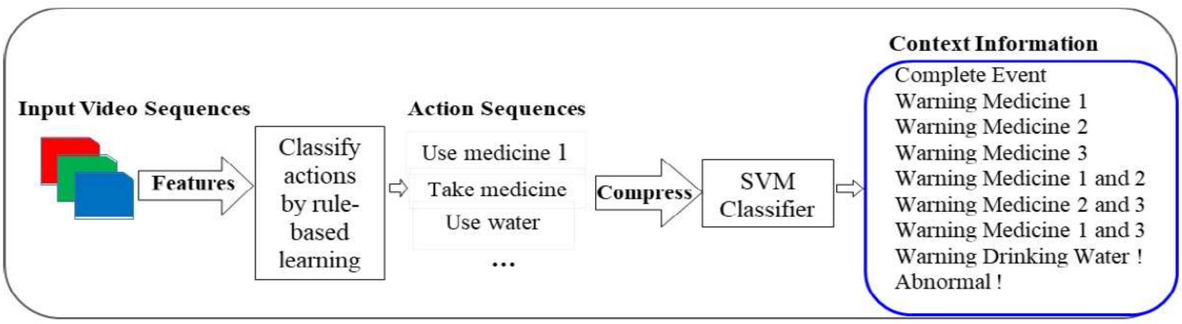


FIGURE 10. The workflow for activity interpretation

compressed action sequence. After applying SVM classifier on this compressed sequence, we finally obtained the context information such as 'Medication Intake is completed' together with date and time. This status can contribute to knowing the caregiver for the assessment and management of taking medicine.

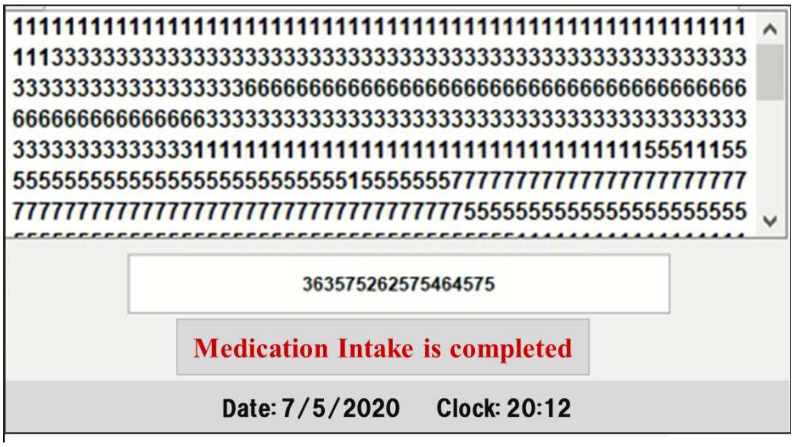


Figure 11. Medication intake history obtained from user interface

3.3.2. *Activity interpretation for meal intake model.* To detect the meal intake actions, we introduce three actions using pixel distance features (r_3r_4, r_3r_5, r_2r_4, r_2r_5 and r_1r_3) as presented in Subsection 3.2.2.

The distance threshold values are empirically chosen using three video sequences. The proposed step-by-step rule-based learning for action classification is as follows:

$$\begin{aligned} \text{Drinking:} \quad & r_1r_3 < 80 \\ \text{Eating:} \quad & (r_3r_4 < 70 \text{ and } r_1r_3 > 250) \text{ or } (r_3r_5 < 70 \text{ and } r_1r_3 > 250) \\ \text{Preparing to eat:} \quad & r_2r_4 < 110 \text{ or } r_2r_5 < 110 \end{aligned}$$

We firstly check to confirm whether conditions for the drinking rule have been satisfied. If not, we confirm whether those for eating or preparing to eat have been met. We also consider transitions, such as preparing for an action. The relationships between the extracted features, classified actions, and classes '1', '2', and '3' are shown in Table 3. The meal intake actions obtained from rule-based learning can be seen in Figure 12. When the person performs an action such as eating or drinking, the pixel distance between the interaction objects shortens. In this table, we see that an eating action is detected by the positions of the hands in approaching the mouth. Likewise, drinking is also detected by movements of the hand and the water cup in approaching the mouth. In the activity interpretation process, the UI output providing a consumption record for eating and drinking is described in Figure 13. This UI displays the quantity of food and water intake, as well as date and time for user consumption. By seeing this intake record, the caregiver can easily notice the requirement for the daily intake.

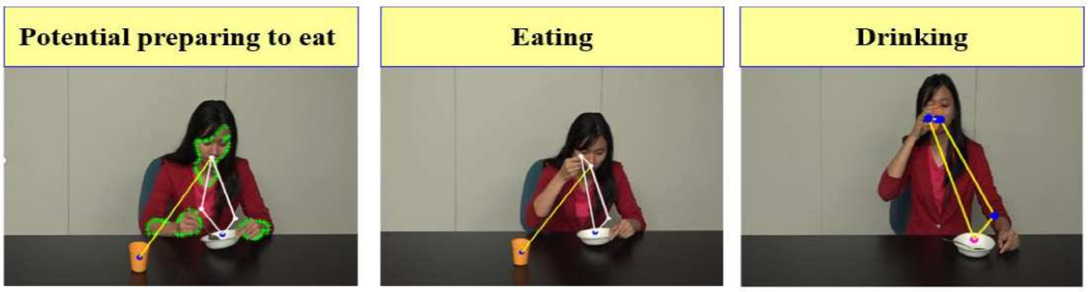


Figure 12. Three meal intake action classes using the proposed method

TABLE 3. Feature patterns and their related meal intake actions

Features for meal intake model (pixel distances)					Actions	Action class
r_3r_4	r_3r_5	r_2r_4	r_2r_5	r_1r_3		
167.8	177.8	103.4	91.5	314.9	Preparing to eat	1
190.2	188.0	89.6	91.9	323.8	Preparing to eat	1
205.1	201.7	89.1	93.4	334.7	Preparing to eat	1
211.1	202.5	85.0	91.1	335.3	Preparing to eat	1
186.6	199.6	104.3	94.4	331.8	Preparing to eat	1
184.7	190.5	103.3	92.7	326.1	Preparing to eat	1
192.5	184.9	91.5	94.4	320.6	Preparing to eat	1
60.0	167.9	225.7	90.5	306.8	Eating	2
65.0	159.5	266.3	92.1	300.2	Eating	2
73.8	158.7	280.5	92.3	300.3	Eating	2
77.5	160.9	288.8	92.2	302.7	Eating	2
78.7	162.6	292.2	91.9	304.1	Eating	2
79.3	164.1	294.4	91.6	305.3	Eating	2
66.8	166.3	275.1	91.6	306.9	Eating	2
260.6	62.2	153.5	334.7	35.1	Drinking	3
261.5	77.8	153.9	354.4	27.8	Drinking	3
264.2	87.6	152.8	368.9	36.3	Drinking	3
268.9	95.0	156.4	384.1	43.4	Drinking	3
272.9	103.6	156.6	401.8	51.4	Drinking	3

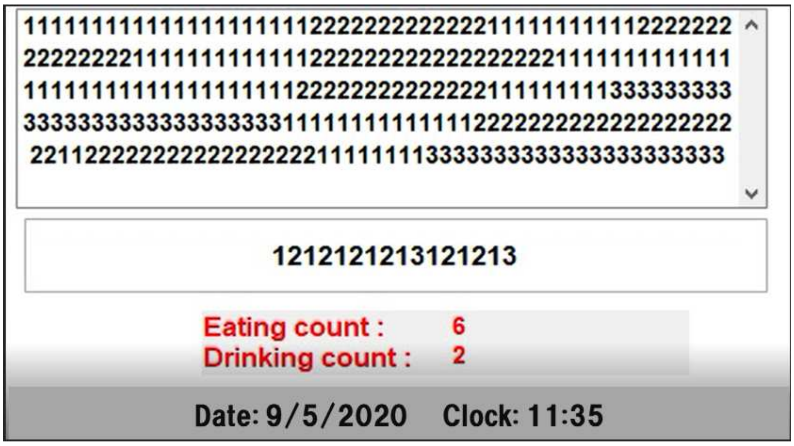


FIGURE 13. Meal intake history obtained from user interface

We have finally adapted the proposed method to detect towel use, as described in Figure 14. Although this action is not as important in health-care monitoring of the elderly, we find that recognizing towel use is useful in correcting errors in interpreting eating actions, as the hand-to-mouth gestures are similar. On the other hand, this proposed interaction will be useful in recognizing other activities of daily living (ADLs).

4. **Experimental Results.** This presents a performance analysis for the results of using our proposed methods in classifying actions. The experimental works are performed using Windows 10, on an Intel ® Core ™ i7-7700 CPU, @ 3.6 GHz with 8GB memory. Table 4 presents accuracy results for medication intake videos, and Table 5 shows the results

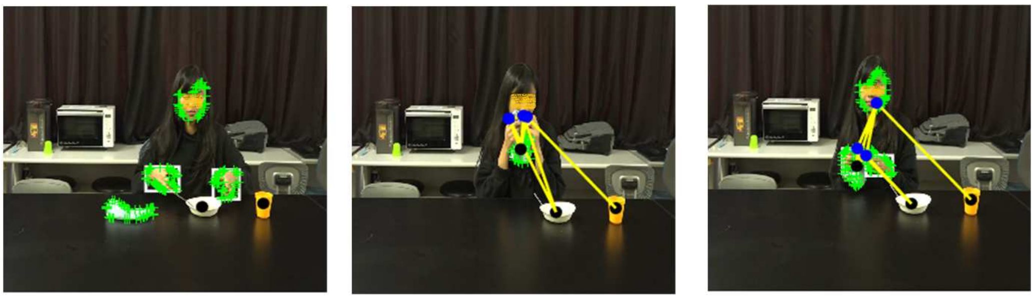


Figure 14. Towel-use action

Table 4. Action classification for medication intake videos

Medication intake videos	**Number of frames processed by 15 fps**	**Number of frames for correct action classification**	**Accuracy of action classification**	
			Proposed method	Previous method
Video 1	571	563	98.6%	95.8%
Video 2	616	606	98.4%	86.7%
Video 3	571	536	93%	75%
Video 4	606	601	99%	98.3%
Video 5	675	672	99.5%	98.2%

Table 5. Action classification for meal intake videos

Meal intake videos	**Number of frames processed by 15 fps**	**Number of frames for correct action classification**	**Accuracy of action classification**	
			Proposed method	Previous method
Video 1	828	819	98.9%	98.9%
Video 2	331	325	98.1%	94.3%
Video 3	334	329	99.6%	97.3%
Video 4	591	586	99.1%	93.6%
Video 5	419	416	99.2%	98.8%

for meal intake videos. The proposed method achieved better accuracy than the previous method in [16] and it has over 97% for both medication intake action and meal intake action classification.

5. **Discussion.** Our proposed monitoring system is of particular interest in determining healthcare status. For long-term elder care, our system focuses on ensuring quality medical and nutritional care, rather than on less important, and sometimes more complicated tasks. Continuity of observation is key to providing quality long-term care. This paper includes extensive revisions and modifications of our previous work in order to make the application more user-friendly. We have simplified the proposed feature extraction method, obtaining promising results for action classification. When tested using five video sequences for each model, our proposed step-by-step conditional rule-based method attained more accurate classification results than with the method in the previous work. However, much more development will be needed to implement in the real-world environment of elderly. As another innovation, our current method includes a new action classification for "using towel". Including this class in action interpretation reduces errors that had resulted from incorrectly classifying those actions as eating, and also presents opportunities for adding additional action classes in the future. When the subject was wearing a short-sleeve shirt, some classification errors occurred related to grasping medicine. The

initial sub-action involved in grasping one medicine was sometimes mistaken for grasping another medicine. Another classification error resulted from an indefinite extraction of body parts. However, as these errors occurred in some number of frames, we could reduce the obtained error count over multiple frames by using occurrence threshold. This improved our assessment of overall system performance. After acquiring an accurate classification of action sequences, the entire work is evaluated by providing the caregiver with useful information along with the history for an observed subject.

6. **Conclusions.** In this paper, we proposed an intelligent monitoring system, especially for elderly healthcare that can provide the contextual information to the caregiver. The system is implemented in a user-friendly interface to show the user history. Such history includes dates and times, as well as the quantity of food consumed, and whether a dose of medication is skipped. The proposed informative system provides reliable messages, and will prove to be a helpful, indispensable companion for caregivers. In the future, Internet of Things (IoT) applications will be generated based on GUI implementations, which can provide offline learning. However, many challenges remain due to environmental variation, such as complex backgrounds, skin detection issues due to race-based variations in skin color, and color detection problems resulting from lighting effects. Therefore, the need is anticipated for additional applications and modifications of this research, such as the use of a pre-trained network for object recognition instead of using color information and more robust skin detection to avoid the limitations. Thus, efforts to attain real-time monitoring in real-world applications will continue.

Acknowledgment. We sincerely appreciate to our lab members for participating experimental videos with us.

REFERENCES

[1] L. Garçon et al., Medical and assistive health technology: Meeting the needs of aging populations, *The Gerontologist*, vol.56, no.Suppl_2, pp.S293-S302, 2016.

[2] S. Koch, Healthy aging supported by technology – A cross-disciplinary research challenge, *Informatics for Health and Social Care*, vol.35, nos.3-4, pp.81-91, 2010.

[3] M. Buzzelli, A. Albé and G. Ciocca, A vision-based system for monitoring elderly people at home, *Applied Sciences*, vol.10, no.1, p.374, 2020.

[4] W. Budiharto, Intelligent controller of electric wheelchair from manual wheelchair for disabled person using 3-axis joystick, *ICIC Express Letters, Part B: Applications*, vol.12, no.2, pp.201-206, 2021.

[5] Y. Lu et al., An artificial intelligence-based system for nutrient intake assessment of hospitalised patients, *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp.5696-5699, 2019.

[6] L. Costa, P. Trigueiros and A. Cunha, Automatic meal intake monitoring using hidden Markov models, *Procedia Computer Science*, vol.100, pp.110-117, 2016.

[7] H. Chidananda and T. H. Reddy, Human eating/drinking activity recognition using hand movements to monitor and assist elderly people, *IOSR Journal of Computer Engineering (IOSR-JCE)*, vol.19, pp.57-63, 2017.

[8] O. Amft, Automatic dietary monitoring using on-body sensors: Detection of eating and drinking behaviour in healthy individuals, *[Zurich]: ETH Zurich*, http://e-collection.library.ethz.ch/eserv/eth:30651/eth-30651-02.pdf, Switzerland, 2008.

[9] E. Gambi, M. Ricciuti and A. De Santis, Food intake actions detection: An improved algorithm toward real-time analysis, *Journal of Imaging*, vol.6, no.3, p.12, 2020.

[10] D. Ameta, K. Mudaliar and P. Patel, Medication reminder and healthcare – An android application, *International Journal of Managing Public Sector Information and Communication Technologies (IJMPICT)*, vol.6, no.2, 2015.

[11] G. A. Bilodeau and S. Ammouri, Monitoring of medication intake using a camera system, *Journal of Medical Systems*, vol.35, no.3, pp.377-389, 2011.

[12] F. M. Hasanuzzaman, Y. Tian and Q. Liu, Identifying medicine bottles by incorporating RFID and video analysis, *Proc. of the IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*, Atlanta, GA, USA, pp.528-529, 2011.

[13] M. Valin, J. Meunier, A. St-Arnaud and J. Rousseau, Video surveillance of medication intake, *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pp.6396-6399, 2006.

[14] P. T. Seint and T. T. Zin, Behavior analysis for nursing home monitoring system, *International Conference on Big Data Analysis and Deep Learning Applications*, Singapore, pp.274-283, 2018.

[15] H. H. Huynh, J. Meunier, J. Sequeira and M. Daniel, Real time detection, tracking and recognition of medication intake, *World Academy of Science, Engineering and Technology*, vol.60, pp.280-287, 2009.

[16] P. T. Seint, T. T. Zin and M. Yokota, Medication and meal intake monitoring using human-object interaction, *2018 IEEE 7th Global Conference on Consumer Electronics (GCCE)*, pp.399-400, 2018.

[17] M. J. Jones and J. M. Rehg, Statistical color models with application to skin detection, *International Journal of Computer Vision (IJCV)*, vol.46, no.1, pp.81-96, 2002.

[18] S. L. Phung, A. Bouzerdoum and D. Chai, A novel skin color model in ycbcr color space and its application to human face detection, *IEEE International Conference on Image Processing (ICIP)*, vol.1, pp.289-292, 2002.

[19] J. Shi and C. Tomasi, Good features to track, *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.593-600, 1994.

[20] G. Kaur, A. S. Arora and V. K. Jain, Multi-class support vector machine classifier in EMG diagnosis, *WSEAS Trans. on Signal Processing*, vol.5, pp.379-389, 2009.