International Journal of Engineering & Technology, 8 (1.4) (2019) 533-544

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Unusual Event Detection Algorithm via Personalized Daily Activity and Vision Patterns for Single Households

Junho Ahn[1], Hwijune Park[1], Juho Jung[1], Gwang Lee[1]*

[1] Computer Information Technology, Korea National University of Transportation, Chungju, Korea
**Corresponding author: Gwang Lee*
**Corresponding author E-mail: gwang@ut.ac.kr*

Abstract

People may be too seriously injured, incapable or endangered in emergency situations in their houses. Roommates or family members who live together in the same house can rescue or call 911 for them in the dangerous situations at home. The growth in the number of single-person households is currently rising over the decades and they have difficulties to get help from other people in case of serious injuries in their houses. There are surveillance video camera systems used which can simply classify the user behaviors to identify accident situations in a limited range of indoor areas where cameras are installed. To deal with this limitation, we propose a fusion algorithm to detect personalized unusual events via daily activity and vision patterns for a single household at home. We designed and implemented the proposed algorithm with the smartphone sensors, and a video camera installed in indoor areas. We evaluated individual activity and vision algorithms, and simulated the proposed fusion algorithm in scenarios.

Keywords: *Unusual event, Single Household, Activity, Vision, Fusion Algorithm.*

1. Introduction

According to the 2017 report from household composition statistics in Eurostat [1], the major common type of household was single-person households with one third (33.1 %) of the total number of households in 2016. The proportions of single-person households without children by Germany, France, Italy, United Kingdom, and Sweden are 40.5 percentage, 34.7 percentage, 33.0 percentage, 31.3 percentage, and 51.8 percentage respectively in 2016. If nobody can come to their rescue or call 911 for them in some emergency situations like serious injuries or incapable incapability happens, they may be continuously wounded or seriously endangered.

There are smartphone applications [2, 3] and camera-based systems [5, 6, 7, 8] to identify emergency situations. However, the mobile applications need the users to carry the phone on the body to detect the users' behavior in indoor areas. The vision systems can classify the limited accident situations in a small range of a camera installed at home, but it cannot cover the whole areas in the house. We define "unusual events" to be infrequently-generated behaviors, such as no movement activity or an extreme stationary behavior in the same location.

According to our previous research [9, 10], we also investigated that people's daily behaviors are frequently repeated in their lives and thus we can predict a person's future behaviors by measuring their daily behaviors. We collected data using sensors on the mobile phone--such as GPS, WiFi, Accelerometer, Proximity, Light, Orientation, and Gyroscope. We found that the accelerometer sensor of the smartphone is efficient to continuously detect user movement with low battery power and high accuracy.

We propose an accurate fusion algorithm to identify unusual events utilizing user movement patterns with the smartphone accelerometer, and with a vision camera installed in the house. The movement classifier with the accelerometer on smartphone detects users' activity behavior (no movement, small movement or slow walking, normal walking or running) to classify if the users are in emergency events. The video camera-based algorithm detects people or home objects to classify unusual situations, but each classifier's accuracy was not suitable to cover the whole areas in the house. We analyzed and designed a method to combine all of these classifications for developing the accuracy performance of the detections. Lastly, we investigated a fusion algorithm to improve the detection performance by combining all identification results.

We describe related works in this area in section 2. Section 3 describes the activity classifier with the accelerometer on the smartphone and the vision method with video camera, and explains the fusion algorithm. And Section 4 describes each classifier's evaluations and simulation results of the fusion algorithm. Our ideas for future work are presented in section 5 while the conclusions are presented in section 6.

2. Related Work

There are human activity detection researches [2, 3, 4, 5, 6, 7] with an accelerometer to identify unusual events. These researches focused on detecting users' falling situations as unusual events with an accelerometer. Other activity researches [4, 5] used the activity sensor to detect movement and activity patterns [6, 7] for identifying unusual falling situations on the smartphone. They investigated to classify falling events of patients and elderly people.

Video camera-based detection classifiers [8, 9, 10, 11, 12] were utilized to detect unusual events with user vision patterns. These researches investigated vision-based user behaviors and extracted meaningful features to classify unusual situations. In particular, a vision-based research [9] identified human postures using silhouettes and CNN. There is a normal behavior detection research to identify pedestrian and non-pedestrian [10]. A vision research [12] detected a person's angle, rate, and velocity with the KNN classifiers to detect abnormal events.

In our previous behavior researches [13, 14], we investigated to identify behavior situations for improving the accuracy performance. These researches fused audio, accelerometer, GPS, and Wi-Fi sensors to improve each sensor's accuracy in normal lives. An individual detection algorithm only using one sensor provides a limited accuracy performance in a variety of domains. We developed the fusion algorithm to improve abnormal event identifications in home settings. We also examined to develop movement detections using the accelerometer on the smartphone for identifying user activity behaviours. [15, 16]

We explored some notification applications that provide event information to a user [17, 18] or that are used to request help when a user is experiencing an emergency event [19]. These applications provides alert information to users to help them determine if they are located in, or passing through, a dangerous place or safe place. These applications show recent alert information (i.e., location and type of crime), within a specified recent period and within the surrounding area of the users, selected by them on the mobile phone application. However, the users need to manually check for this information and it does not automatically provide alert information periodically when they move to other locations. SOS Emergency App [10] and Medical Alert [11] provide a functionality for the users to get help by pressing a single button when they find themselves in emergency situations. Users using these applications can easily obtain help, but it requires a manual method to get help by pressing a single button. If the users are not physically capable of pressing the button during the emergency situation, they cannot be rescued.

Offender Locator Lite [19] and Family Locator - Phone Tracker [20] applications are able to share users' locations among family members using the mobile phone. Additionally, the Offender Locator Lite application provides sex offenders' locations to users who are in proximity to sex offenders' homes, based on the users' locations determined from the GPS sensor on the mobile phone. With both applications, family members are able to track other members' locations, but both require regularly expending large amounts of battery power because they use a duty-cycling method which requires reading sensors in fixed small intervals of time. This method has an energy-accuracy trade-off problem, which means it must expend a great deal of power to increase accuracy and that when saving power, accuracy is decreased. To solve this limitation of the duty-cycling method, adaptive algorithms [21, 22] have been proposed as a solution. A-Loc [21] continually fine tunes the energy consumption to meet changing accuracy requirements by using the GPS, Wi-Fi, Bluetooth, and Cell-Tower sensors on the mobile phone. A user chooses a destination and the algorithm adjusts to use the most optimal sensor, according to the destination's distance from the user. If a user is far from the destination, the algorithm uses a cell tower or Wi-Fi to localize the user and to conserve battery power. However, if the user is close to the destination, the algorithm uses the GPS sensor to obtain a more accurate location, but it consumes considerable battery power. This algorithm is also used to detect a user's daily walking path and gait, but as noted, it requires the user to first select a destination in order to use the adaptive algorithm. This algorithm is also unable to provide accurate localization data the farther a user is from the destination. RAPS (rate-adaptive positioning system) [22] localizes a user's location without requiring the user to select a destination, by using an adaptive algorithm supported by mobile sensors such as GPS, Wi-Fi, Bluetooth, Cell-Tower, and Accelerometer. This algorithm has increased location accuracy and consumes less battery power, while also measuring the user's walking path and pattern. However, the RAPS algorithm cannot be adjusted to increase the localization accuracy of the user when he is in a dangerous place, and consume lesser battery power when he is in a safe place.

3. Algorithm

3.1. Location Detection

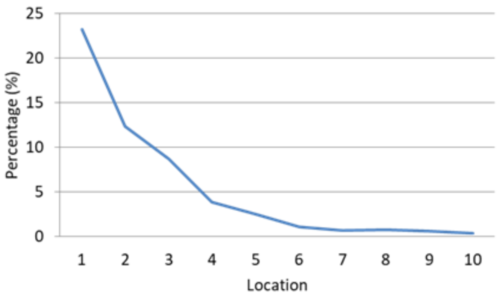

Fig. 1: Experiment of user visited locations

We estimate the user locations to check if the single-person households are in or near the house where video cameras are installed. We conducted basic GPS connection experiments using the smartphone for the user location detection.

The location data our mobile application is stored on the phone whenever the user moves from one location to another, and this information is used to build a historical location-based map. The historical map is then used to estimate whether or not the mobile user is in a habitually or frequently visited place. We used the Gaussian (Normal) distribution to aggregate historical data, based upon location data points collected from GPS or Wi-Fi sensors. From this aggregated historical data we determined an acceptable distance range for each data point recorded by the sensors. We used the ranges to determine a distance traveled measurement and a circle area covering range. We calculated this distance measurement, from each data point to the next one as, for example: (x1, y1) to (x2, y2), and then took this distance as the radius of a circle and drew an area covering range around this distance measurement. From these calculations we could thus determine a walking or driving distance and a covering range around that distance and location measurements. When a user frequently visits in the same place, the size of the covering range is changed according to the next GPS connection. The number of covering ranges is increased whenever the user passes through the place. We measured Gaussian distribution using the multiple ranges created in the same place. In our previous research [13, 14], we found that the subjects spent 92% of their one week's time period in four main repeatedly visited places as shown in Figure 1. The 20 subjects spent two thirds (65%) of their total week's time in the first ranked location, 14% of their time in a second ranked location, and 8% and 5%, respectively in the third and fourth ranked top locations visited.

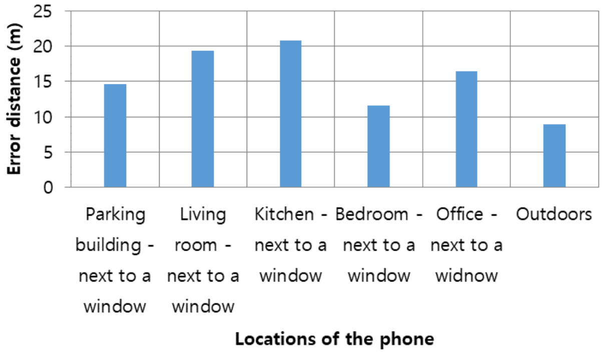

Fig. 2: Experiment using GPS connection time on the HTC Nexus One

We conducted an experiment to determine whether a user is located in an outdoor area or not by using the GPS sensor on the HTC Nexus One as shown in Figure 2. We implemented an application that obtains the phone's location in both outdoor and indoor areas. Figure 2 shows the average of distance errors using the GPS using the phone at home, office and outdoor area. During the experiment, we found that when the phone was in an outdoor area, the GPS sensor could connect to the satellites correctly. Usually when the phone was in an indoor area, it could not receive any signal from the satellites. However, when the phone was positioned close to a window in an indoor area, it sometimes obtained its location correctly. Therefore, we determined that if the phone was not receiving any signal from the satellites, we could affirm that it was in an indoor area, but if it connected to them, we could not determine whether it was located indoors or outdoors. Thus, we decided that the GPS sensor can be used only as a partial indicator to determine if a user is located indoors or outside.

Wi-Fi works indoors, which we have verified to be practically true in typical indoor environments. We could track the user location in an indoor area. There may be occasional wireless disconnection, so we utilized the last GPS connection as the user location and the algorithm recognize the user is in an indoor nearby. If the mobile device is disconnected/off near the user home, we determined the user is at home.

3.2. Activity Classifier

Any data generated by our application during the live system demonstration for a participant is stored on either the smartphone. Further, our smartphone demonstration application collects any activity information from the participant during the demonstration.

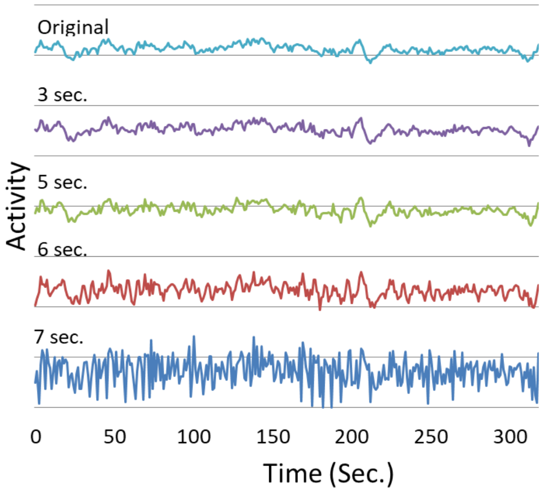

Fig. 3: Comparing with Activities depending on a size of windows

In order to measure a user's average normal movement impact (e.g., slow walking, normal walking, running) with accelerometer data, we first tried to measure the data in one-second intervals, but discovered that the amplitude, as shown in Figure 3, fluctuated greatly and contained a lot of unnecessary "noise" in the data. However, we compared this time-interval measurement to other averaged timespan windows of 3, 5, 6 and 7 seconds. We found that after the 3-second window measurement, the amplitude lines for the other windows were very similar and fluctuated much less than the original one-second and the 3-second windows. Therefore, we chose the 5-second size window as a reasonable window size for our accelerometer measurement, so that when we used this window, the data are normalized. Although using the 5-second window normalized data is sufficient for determining average normal movement impact of the user, it alone does not allow us to detect quick, sudden high impact movement, such as falling. The averaged normalized data smooths out any high impact activity measurements. Thus, we also use the raw one-second interval data in combination with the 5-second window normalized data to detect such high impact movements, like sudden falling.

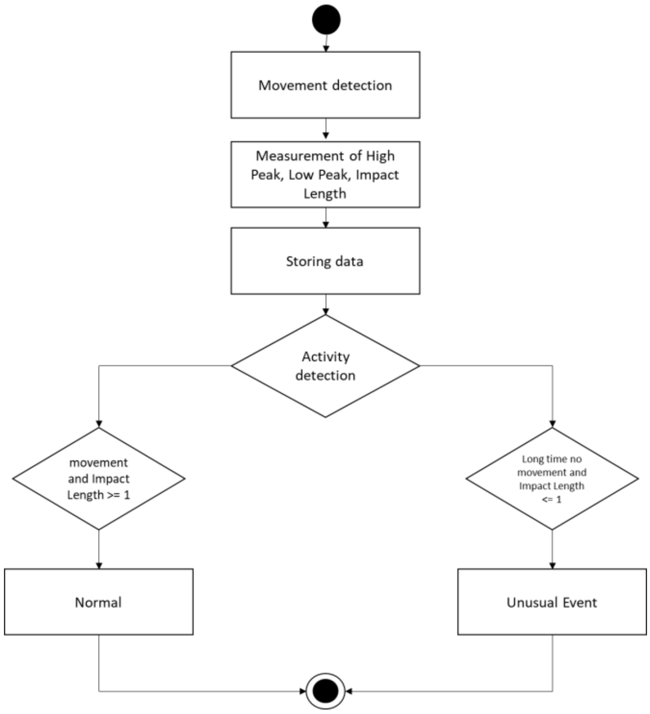

Fig. 4: Flowchart diagram for the activity classifier.

Our application identifies the movement of the mobile accelerometer sensor on the smartphone as shown in Figure 4. It measured the movement with a multistep signal processing method. We calculated the sum of the x-axis, y-axis, and z-axis accelerometer values and measured the normalized average of accelerometer signals. We classify peak values with the average type of accelerometer-based activity. We examine movement strength using impact of between positive and negative peaks. The classifier detects the activities of smartphone users with threshold detection scheme in the processed accelerometer signal trace. Based on this detection method, we realized that if there is no movement in a long time it is classified as unusual events. In order to determine whether the phone is completely immobilized, we compared the sensor data from two different stationary events: a phone placed on desk, and a user sitting on a chair, as shown in Figure 5. From this experiment, we observed that when a user is stationary (e.g., sitting on a chair) and is carrying the phone on his or her person (e.g., in a pocket), there is still very minor movement activity occurring which the phone will measure through automatic activation of the accelerometer. The accelerometer will not activate automatically, however, in the cases when the phone is placed in a stationary location, such as a table, and is not carried anywhere on the user's body. We implemented a mobile application and collected the user’s daily activity using the accelerometer sensor on the phone as shown in Figure 6.

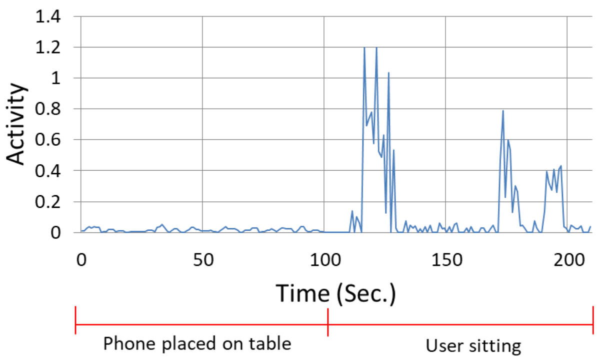

Fig. 5: Stationary activity on the mobile phone.

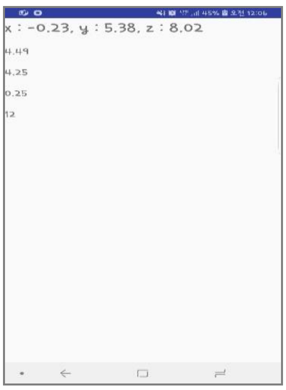

Fig. 6: The mobile application for the activity classifier.

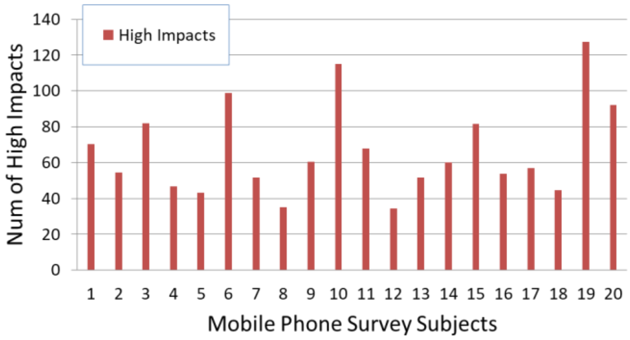

Fig. 7: Basic experiment of the average number of high impact activities per day according to survey subjects.

High impact activities occur in normal life situations and events. We measured the number of initial high impact events that occurred for users of our mobile phone survey. We averaged the number of initial high impact events that occurred for each user, and have summarized these data for each individual phone survey subject as shown in Figure 7. The number of high impact events shown in this Figure represents only the first high impact measurement of the accelerometer; subsequent continuous accelerometer readings of these initially identified high impact events were ignored. The experiment shows the average and the standard deviation of the initial high impact activity measurements calculated across all 20 survey subjects' data. The total average number of initial high impact activities is 62.41 for all subjects and there was much variation across phone survey subjects in the number of high impact activities they performed. Based on this analysis, the high impact is not strongly relevant to unusual event detection using smartphone.

3.3. Vision Classifier

We designed and improved the existing vision algorithm for detecting user behavior events. Our proposed algorithm is experimented with a SNH-V6410PN video camera as shown in Figure 8. The existing algorithms like Haar cascade, Hog cascade, SSD mobinet, faster rcnn inception, etc. focus on detecting object or human using vision data. We also detected objects based on the existing Faster-Rcnn algorithm, and extended the simple object approach to the behavior pattern approach. We investigated detection performances of the existing algorithms such as Haar cascade [8], Hog cascade [8], SSD mobilenet [23], and Faster rcnn inception [24] to identify objects in home settings. We analyzed 100 or more video data collected from CCTV and measured each algorithm's accuracy performance as shown in Table 1. Overall, we realized that the Faster-rcnn inception algorithm is better than other algorithms. We extended the Faster-rcnn inception method to our classifier for the vision algorithm to classify unusual situations.

Fig. 8: Experiment video camera: SNH-V6410PN.

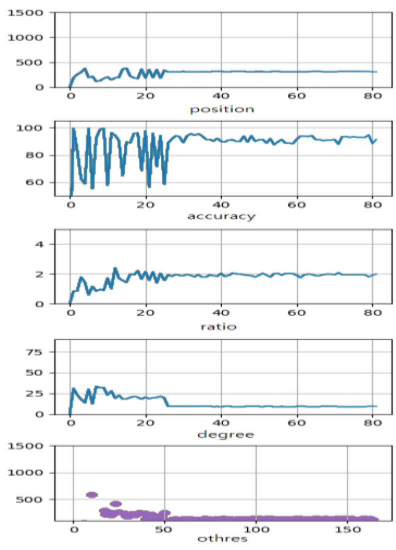

Fig. 9: Position, accuracy, ratio, degree, and others for human and Object detection using Faster r-cnn inception algorithm.

We analyzed human and objects at home using Faster-rcnn inception algorithm as shown in Figure 9. We measured transition of human position accuracy through ratio, degree, others according to time. We could do a variety of experiments in a normal or unusual life. By using this method, figure 10 shows an example of object and human detections with accuracy percentage in indoor areas.

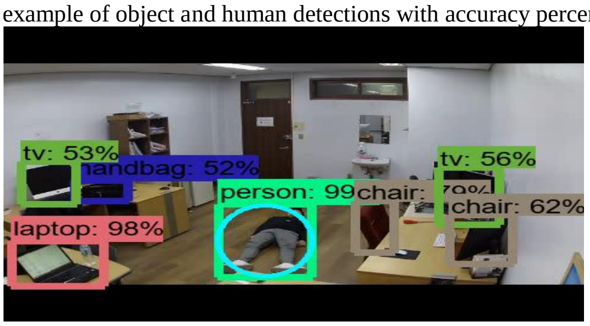

Fig. 10: An example of unusual event classification using object and Human detection in the indoor area.

Table 1: Performance results of the existing vision classifier in indoor areas

Algorithm	Recall	Precision	Accuracy
Haar cascade	0.48	0.096	0.09
Hog cascade	0.52	0.36	0.34
Object Detection ssd_mobilenet	0.92	0.68	0.68
Object Detection Faster-rcnn-inception	0.98	0.95	0.94

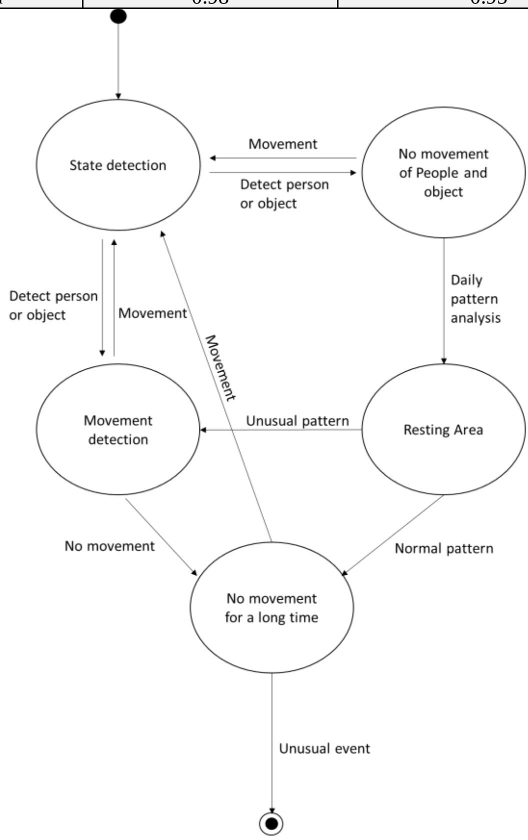

Fig. 11: State chart diagram for the vision classifier to detect single household events

Fig. 12: Human and object detection in a normal life using our vision classifier.

The vision classifier detects movements using video data collected from the camera as shown in Figure 11. It identifies whether the movement things are objects or human as shown in Figure 12 in a normal life. We analyzed that the object movement at home can be an indicator to check if the single person is normal in the life. In other words, the vision algorithms sometimes may not detect human moving in a camera range because of non-human posture like four foot pose, body hidden by furniture, a part of body detection, and so on. In these cases, although the person is moving in a camera detection area, the vision algorithm cannot identify the human correctly. However, if the person uses cups or books in the life, the objects' locations are moved and the vision classifier can detect the object movement. Thus, if an object is moved to other locations or has disappeared, we analyzed that the person is in a normal life.

Fig. 13: Indoor object detection using our vision classifier (upper one: raw camera image, lower one: frequently used furniture)

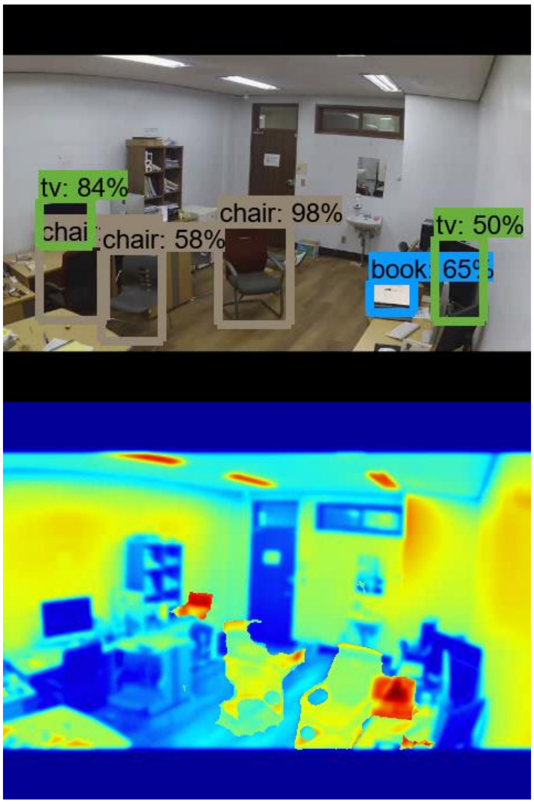

Fig. 14: Indoor detection screens using our vision classifier (upper one: furniture detection accuracy percentage, lower one: daily behavior-based staying locations)

Figure 13 shows that there are some furniture which can be used to study, play games or watch TV in indoor area. Using the Faster-rcnn-inception method, we detected them and classified frequently used furniture based on user daily behaviors as shown Figure 14.

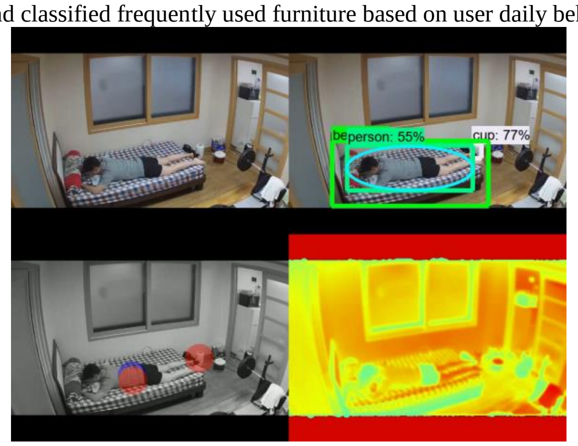

Fig. 15: Normal stationary behavior on the bed and detection using our vision classifier

In a non-movement state based on the detected person's location, our vision classifier identifies whether a person is normal or unusual in the life. When there is a person detected in a camera range, and if the person never moves after a long time in the area, the proposed algorithm classifies whether the person is in a resting furniture like a chair, a bed, a couch, etc. where the person took time in a normal life as shown in Figure 15. We also measured user daily behaviors to find normal locations to take time in a normal life. The single person stayed in the location for a specified timeframe, we detected it as normal. Otherwise we detected that it is unusual in case that the person is located on the table, under the table,on the floor of the room, etc. where one never or infrequently took time in the normal life as shown in Figure 16.

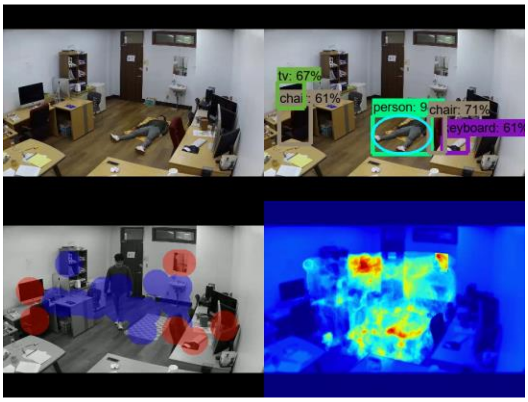

Fig. 16: Unusual stationary behavior in an infrequent location and detection using our vision classifier.

Although the person is in the frequently located places in daily behavior or in the resting furniture, if the time is longer than user's daily staying time, our algorithm detected that it is unusual.

3.5. Smartphone Function Classifiers

When we analyzed based on activities of users who were stationary during these segments, we identified two activities: user game-playing and phone calling activities. For example, Figure 17 displays a comparison of user-movement activity data measured when a user is stationary contrasted to data measured when the same user is holding their devices and playing a game on their smartphone. This example illustrates the difficulty in measuring any user-phone interaction activities accurately, such as texting, web-surfing, skyping, etc. In addition to our identification of the data generated with user-phone interaction activities, such as game-playing, texting, etc. We found that these activities are normal events and these mobile sensors can be used to identify events. Additionally, when users charge their smartphone, we could not recognize normal or unusual event detection.

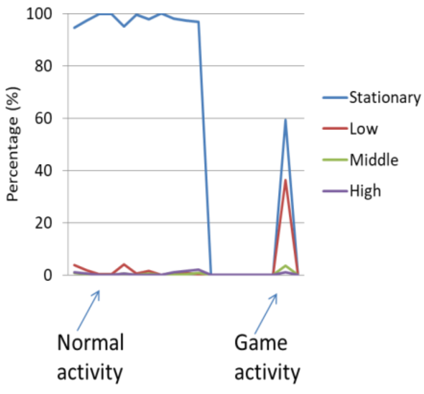

Fig. 17: Example of noise data generated on mobile phone when user is stationary: Game activity

3.5. Other Classifiers

We also measured a light strength using a light sensor on the HTC Sensation to decide if a user is in an outdoor area during the daytime. The light strength of the sun is much stronger than the strength of an electric light. The light sensor of the phone can correctly read this strength and can determine if the scanned light is sunlight or light generated by electricity. Figure 18 shows the average of a light strength according to a location. We performed the experiment in an outdoor area on a sunny day (below a big tree, a small tree, direct sunlight), on a cloudy day, and in an indoor area (inside and next to a window). As a result of this experiment, we accurately determined that a user is located in an outdoor area during the daytime if the light sensor detects a light strength of more than 1028 Lux. Therefore, the light sensor can be used as another indicator for determining outside versus inside areas during the daytime. Although the smartphone's light sensor can detect if the user is an indoor area, the user normally carried on their phone on the pocket or a bag. Thus, we mainly focused on the activity and vision classifiers to detect user behaviors in indoor areas.

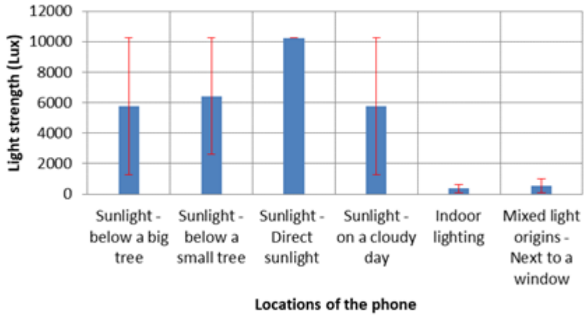

Fig. 18: Experiment using a Light sensor on the HTC Sensation

3.6. Fusion Classifier

Our fusion algorithm mainly combined the vision detection classifier with a video camera installed in houses and the behavior activity classifier with the smartphone accelerometer. We combined the various classifiers to detect whether single households are normal or unusual in their lives. We evaluated the fusion algorithm for detecting user normal behaviors via the activity and the vision algorithms in sixteen scenarios. We designed the scenarios based on the Youtube CCTV video data in normal lives and analyzed the detection performance for detecting user behaviors using our fusion classifier in the experiment result section.

4. Experiment Results

We evaluated the accuracy performances of both the individual vision classifier and individual activity classifier and measured the performance of the proposed fusion algorithm in scenarios.

User daily activity behavior at home can be used to predict unusual event in a daily life. We investigated CCTV video data collected from Youtube and analyzed 20 activity behaviors in a normal life. We collected experiment data from 5 people using the 20 normal behaviors in the house. Our proposed activity classifier identified the following behaviors: holding the smartphone and watching TV, Touching the phone and watching TV, putting the phone on the desk and watching TV, holding the phone and walking, putting the phone on the pocket and walking, putting the phone on a desk and walking, putting the phone in the back and walking, holding the phone and no movement, putting the phone on the pocket and no movement, holding on the body and moving, holding on the body and no movement, putting the phone on the backpack and no movement, putting the phone in the handbag and no movement, holding the phone and staying on a sofa, putting the phone in the pocket and staying on a sofa, playing games on the bed, infrequently touching the phone and being on the bed, putting on the phone in the backpack and sometimes moving, putting on the handbag and sometimes moving in indoor areas. Table 2 shows the experiment results based on these behaviors for measuring the accuracy performance to detect normal events. The overall accuracy was 0.833, and the recall and precision were 0.794 and 1.00 respectively.

Table 2: Recall, precision, and accuracy performance of our activity classifier for user behaviour detection in indoor areas

Classifier	Recall	Precision	Accuracy
Activity classifier performance	0.794	1	0.833

We measured the accuracy performance of our proposed vision algorithm by comparing the existing vision detection algorithm in indoor areas. We collected samples of CCTV video data from 60 Youtube videos of people involved in actual normal life scenarios, etc., and also recorded 71 video data in normal lives collected from our cameras. Our vision classifier identified lying down situations in a camera range. Our classifier found the status with higher performance of 96.1 recall, 94.3 precision, and 93.1 accuracy measurement than the performance of the existing algorithm in houses as shown Table 3. Additionally, we analysed performance of object movements, resting areas or furniture, a half of body, and both a half of body and object detection. The performance object movement classification was 97.1 recall, 95.7 precision, and 94.5 accuracy in the houses. The accuracy performance of resting area or furniture, half body detection, and mixed one was 90.3, 92.1, and 90.3 accuracy for the vision detection. We found that our vision algorithm for detecting normal events was performed overall higher than the existing algorithm.

Table 3: Recall, precision, and accuracy performance of our vision classifier for user behaviour detection in indoor areas

Behavior	Detection	recall	precision	accuracy
Lying down	Faster-rcnn-inception	96.1	87.7	87.6
	Proposed Algorithm	96.1	94.3	93.1
Object Movements	Faster-rcnn-inception	96.1	87.7	87.6
	Proposed Algorithm	97.1	95.7	94.5
Resting area or furniture	Faster-rcnn-inception	82.2	87.9	78.3
	Proposed Algorithm	91.9	95.0	90.3

Half body detection	Faster-rcnn-inception	96.2	91.2	89.0
	Proposed Algorithm	96.2	94.5	92.1
Half body + Object detection	Faster-rcnn-inception	91.9	87.9	78.3
	Proposed Algorithm	91.9	95.0	90.3

We conducted experiments for detecting user normal behaviors using our fusion classifier based on the activity and the vision algorithms in sixteen scenarios. We made the scenarios based on the Youtube CCTV video data in normal lives and measured the detection performance for detecting user behaviors using our fusion classifier. Table 4 shows scenarios carrying on a smartphone within a camera detection range: walking, running, lying on one's side, lying on one's side back, and lying on one's face back.

Table 4: Recall, precision, and accuracy performance of the fusion classifier in scenarios carrying on a smartphone within a camera detection range

Scenarios carrying on a smartphone within a camera detection range	Vision Classifier	Activity Classifier	Fusion Classifier
Walking	O	O	O
Running	O	O	O
Lying on one's side	O	O	O
Lying on one's side back	X	O	O
Lying on one's face back	O	O	O

Table 5 shows scenarios carrying on a smartphone and with limited video camera detections: movement & out of a camera range, movement & partial body recording (an upper body, a lower body, and one third on body), and movement & turning off the light. The activity classifier detected the user movement, but vision classifier limitedly identified the user behavior due to the partial video recording of the user body in these scenarios.

Table 5: Recall, precision, and accuracy performance of the fusion classifier in scenarios carrying on the smartphone with limited normal video camera detections

Scenarios carrying on the smartphone with limited normal video camera detections	Vision Classifier	Activity Classifier	Fusion Classifier
Movement & Out of a camera range	X	O	O
Movement & Video recording of an upper body	O	O	O
Movement & Video recording of an lower body	O	O	O
Movement & Video recording of one third on body	X	O	O
Movement & Turning off the light	O	O	O

Table 6 shows scenarios with objects carrying on a smartphone within a camera detection range: sitting in a chair, no movement & sitting down on a chair, sitting with back turned on a chair, and falling down & leaning in an object. The activity classifier detected the user behavior, and vision classifier identified the user behavior and objects in these scenarios.

Table 6: Recall, precision, and accuracy performance of the fusion classifier in Scenarios with objects carrying on smartphone in a camera range

Scenarios with objects carrying on smartphone in a camera range	Vision Classifier	Activity Classifier	Fusion Classifier
Sitting in a chair	O	O	O
No movement & Sitting down on a chair	O	O	O
Sitting with back turned on a chair	O	O	O
Falling down & Leaning in an object	O	O	O

Table 7 shows scenarios without the smartphone within a camera detection range: movement & charging the smartphone and movement & putting the smartphone on a table. The activity classifier does not detect the user activity, and vision classifier identified the user behavior in these scenarios. Although the activity classifier does not detect user behaviors when the users do not carry on their phones, we could detect that the user does not carry their phone and it is charging somewhere using the status sensor on the smartphone.

Table 7: Recall, precision, and accuracy performance of the fusion classifier in scenarios within a camera detection range without the smartphone

Scenarios within a camera detection range without the smartphone	Vision Classifier	Activity Classifier	Fusion Classifier
Movement & Charging the smartphone	O	X	O
Movement & Putting the smartphone on a table	O	X	O

We found that our proposed fusion classifier detected all of the user behavior based on the activity and the vision measurement the high accuracy performance in sixteen scenarios.

5. Future Work

Assessing classifiers built to detect unusual events is hard, because one needs large volumes of test data in order to have enough rare events to comment on the efficacy of the classifiers. We are in the process of collecting more mobile user data, but it's unclear how much more data will be needed to complete the validation. We expound here upon some experiences we have had in this regard. Building a danger notification system is harder than building a system, because the penalty for false positives, e.g. notifying the police or parents, is much higher than the penalty for simply recording extra information either on the mobile or the cloud. Moreover, the difficulty of differentiating which among the unusual events are actually dangerous, such as an epilepsy compared to dancing - which might both reveal a similar set of readings - is much more challenging, because it involves developing highly sophisticated classifiers that need detailed and complex contextual clues. To help gain some insight on dangerous situations, in the same phone survey as mentioned earlier, in our previous survey research, we also asked users to rate every 30 minutes the "danger level" of a visited place, using a 5-point scale (1=very safe to 5=very dangerous), i.e. how safe they felt in these places.
The definition for "danger level" for this survey question was open-ended and subjective, according to an individual user's perspective and how safe they felt in a particular location. As part of the paper survey after the phone traces were collected, participants were asked to assess their estimated feeling of safety in these places. Also, more importantly, we included other questions on the paper survey about

how aware the participant is of publicly posted crime statistics and sex offenders' house locations. The results of the survey were that all participants selected only the options: "very safe", "safe", or "somewhat safe" for their ratings of every place at which they recorded the frequency of their visit. None of the participants chose "dangerous" or "very dangerous" in their assessment of how safe they felt in any location, because none of them experienced a dangerous situation, such as an accident and crime, during the experiment period. Although they didn't rate any situation or location as dangerous during this period, they did provide data on varying feelings of safety for three of the five possible ratings that we could compare with actual locations visited. We found three different types of users in our 20-person survey experiment that we could group according to their assessed safety level ratings for various locations and times of visits.

In the first user group type, the subjects' safety feeling ratings changed according to how commonly or frequently visited a location was for them, regardless of the time of day. In the second user group, subjects clearly felt varying levels of safety in different locations, according to the whether they were visiting that location in the daytime or the nighttime. For purposes of this survey experiment we defined day-time as the hours from 8am to 6pm, and we defined nighttime as the hours from 6pm to 8 am. In addition to these users' feelings of safety varying according to daytime or nighttime in the same locations, their safety ratings also varied somewhat according to type of location (e.g., frequently visited, etc.), as did the safety ratings of Group 1 Users. In the third group of users, subjects felt very safe regardless of the time of the day or the frequency of visitation for a location. After reviewing the results of this survey experiment, we found that users' safety feeling ratings varied widely, according to different types of users, and that they were not informed by any factual knowledge about the locations visited (e.g., by public crime or accident data). From the phone survey data, it became apparent that the same locations could be evaluated very differently by different types of users. Thus users' safety ratings for a location are often very subjective, and therefore not very helpful in predicting locations where there are higher probabilities of accident and crime situations occurring. According to accident and crime statistics from a police database, the number of accidents and crimes are different depending on a region and are more likely to occur in the afternoon time than in the evening or night time. We noted that our phone survey subjects' feelings of safety or danger were quite unmatched to findings from statistical data of actual accidents and crimes. When the participants of our phone survey experiment finished data collection on the mobile phone, they also were asked to complete a paper survey. On this questionnaire, we asked the subjects if they had ever seen publicized crime rates, provided by the government on the web or in the papers, for their geographic location–their neighborhood, their company, or school. All of the subjects indicated that they had NOT seen any publicly posted crime rates covering the area where they live, and many of them had not known that these crime rates were available. As a result, a danger notification application should be very careful to not be misled by the perceived danger of users, but rather should focus on the ground truth derived from actual danger posed to users. Our adaptive fusion algorithm extends the longevity. However, in the typical use case, users may simply be recharging their phones at night, in which case extending the lifetime of the phone beyond 12 hours does not accrue much benefit. Instead, we could use the energy savings afforded by Adaptive Fusion to increase the sampling rate during normal hours of user activity, and thus improve the accuracy of our unusual event detection while maintaining a lifetime long enough to last a day. That is, our Adaptive Fusion approach allows us the freedom to tradeoff extending the energy lifetime of the application for increasing the sampling rate and hence the accuracy of classification. We intend to experiment with the case of increasing the sampling rate in future work. For the purposes of this paper, we decided that those three sensor dimensions would be sufficient to demonstrate the concept of detecting most unusual events of interest relevant to mobile devices. However, we intend to investigate the use of the camera, and other sensing modalities such as light, compass, and gyroscopes in improving the detection of unusual events. Incorporating user feedback and input into the design is a direction we intend to explore. Our idea here is two-fold. First, we'd like feedback to adjust the thresholds for classification. For example, if a user is at a party or visiting a bar, the noise level in that environment will likely be very high and somewhat unpredictable, which would decrease the audio measurement accuracy. On the other hand, if the user is at home, the measurement accuracy will be very high and predictable. The audio detection measures the pitch level of human voices, and thus is very dependent upon the number of people and activity occurring in the environment of the user's mobile phone. Thus, in order to better classify the voice sounds in different locations and situations, whenever the audio classification algorithm detects a higher number of high-pitched voice sounds, we could survey the user within the application. If a user said that the "abnormal classification" was incorrect, we would adjust the threshold of high-pitched audio sounds allowed to fit the current environment. If they said it was correct, we would not need to make any adjustments to the threshold. Second, we consider the idea of incorporating an emergency button that the user could push to force recording. Finally, our current design only uses sensor data available on the mobile (along with user history) to efficiently determine an unusual event. Since the system is cloud-friendly, we can improve efficiency of adaptive fusion by for example considering whether the current location is a crime-ridden area - information that can be provided by the cloud. This would provide a cloud driven criterion for adaptively invoking various sensors.

Until now, these applications and algorithms have been successfully implemented and developed to help users partially avoid dangerous places and be rescued in emergency situations. However they have not been able to provide optimal and efficient protection for users of the mobile phone who find themselves in crime and emergency situations. I propose an emergency application for the mobile phone which will inform users of dangerous places near their current location, detect possible dangerous situations (e.g., crime locations) or emergency situations (e.g., car accidents, assaults, rape, etc.), and record the user's location, activity, and audio like a black box recorder. When a user walks or moves through a dangerous place where crimes or accidents are known to frequently occur, the proposed system will notify the user of this fact, by using statistic and real-time information. If the user is involved in a crime or emergency situation, the application will detect that such a situation is occurring and automatically call 911 and/or relatives so that the user can receive help. The application also records the emergency situation as it's occurring in order to save a record of the clues or evidence regarding the crime or accident. The proposed system provides these three types of functionality (user notification, emergency detection, black box) by using mobile sensors installed on the phone without any additional devices. However, normally the implementation of these three functionalities would require an extensive amount of battery power to provide these features. To overcome this battery limitation on the mobile phone, this application will also include an efficient and adaptive sensing algorithm, which can optimize the use of battery power to increase the accuracy of detection when users are in dangerous/emergency situations and save battery power when they are in safe situations.

A possible scenario illustrating the use of this system might be as follows. Suppose a student walks to school every weekday morning along a particular street in his neighborhood. The mobile phone learns, by recording, the user's location, the time, and this daily walking pattern. To save battery power, the application is set to collect these measurements infrequently in normal situations such as this one. Then suppose one day, the student is walking down this same street with the same walking gait, and a kidnapper grabs the student and pulls him into a car and drives away down a different street. The system will recognize that the student's usual walking pattern and direction have drastically changed and then will begin to frequently access the GPS and audio sensor. These sensors are able to detect that the student is now proceeding down an unknown path and that he is screaming. The application automatically notifies 911 and relatives who have been previously entered into the application. It also records the voices of any conversation, the location, and all of the student's movement activities—thus operating like a black box recorder to save clues and evidence relevant to this crime. Law enforcement personnel in this identified area can then be notified quickly by 911 as well as relatives to hopefully come to the rescue of this student as soon as possible. Therefore, the system is able to utilize these three functionalities to rescue users from crime or accident situations with the use of only a mobile phone that requires no additional devices.

6. Conclusion

We have presented the fusion algorithm to detect personalized unusual events via daily activity and vision patterns for a single household member at home. We evaluated the accuracy performances of both the individual vision classifier and individual activity classifier and measured the performance of the proposed fusion algorithm in scenarios. We hope to extend and implement that a practical application consists of a mobile client component that adaptively fuses location, activity, and vision to efficiently detect unusual events, and then log them both locally on the mobile and to a cloud-based server component.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF), grant funded by the Korea government (MSIP; Ministry of Science, ICT & Future Planning) (No. 2017R1C1B5017847). This was also supported by Korea National University of Transportation in 2018

References

[1] Eurostat, "Household composition statistics", European Union (EU) Labour Force Survey (LFS), 2017, available online: http://ec.europa.eu/eurostat/statistics-explained/index.php/Household_composition_statistics
[2] Monisha Mohan, Arun P.S, ACCELEROMETER–BASED HUMAN FALL DETECTION AND RESPONSE USING SMARTPHONES, International Journal of Computer Engineering In Research Trends,5, 2017
[3] Zishan Zahidul Islam, Syed Mahir Tazwar, Md. Zahidul Islam, Seiichi Serikawa and Md. Atiqur Rahman Ahad, Automatic Fall Detection System of Unsupervised Elderly People Using Smartphone, IIAE International Conference on Intelligent Systems and Image Processing, 7, 2017
[4] Beauvais, S.B.; Rialle, V.; Sablier, J. Myvigi: An Android Application to Detect Fall and Wandering. In Proceedings of the sixth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, UBICOMM 2012, Barcelona, Spain, 23–28 September 2012; pp. 156–160.
[5] Zhang, T.; Wang, J.; Liu, P.; Hou, J. Fall detection by embedding an accelerometer in cellphone and using KFD algorithm. Int. J. Comput. Sci. Netw. Secur. 2006, 6, 277–284.
[6] Miluzzo, E.; Lane, N.; Eisenman, S.; Campbell, A. Cenceme—Injecting sensing presence into social networking applications. In Proceedings of the 2nd European conference on Smart sensing and context (EuroSSC '07), Lisbon, Portugal, 21–23 March 2007; pp. 1–28.
[7] Miluzzo, E.; Lane, N.; Fodor, K.; Peterson, R.; Lu, H.; Musolesi, M.; Eisenman, S.; Zheng, X.; Campbell, A. Sensing meets mobile social networks: The design, implementation and evaluation of the cenceme application. In Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems, SenSys '08, Raleigh, NC, USA, 4–7 November 2008; pp. 337–350.
[8] Glen Debard, Marc Mertens, Toon Goedeme, Tinne Tuytelaars and Bart Vanrumst Three Ways to Improve the Performance of Real-Life Camera-Based Fall Detection Systems Journal of Sensors 2017
[9] Miao Yu, Liyun Gong, Stefanos Kollias Computer vision based fall detection by a convolutional neural network ACM, 2017
[10] Philip Geismann. Georg Schneider, A Two-staged Approach to Vision-based Pedestrian Recognition Using Haar and HOG Features,IEEE, 6, 2008
[11] Fouzi Harroua , Nabil Zerroukib , Ying Suna , Amrane Houacineb,Vision-based fall detection system for improving safety of elderly people,IEEE Instrumentation and Measurement Society, 21, 2017
[12] Koldo de Miguel, Alberto Brunete, Miguel Hernando and Ernesto Gambao, Home Camera-Based Fall Detection System for the Elderly, Multidisciplinary Digital Publishing Institute (MDPI),Sensors,21, 2017
[13] Junho Ahn, Richard Han, myBlackBox: Blackbox Mobile Cloud Systems for Personalized Unusual Event Detection, Sensors, 2016, Volume 16, Issue 5, 753; doi:10.3390/s16050753, 20 pages.
[14] Ahn, J.; Han, R. Personalized behavior pattern recognition and unusual event detection for mobile users. Mobile Information Systems, 2013, 9, 99–122.
[15] Junho Ahn, James Williamson, Mike Gartrell, Richard Han, Qin Lv, and Shivakant Mishra. Supporting Healthy Grocery Shopping via Mobile Augmented Reality. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 12, 1s, Article 16 (October 2015), 24 pages.
[16] Ergashev Odiljon, Hwijune Park, Junho Ahn. Personalized Unusual Event Detection Algorithm at Smart Home via Daily Activity and Vision Pattern, The 15th 2018 International Interdisciplinary Workshop Series, 2018
[17] SOS Emergency App. https://play.google.com/store/apps/details?id=com.sosmate.
[18] Medical Alert https://play.google.com/store/apps/details?id=com.hnyapps.medicalalert
[19] Offender Locator Lite.
https://itunes.apple.com/us/app/offender-locator-lite/id322664011?mt=8
[20] Family Locator - Phone Tracker.
https://play.google.com/store/apps/details?id=com.sygic.familywhere.android
[21] K. Lin, A. Kansal, D. Lymberopoulos, and F. Zhao. Energy-accuracy aware localization for mobile devices. In Proceedings of ACM MobiSys '10, San Francisco, California, USA.
[22] J. Paek, J. Kim, and R. Govindan. Energy-efficientrate-adaptive gps-based positioning for smartphones.In MobiSys'10, June 2010.
[23] Huang J, Rathod V, Sun C, Zhu M, Korattikara A, Fathi A, Fischer I, Wojna Z, Song Y, Guadarrama S, Murphy K., Speed/accuracy trade-offs for modern convolutional object detectors. CVPR 2017, https://github.com/tensorflow/models/tree/master/research/object_detection
[24] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, 2016