Received September 7, 2017, accepted October 7, 2017, date of publication November 8, 2017, date of current version March 9, 2018.

*Digital Object Identifier 10.1109/ACCESS.2017.2771389*

# Development of Home Intelligent Fall Detection IoT System Based on Feedback Optical Flow Convolutional Neural Network

**YI-ZENG HSIEH[1] AND YU-LIN JENG[2]**
[1]Department of Electrical Engineering, National Taiwan Ocean University, Keelung 202, Taiwan
[2]Department of Information Management, Southern Taiwan University of Science and Technology, Tainan 710, Taiwan

Corresponding authors: Yi-Zeng Hsieh (yzhsieh@mail.ntou.edu.tw) and Yu-Lin Jeng (jackjeng@stust.edu.tw)

This work was supported by the Ministry of Science and Technology, Taiwan, China, under Grant 106-2511-S-019-003 and Grant 105-2511-S-218 -003 -MY2.

**ABSTRACT** Fall events are important health issues in elderly living environments such as homes. Hence, a confident and real-time video surveillance device that pays attention could better their everyday lives. We proposed an optical flow feedback convolutional neural network according to the video stream in a home environment. Our proposed model uses rule-based filters before an input convolutional layer and the recorded optical flow for supervising the optical flow of variation. Detecting human posture is a key factor, while fall events are like a falling posture. By sequencing frames of action, it is possible to recognize a fall. Our system can clearly detect the normal lying posture and lying after falling. Our proposed method can efficiently detect action motion and recognize the action posture. We compared the performance with other standard benchmark data sets and deployed our model to simulate a real-home situation, and the correct ratio achieved 82.7% and 98% separately.

**INDEX TERMS** IoT, fall detection, convolutional neural network, optical flow.

## I. INTRODUCTION

Every day the Internet of Thing (IoT) is the interaction of omnipresent sensors and machines to connect tangible and invented things among many organizations. In today's linked world, there are many techniques of connected machines, such as, 4G, Bluetooth, NFC, RF and Wi-Fi. Omnipresent calculations contrast with normal calculations in two principle areas [1]: (1) the outstanding move of calculating from normal objective calculators to mobile devices, and (2) the forward direction of the active interaction of calculating machines with their matching objectives and circumfluent frameworks, which are frequently short of distinct agent influence. A principal view of these favors can be caught by the research of Mark Weiser. In his survey [2], he proposed that "The greatest techniques are those that vanish. They braid themselves into the material of everyday life until they are invisible from it." Humans live in an extending linked and self-working community. We are exploring recording and computing services in our most individual surroundings: the home. We are exploring models of smart homes to execute health assistance for disabled persons or the elderly. Most disabled or older people would favor using non- invasive scientific devices to assist them with their everyday actions. Such methods of monitoring devices for disabled or older people to command are not received and in most states they are eliminated wholly [3]. The study and evolution have concentrated on the operation of distinct low-light scientific machines which are immediately usable [4]–[6]. Nevertheless, many surveys have concentrated on possibility methods. However, there are some constraints with probability models. First, the main grade focus on an event determined by the main grade in the opposed case. Second, the standard possibility method cannot shape into simple in a normal way [7]. A consistent distribution of possibility method on data set advanced present arbitrarily situations than simple methods. Simple methods describe the truth that, for an operator, every probable result of case is the same reasonable, while it has no fact that is availability to help any of them by requiring data. Movement detection direct to exactly recognize person's everyday movements established on a prescribing activity function [8]. It is a very important survey subject in the domain of omnipresent calculating and generally used in lot of human-orientation applications, like health and fitness surveillance system [9]–[17], helpful living devices [18]–[24], environment-based contest and relaxations [25]–[27], communal networking [28], [29] and contest following [30]–[32]

To detect movements, visible sensors are often allocated in surroundings, fixed on items, or placed on the human body to repeatedly gather sensor inspections. Afterwards, based on pre-described model detection methods, the kinds of movement are recognized at an accumulation for superior layer employments. These sensor-based methods are called normal movement detection models. They can be partitioned into three classes: (1) wearable sensor-based models [33], which use activity sensors to perceive the activities of body parts, such as [11], [12], [34]–[39]; (2) CMOS sensor based models [40], which use CMOS sensors to describe the stream series and detect movements utilizing image processing methods and which include RGB stream (e.g. [41], [42]), depth stream (e.g. [26], [43]) or RGB-D stream (e.g. [44], [45]); and (3) surroundings parameter-based models, which adopt visible sensors to gather movements from the position of used objects or variations of the surroundings, such as [23], [24], and [46]. Nevertheless, normal movement detection models get better results and are wholly approved, however they require special realizing devices and boost some interests for instance secrecy, energy reduction and allocation amount. Many existences of fall event recognition and warning methods can be categorized into three classes: wearable-based, surrounding-based, and vision-based methods. Wearable sensor-based methods usually depend on G-sensors that are joined to the user's body [47]. Surrounding sensor based fall event recognition methods utilize outer sensors inserted in the surroundings and contain pressure inductors, acoustic inductors, and EMG sensors etc. [48], [50]. With the evolution of image processing, image and stream-based models have become common in fall event detection devices [51], [52]. This type of model is non-obvious and suitable for the elderly.

In the observable fall event recognized devices, the station of the view of the seniors can be transferred to a suitable objective if a fall event is caught and warning signals are launched. Fall event recognition machines usually concentrate on a single individual. If the position has two or more people, separation and evidence devices will be adopted to focus on each human and recognize and trace them separately [53].

Since 2D images are the projection of 3D objectives, the issue of phenomenon transformation may take place in detecting fall events [54]. So as to handle the issue, this paper proposed a new fall event recognition model established on treating the silhouette in the depth image. The following is as: (1) depth raw data are pre-operated by a median filter for two objective appearances and background objectives; (2) the shape of the shifting person in depth image is performed by the reduction algorithm of background images; (3) the floor level formula is calculated by the least square algorithm and a disparity image transformed from the depth data; and (4) shape knowledge of the body in depth data is examined by a set of instant services, and the parameters of ellipses are measured to decide the orientation and situation of the user. The means of the body and the angle between the shape of the body and the lower level are measured for fall event recognition. When two surpass some thresholds, a fall event will be recognized. The most principal benefit of bio-signal dominated devices over other kinds of control devices, for instance, body forced mechanical machines, is the hands-free command of a person's objective. They supply more closed services and apparent vision [55]. Concentrating on electromyography and electroencephalography signals, a lot of possible real-world utilization of these two biological sequential data have been described, including versatile service rehabilitation, sensible assisted wheels, step production, scratching commands, and gesture-based input, etc.

The predominant algorithm for observing community position conversation employments has been surveys [56], [57], [60] [62]–[65]. The position situations are proposed to perform an approximate position requirement offered a definite environment for example appointment space, or advantage community influence. Position examples like WatchMe [62] often execute leader surveys to search the device features that should be connected within their modules. Then, research like Guide Me [60] adopt occurrence patterns to develop the knowledge expression to their customers. There are a number of common business position employments [61], [62] utilized to help social influence. Projects such as Dodgeball [62] supply a function reward for a particular case, but are finite as a strong normal intention community position module. Context communication with the user is restricted not only by the device itself, but by the user's smart devices.

The percentage of people are over the age of 65 in the European Union, China, and the United States are 30%, 30% and 20.2%, respectively. There is a large population over the age of 65 in the world [63]. The distribution of distinct healthcare design is forced, and most of the models have been developed according to the traditional clinical method in which case are analyzed and considered in per-acute situations in a clinical setting, and more and more residence-based method that is expressed to some case in their private space [63]. This residence-based method betters the attributes of a patient's existence, and is capable of being alive extensive in a well-known context not changing their main living space. In addition, the residence-based method lessens the common charge for supporting patient health services. For many elderly living alone, the hazard of falling and sustaining an injury is a major subject to handle. A fall is defined as "an occurrence which effects a person to rest unintentionally on the ground or floor or other lower level" by the World Health Organization, and the number of people aged 65 and over 70 who fall every year is about 28-35% and 32-42%, respectively [64]. Fall cases result in bodily injury and a mental influence of being scared of falling once more. Ordinary, the reliance of physical activity is less and less [65]. In addition, there are many studies describing the decreasing life anticipation of people recovering from a fall.

It is hazardous for people living alone if they are delayed help after a fall. In the home, a low-cost, inconspicuous device that is able to recognize the falling posture of the elderly could assist in lowering the occurrence of delayed support. Many models have been studied and developed for recognizing the falling posture of the elderly. Many researches contain wearable devices based on touching a switch before falling, and then utilize sensors such as G-sensors [66]–[69]. Nevertheless, wearable devices have to be worn properly and need an electric cell, and they could be easily be ignored by the patient. Concerning wearable machines needing behavior on the component of the user losing feeling after falling would avoid using. Investigations have also suggested the elderly prefer invisible sensors [70].

These devices would directly help the elderly by allowing them to keep on living independently, free of the requirement to move to accepted protection and eventually lessening the affective and economic loading for the elderly and their relatives. There are also definite social and saving influences by lessening the finance required to consider fall events.

It is also important to observe that faithful supervising schemes are favorable not only for detecting fall events, but also for calculating the living character of a person. This contains what action the elderly has, their movements in distinct sections of the building, and what motions are involved, especially basic postures like eating and personal health condition. Abnormalities of the basic living model of the elderly can indicate recognition worsening health situations, suppling the chance for instant and essential care [71].

## II. RELATED WORKS

There are numerous studies exploring image pattern models to detect fall events. Many studies extended from single charge-coupled fixed devices [72], [73] or multiple image sensors mounted [74], [75] around a space to create 3-D modeling of fore-space targets [76], [78]. Single image sensor devices depend on image regions to extract features from contours, such as bounding shell percentages. Multi image sensor devices obtain patterns such as speed from 3-D objects built on manifold contours. However, traditional image sensor-based systems incur many constraints.

Kepski and Kwolek [79], [86] proposed deriving the floor surface spontaneously. The patient's outline is recognized through the related depth appearance, which is regularly renewed. An invented case around the human was measured, and features found on the outline and distance were measured and adopted for distinct machine learning models. Thirty-five young users were adopted to calculate the models, and were gathered in view of two depth sensors.

As soon as a fall event is found, the event must be verified by the user through a voice identification model and the multi-microphone array. A model using depth sensor to analyze human silhouette and RGB sensors is presented in [80]. A parameter-based background extraction is used for deriving the features, and the human is recognized in view of skin tone images. The human is followed through the equivalent of its core and, if a huge standing movement detection is exhibited, a second engine based on the direction of the major principle of the human silhouette is used to distinguish between falling or squatting movements.

Amini *et al.* [81] mentioned a contrast with heuristic and machine learning models for fall notification with depth sensors. The heuristic algorithm is in view of the skeleton dataset, and the 3D position of the head junction are followed. A fall event is recognized by installing a threshold on the speed and G-sensor of the top junction, along with a short length between the top and the low level. The machine learning model is established using an AdaBoost classifier together with a set of weighted feeble supervised models to be a closing boosted supervised model. Just the speed and the user's top length to the low level have been thought of the machine learning methods. For both heuristic and neural network models, the information data was recorded by 11 young users. Each user portrayed six true and six negative fall event cases, which concluded by falling or sitting on the low level. The rule-based models approached a correct ratio of 95.42% for fall recognition, while the machine learning method has a less correct ratio (88.33%) because of the finite number of user cases. The model proposed in [82] uses only the pulled out silhouette data and is the best at recognizing fall events associated with weight moving events. Characteristics measured from silhouettes think of tallness, speed of the upper principle, principle direction and its varieties, and projection of the center of the body on the low level. A linear support vector machine withdrawn features and supervised neural network fall cases from non-falling cases, in which non-falling actions consist of walking, standing, sitting and sleeping. Zhang *et al.* [83] utilized the junctions of high point and body, which are properly recognized if a human is standing or sitting, and incorrectly calculated if a fall event occurs. They established a mechanical character vector of the position between pairs of junctions on distinct silhouettes, using the least and greatest values of the tallness of the human within a series of frames. They could recognize five fall events associated with movements using RGB pixels to get the human outline if the silhouette is unavailable. Dai *et al.* [84] selected hidden Markov models to create the model of temporal series of positions which form movements.

All the time sequence series from seven movements are divided into sets to withdraw in connection with movements. Eventually, a trained HMM model is developed for every movement, and the movement class proportional to the pattern reaching the maximum similarity is the detected class. Alazrai *et al.* [85] presented a sight-unlike Motion-Pose Geometric Descriptor (MPGD) calculated from the silhouette positions of junctions, which is capable of catching the movement and position of human body parts while keeping the sequential formation of the shifting parts. The recognition of fall event structures contains two group tiers. The first tier is a set of SVMs which define the condition of the person on each image. At the second tier the constraint dynamic time warping (cDTW) method is utilized to organize the entire series of conditions into falling or non-falling cases. This model performed good outcomes in the categorization of four movements.

The last, research in object recognition is steered by the achievement of region–based models [86] and proposed region convolutional neural networks (RCNN) [87]. Even though region-based CNNs have improved [87], their amount has been less because of apportioning convolutional layers through methods [88]. The newest type, Fast R-CNN [89], adopt change deep methods [90], while neglecting the time cost of region-based models. The region-based models depend on reasonable characters and economizing on consequence rules. The exclusive search method [86], one of the top common models, commercially combines high level pixels found on guided features of the low layer. Contrasted with effective recognition models [89]. Generally, [91] proposes the best concession among method property and speedup, at 0.2 seconds for one frame. However, the region-based model method spends time to compute the recognition learning methods. The individual should be concerned that most of the speedy region models of deep learning neural networks use GPUs computation, while the region-based networks adopted in advanced survey are built on the CPU. An apparent method to speed up proposed mathematics is to build it again on the GPU. This can be an efficient direction explanation, but building it again neglects the low quality of the stream recognition model and hence loses a major chance for sharing calculations.

It is good distinguishing motion features among with the time sequence and object domain dimensions must be noticed. Therefore, using filter containing series motion characters to compare with traditional convolutional filters at one frame can include location information as the input feature, and now the time sequence feature is better to share their information. Because of the time sequence deep learning convolution model, most features have adaptability of structures to describe the object motion stream data. Considering a time sequence object of structure produces more filters containing spatial and temporal features from the continuing stream and evaluating spatial and temporal feature convolution and down sampling apart from feature domain. The latest feature domain information is received by connecting temporal and spatial knowledge. To increase the efficiency of temporal and spatial knowledge networks, it is necessary to increase the networks with assisting results to calculate feedback while the super layer generates motion information and coordinates the results of a different network structure for predicting decision making.

In this paper, we proposed a purpose objective detection by an optical flow feedback-based model, calculating framework with a feature feedback based deep neural network, giving rise to a refined and efficient result when the proposed computation has more confidence from the computing time of the object recognition network. In addition, we also proposed a feature feedback mechanism scheme (FFMS) that would apportion the feature of the convolutional layers with the best object recognition models [88], [89]. By apportioning features at the test moment, the peripheral time computation was efficient. Our proposed detection was the convolution layer combined with optical flow feedback and adopted by region-based models. When the object is moving or takes some movements, our proposed method can also be used for generating region for objective by optical flows. Based on these convolutional characteristics, we built a feature feedback mechanism scheme by adding a few feedback characters to the existing extra feature layers that together recovered the object of the region boxes and detected object signs at each position on a normal network. Hence, the FFMS layer was a type layer of full convolution [92], and the training scheme was also learned for the purpose for producing a moving object detected model. FFMS are built to effectively forecast boundary models with a varying scope of degree and view rates. To compare with common models [88], [89], [93], [94], [98] that utilize optical flow feature or optical flow of filters, the feature is called the optical flow feedback boundary that helps with moving object detection at multiple optical flow statistics. The proposed method is as type of optical flow feedback from sequential frames, which avoids listing the time consumption computation for stream video to get filters with optical flow statistics. Our proposed method executes better performance, while training step and testing step adopting feedback optical frames take advantage of computing efficiency. To combine FFMS with objective detection model [89], the proposed learning method chooses among fine-feedback features for the object detection task, at that time re-tuning the method for sharing features. Our mechanism immediately centralizes and generates a basic machine learning method with the convolutional layer to consider the optical feedback. From the result our proposed model generated object movement recognition with more confidence than the existing standard baseline comparative algorithms. In addition, our proposed scheme yielded a more efficient computing cost. The application was implemented in human fall event detection using RGB cameras in an IoT smart home environment.

## III. FEEDBACK OPTICAL FLOW CONVOLUTIONAL NEURAL NETWORK

The proposed motion object recognition model consists of two learning phases. The first detector method is based on interest of point statistics histograms to retrieve efficient features as the input to the convolutional layer. This step is a backpropagation computation connection layer to determine the motion object boundary, and the next scheme adopts the found object's boundary. Our proposed method like Figure 1 is an independent general learning method for recognizing object information.

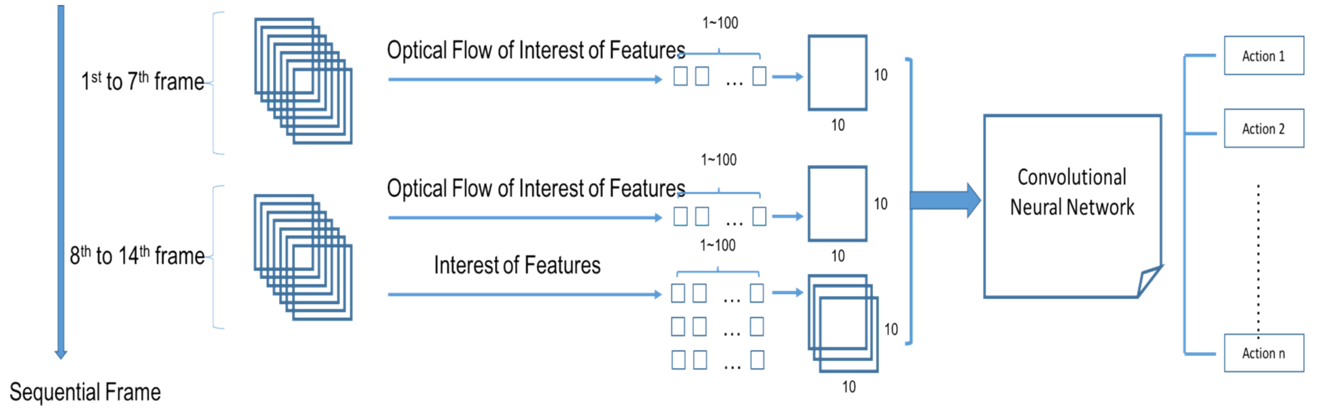



**FIGURE 1.** Feedback optical flow convolutional neural network.

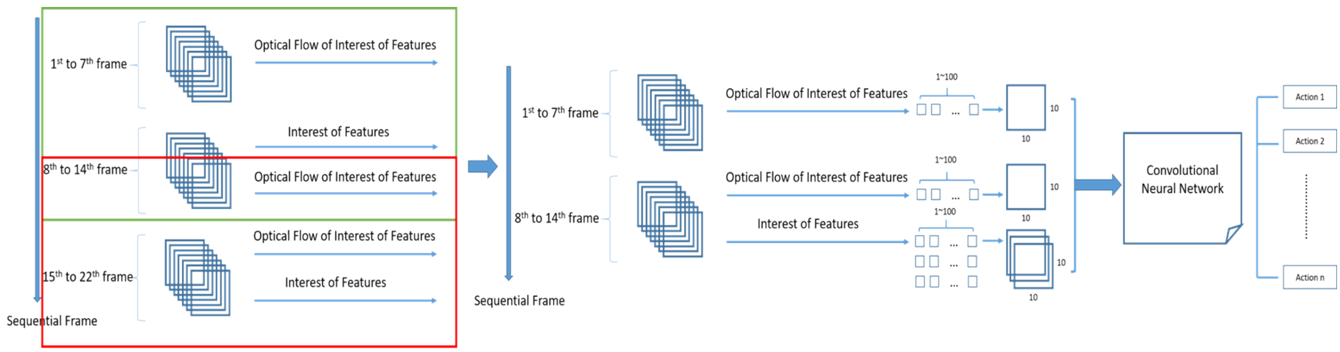



**FIGURE 2.** Feature feedback mechanism scheme.

### A. FEATURE FEEDBACK MECHANISM SCHEME (FFMS)

The feature feedback mechanism scheme (FFMS) uses interest of optical flow vector to calculate statistic Euclidean distance of histogram and to choose the better points and then boundary is according to optical flow features to retrieve the object. This paper proposed using the optical boundary region to retrieve the meaningful points as input into a fully back-propagation network [93]. Hence, we adopt the optical flow feedback scheme to identify the object when it is moving. Our model built the sharing mechanism to tell the next frame the same object moving information such as Figure2.

In order to provide optical flow boundary goals, the small boundary is ignored when the object is a human, and hence the temporal information can detect human movement based on the temporal and spatial information when the former frame information is recorded. FFMS takes advantage of the n former optical flow points as part of the input as temporal information into a fully connected network. Therefore, the input features combine with the interest of points, the optical flow of the present frame and the optical flow of the next frame. Our proposed feature type is divided into two types of retrieved information. The convolutional layer is one type, and our proposed does not consume additional memory space to train the other neural network. The retrieved features for each RGB channel list are 100, 100 and 100 respectively. The first 100 features are the interest of points, the second 100 features are the optical flow of the interest of points, and the final 100 feature is the optical flow of the former frame; therefore, the total dimension is $100\times3+100+$ 100, as shown in fig.2. Additionally, our proposed detection introduced the large motion detector like fall events, in which sudden movement must be considered and the large motion may appear at the same time, for example, persons walking in the field. To solve this problem, our proposed method adopted the rule based mechanism when the continuous frame from the variation of the cross product of two optical flow exceeded the threshold. Our proposed method considered this situation first to avoid unexpected events, for instance, fall events at home.

### B. TWO STEPS RULE BASED MOTION DETECTION

Because of 3DCNN [97], [99] and the dimensions considering RGB and optical flow data, its computing cost is expensive. However, its correct ratio in video stream for human motion detection is as good enough as a normal training model. Consequently, the benefits of 3DCNN is adopted without exploring the training weight mechanism. Our mechanism is like a 3DCNN structure, but we adopted the fast region based CNN retrieving feature method. First, we used the fast region CNN retrieving

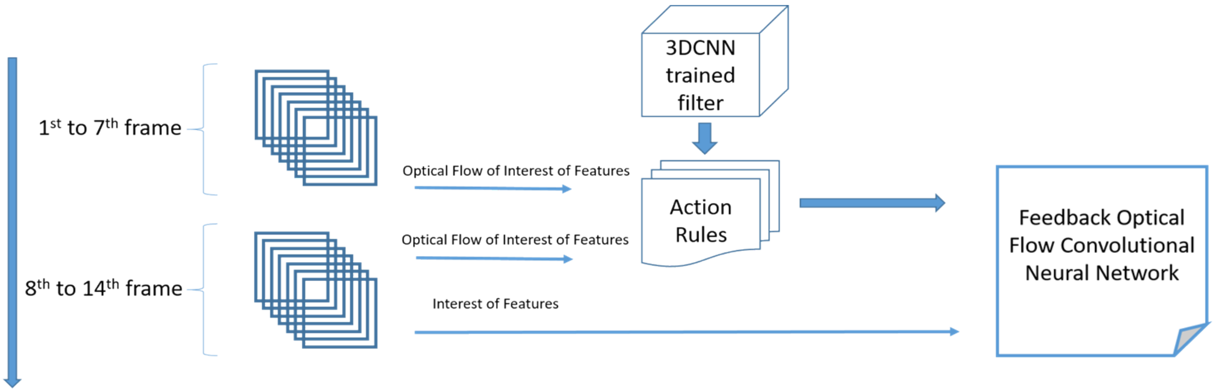



**FIGURE 3.** Features action rules.

**TABLE 1.** The correct ratio of action recognition set.

| Algorithm | boxing | handclapping | handwaving | jogging | running | walking | Average |
|---|---|---|---|---|---|---|---|
| [97] | 90 | 94 | 97 | 84 | 79 | 97 | 90.2 |
| [96] | 97.9 | 59.7 | 73.6 | 60.4 | 54.9 | 83.8 | 71.7 |
| [100] | 93 | 77 | 85 | 57 | 85 | 90 | 81.2 |
| [101] | 98 | 86 | 93 | 53 | 88 | 82 | 83.3 |
| [95] | 92 | 98 | 92 | 85 | 87 | 96 | 91.7 |
| [102] | - | - | - | - | - | - | 92.7 |
| Our method (ten-folds) | 92.5±2.5 | 91.1±3.8 | 97.7±2.5 | 90.2±1.4 | 86.1±3.1 | 98.3±0.5 | 92.65±2.3 |

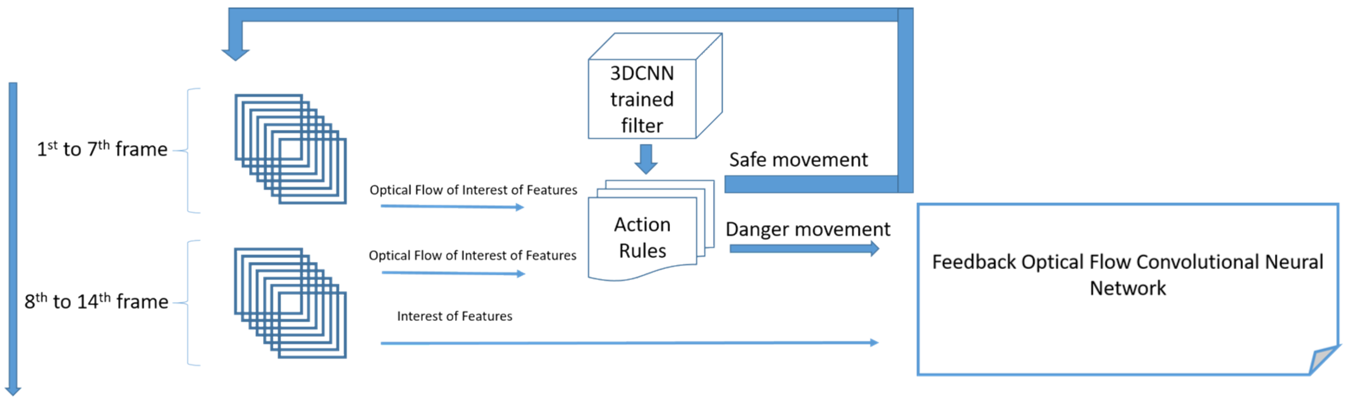



**FIGURE 4.** Falling event detection with action rules based system.

feature model to build the interest of point reference information. Second, the retrieved information was adopted as the spatial information of 3DCNN object detection, and the original convolutional layer optical flow input of 3DCNN was as the original. Thus, the deep neural network detecting object of our proposed method considered fast region CNN [89].

For 3DCNN, our method utilized the temporal features. Boundary-based models are time expensive, and the fast region CNN can solve this problem. In general, the model trained 3DCNN and fast region CNN separately, we adopted the trained parameters of optical flow information of 3DCNN, and the spatial part was gathered from the fast region CNN. Hence, our model did not change the original structure. Consequently, selection of the useful parameter information was considered. Our method had two ways to discuss this problem.

1) Choosing well-known parameters: The 3DCNN used the temporal optical flow information chosen to initially learn well-known knowledge for our proposed method. Hence, the trained process of the network did fine-tune the parameter by 3DCNN. Then, the spatial

information was found by the fast region CNN and the initial spatial features of our model.

2) Motion rule-based detection: In order to satisfy specific motion events, our model considered movement that is dangerous at home, such as fall events. In the training iteration, before inputting into the convolutional layer, the Euclidean distance of two consequences images must be considered. To enforce some dangerous situations at home, the motion event must be filtered by the rule based motion. Therefore, our method adopted the 3DCNN temporal information to construct the rule-based motion event. The backpropagation method was also used to train our deep network, and it was simplified to train the model and speed up the performance when a dangerous situation took place. Then, the detector operator utilized the fast region CNN trained parameters to consider the detected object's position.

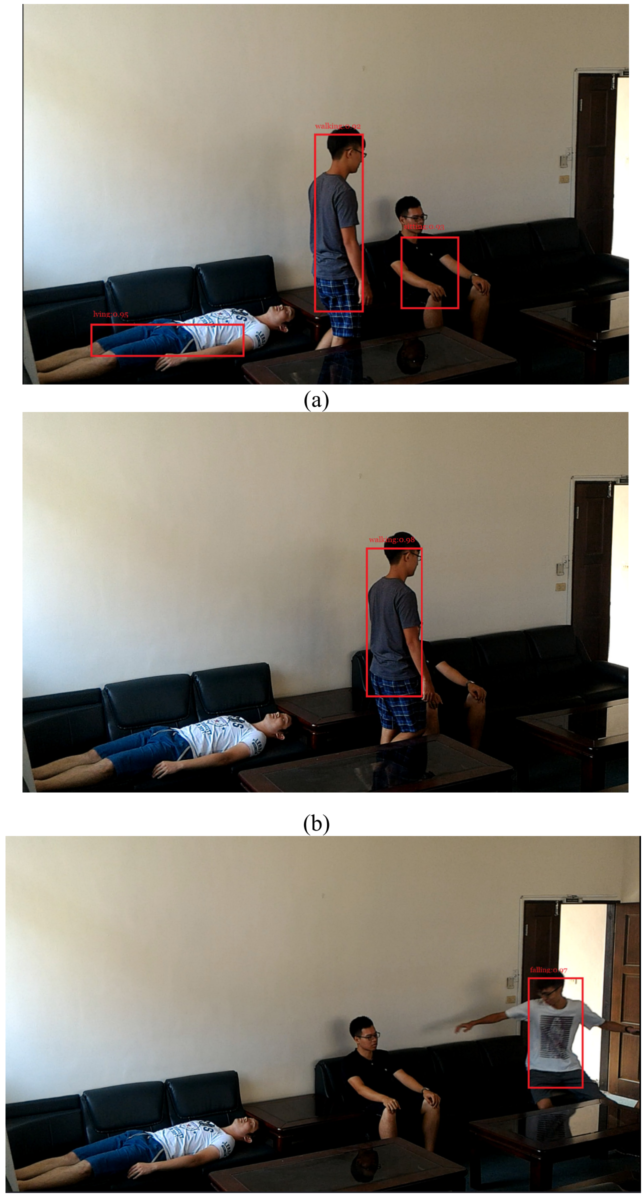
## IV. EXPERIMENT

Our proposed model tested the KTH [96] dataset to evaluate the performance for detecting action motions by adopting monitoring stream videos. At the same time, we also simulated a real life environment to check the proposed method according to real life video streams.

### A. ACTION RECOGNITION ON THE BENCHMARK DATA

In order to compare action recognition with 3DCNN and our model, the KTH data was adopted to calculate the motion detection and compare performance. The data consisted of six action types and there were 25 subjects. First, the HMAX model was built in general, and our proposed method also adopted nine continuous frames as input to retrieve the foreground objects [95]. The proposed method resembled 3DCNN in that less memory was used, as each frame reduced the resolutions from the input stream to $80 \times 60$. We used a similar architecture as shown in Figure 3, with the sizes of kernels and the number of feature maps in each layer modified to consider the $80 \times 60 \times 14$ inputs. First, our proposed method retrieved the interest of features in each frame, and then chose the 100 features in these frames. Next, the 100 features of the optical flow in first seven frames was calculated to adopt and also the $8^{th}$ to $14^{th}$ were taken. The adopted three convolutional layers used kernels sized $3 \times 3$. There were classes, including boxing, handclapping, hand waving, jogging, running and walking. From comparison with [95], the experiment also arbitrarily adopted 16 subjects for training the data, while the other nine subjects were used for testing. In ten-folds of the recognition performance is recorded in Table 1, and this was also compared with some of the models in ten-folds. The average accuracy of the proposed model was $92.65 \pm 2.3$ percent, as shown in Table 1.

Our model utilized the kmeans method to cluster the 3DCNN motion and generate the temporal rules, as shown in fig.3. Hence, our model adopted these motion rules to select the main action without computing each frame. The same parameter was as in [95]. Our proposed model utilized 16 subjects for training data and the other nine subjects for testing. The ten folded experiment was used, and the performance with other methods is listed in Table 1. Our proposed method reached 82.7 percent for final results and had better time computation. In addition, the comparison with the other models in Table 1 adopted distinct training/testing ten folded methods.



(a) (b) (c)

**FIGURE 5.** Real life posture detection.

### B. REAL WORLD FALL EVENT EXPERIMENTS

It is hazardous for people living alone if they are delayed help after falling. In the home, a low-cost, inconspicuous device that can recognize the falling posture of the elderly could assist in lowering the occurrence of delayed support. Many models have been studied and developed for recognizing the falling posture of the elderly. Older adults living alone are at great risk of delayed assistance following a fall. A low-cost, unobtrusive system capable of automatically detecting falls in the homes of older adults could help significantly reduce the incidence of delayed assistance after a fall. Falls are a major health problem in the elderly population. Therefore, a dedicated monitoring system is highly desirable to improve independent living. This experiment presented a video-based fall detection system in an indoor environment using our proposed model. The proper recognition of human posture is significant in order to detect fall events such as Figure 4. Based on our model, rule-based knowledge was adopted in sequence frames. Detecting posture is the main problem in distinguishing fall or non-fall events. A posture that immediately changes to a lying posture could represent the occurrence of a fall event. Sleeping is different from the falling posture, because the falling posture is an immediate and large movement. Our proposed model, the front rule layer detected different posture situations if the posture is falling or not. A video stream was used and the image was RGB based. The front object was subtracted from the background image, because the home environment does not change very often, so the background image was updated every three minutes to lessen the computation cost. The average for processing frames was 25 frames/second. The result is shown in Figure 5.

## V. CONCLUSION

It is very hazardous for elderly people living alone if they are delayed help after falling. In the home, a low-cost, inconspicuous device that can recognize the falling posture of elderly patients could lower the occurrence of delayed support. We proposed the fall event detection using an action rules-based system. Our proposed model can achieve immediate monitoring of fall events to prevent dangerous situations in the home environment.

## REFERENCES

[1] S. K. Das, A. Bhattacharya, A. Roy, and A. Misra, "Managing location in 'Universal' location-aware computing," in *Handbook of Wireless Internet*, B. Furht and M. Ilyas, Eds. Boca Raton, FL, USA: CRC Press, 2003, ch. 17, pp. 407–425.

[2] M. Weiser, "The computer for the 21st century," *SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 3, no. 3, pp. 3–11, Jul. 1999.

[3] (2004). *Marilyn Cash, At Home With AT (Assistive Technology), Research Report*. [Online]. Available: http://www.dementia-voice.org.uk/Projects/At Home with AT main.pdf

[4] C. Nugent *et al.*, "Home based assistive technologies for people with mild dementia," in *Proc. 5th Int. Conf. Smart Homes Health Telematics*, 2007, pp. 63–69.

[5] R. Orpwood, C. Gibbs, T. Adlam, R. Faulkner, and D. Meegahawatte, "The design of smart homes for people with dementia—User-interface aspects," *Univ. Access Inf. Soc.*, vol. 4, no. 2, pp. 156–164, 2005.

[6] D. Dubois, A. HadjAli, and H. Prade, "A possibility theory-based approach to the handling of uncertain relations between temporal points," *Int. J. Intell. Syst.*, vol. 22, no. 2, pp. 157–179, 2007.

[7] D. Dubois and H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. New York, NY, USA: Plenum, 1988.

[8] E. Kim, S. Helal, and D. Cook, "Human activity recognition and pattern discovery," *IEEE Pervasive Comput.*, vol. 9, no. 1, pp. 48–53, Jan. 2010.

[9] Q. Li, G. Zhou, and J. A. Stankovic, "Accurate, fast fall detection using posture and context information," in *Proc. 6th ACM Int. Conf. Embedded Netw. Sensor Syst.*, 2008, pp. 443–444.

[10] Q. Li, J. A. Stankovic, M. A. Hanson, A. Barth, J. Lach, and G. Zhou, "Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information," in *Proc. 6th Int. Workshop Wearable Implant. Body Sensor Netw.*, 2009, pp. 138–143.

[11] T. Hao, G. Xing, and G. Zhou, "iSleep: Unobtrusive sleep quality monitoring using smartphones," in *Proc. 11th ACM Conf. Embedded Netw. Sensor Syst.*, 2013, Art. no. 4.

[12] X. Qi, M. Keally, G. Zhou, Y. Li, and Z. Ren, "AdaSense: Adapting sampling rates for activity recognition in body sensor networks," in *Proc. IEEE 19th Real-Time Embedded Technol. Appl. Symp.*, Apr. 2013, pp. 163–172.

[13] Y. Tang, S. Wang, Y. Chen, and Z. Chen, "PPCare: A personal and pervasive health care system for the elderly," in *Proc. 9th Int. Conf. Ubiquitous Intell. Comput.*, Sep. 2012, pp. 935–939.

[14] L. Hu, Y. Chen, S. Wang, and Z. Chen, "b-COELM: A fast, lightweight and accurate activity recognition model for mini-wearable devices," *Pervasive Mobile Comput.*, vol. 15, pp. 200–214, Dec. 2014.

[15] T. V. Duong, H. H. Bui, D. Q. Phung, and S. Venkatesh, "Activity recognition and abnormality detection with the switching hidden semi-Markov model," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 838–845.

[16] Y. Jia, "Diatetic and exercise therapy against diabetes mellitus," in *Proc. 2nd Int. Conf. Intell. Netw. Intell. Syst.*, Nov. 2009, pp. 693–696.

[17] J. Yin, Q. Yang, and J. J. Pan, "Sensor-based abnormal human-activity detection," *IEEE Trans. Knowl. Data Eng.*, vol. 20, no. 8, pp. 1082–1090, Aug. 2008.

[18] M. Keally, G. Zhou, G. Xing, J. Wu, and A. Pyles, "PBN: Towards practical activity recognition using smartphone-based body sensor networks," in *Proc. 9th ACM Conf. Embedded Netw. Sensor Syst.*, 2011, pp. 246–259.

[19] M. Keally, G. Zhou, G. Xing, and J. Wu, "Remora: Sensing resource sharing among smartphone-based body sensor networks," in *Proc. IEEE/ACM 21st Int. Symp. Quality Service*, Jun. 2013, pp. 1–10.

[20] S. Wang, Y. Chen, and Z. Chen, "Recognizing transportation mode on mobile phone using probability fusion of extreme learning machines," *Int. J. Uncertain. Fuzz. Knowl. Based Syst.*, vol. 21, pp. 13–22, Dec. 2013.

[21] L. Hu, Y. Chen, S. Wang, and L. Jia, "A nonintrusive and single-point infrastructure-mediated sensing approach for water-use activity recognition," in *Proc. 11th IEEE/IFIP Int. Conf. Embedded Ubiquitous Comput.*, Nov. 2013, pp. 2120–2126.

[22] T. L. M. van Kasteren, G. Englebienne, and B. J. A. Kröse, "An activity monitoring system for elderly care using generative and discriminative models," *J. Pers. Ubiquitous Comput.*, vol. 14, no. 6, pp. 489–498, 2010.

[23] M. Buettner, R. Prasad, M. Philipose, and D. Wetherall, "Recognizing daily activities with RFID-based sensors," in *Proc. 11th Int. Conf. Ubiquitous Comput.*, 2009, pp. 51–60.

[24] P. Hevesi, S. Wille, G. Pirkl, N. Wehn, and P. Lukowicz, "Monitoring household activities and user location with a cheap, unobtrusive thermal sensor array," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2014, pp. 141–145.

[25] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland, "Pfinder: Real-time tracking of the human body," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 780–785, Jul. 1997.

[26] J. Shotton *et al.*, "Real-time human pose recognition in parts from single depth images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 1297–1304.

[27] F. Huo, E. Hendriks, P. Paclik, and A. H. J. Oomes, "Markerless human motion capture and pose recognition," in *Proc. 10th IEEE Workshop Image Anal. Multimedia Interact. Services*, May 2009, pp. 13–16.

[28] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 935–942.

[29] A. T. Campbell *et al.*, "CenceMe: Injecting sensing presence into social network applications using mobile phones (Demo Abstract)," in *Proc. 9th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2008, pp. 1–2.

[30] Y. Ke, R. Sukthankar, and M. Hebert, "Spatio-temporal shape and flow correlation for action recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.

[31] W.-L. Lu and J. J. Little, "Simultaneous tracking and action recognition using the PCA-HOG descriptor," in *Proc. 3rd Can. Conf. Comput. Robot Vis.*, 2006, p. 6.

[32] Y. Ohgi, M. Yasumura, H. Ichikawa, C. Miyaji, "Analysis of stroke technique using acceleration sensor IC in freestyle swimming," in *Proc. Eng. SPORT*, 2000, pp. 503–511.

[33] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1192–1209, 3rd Quart., 2013.

[34] H.-T. Cheng, F.-T. Sun, M. Griss, P. Davis, J. Li, and D. You, "NuActiv: Recognizing unseen new activities using semantic attribute-based learning," in *Proc. 11th Annu. Int. Conf. Mobile Syst., Appl., Services*, 2013, pp. 361–374.

[35] K. Aminian, F. Dadashi, B. Mariani, C. Lenoble-Hoskovec, B. Santos-Eggimann, and C. J. Büla, "Gait analysis using shoe-worn inertial sensors: How is foot clearance related to walking speed?" in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2014, pp. 481–485.

[36] T. Maekawa, Y. Kishino, Y. Sakurai, and T. Suyama, "Activity recognition with hand-worn magnetic sensors," *Pers. Ubiquitous Comput.*, vol. 17, no. 6, pp. 1085–1094, 2013.

[37] O. Yürüten, J. Zhang, and P. H. Z. Pu, "Predictors of life satisfaction based on daily activities from mobile sensor data," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 2014, pp. 497–500.

[38] D. Schuldhaus, H. Leutheuser, and B. M. Eskofier, "Classification of daily life activities by decision level fusion of inertial sensor data," in *Proc. 8th Int. Conf. Body Area Netw.*, 2013, pp. 77–82.

[39] F. Mokaya, B. Nguyen, C. Kuo, Q. Jacobson, A. Rowe, and P. Zhang, "MARS: A muscle activity recognition system enabling self-configuring musculoskeletal sensor networks," in *Proc. 12th Int. Conf. Inf. Process. Sensor Netw.*, Apr. 2013, pp. 191–202.

[40] S.-R. Ke, H. L. U. Thuc, Y.-J. Lee, J.-N. Hwang, J.-H. Yoo, and K.-H. Choi, "A review on video-based human activity recognition," *Computers*, vol. 2, no. 2, pp. 88–131, 2013.

[41] X. Ren and C. Gu, "Figure-ground segmentation improves handled object recognition in egocentric video," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 3137–3144.

[42] R. Messing, C. Pal, and H. Kautz, "Activity recognition using the velocity histories of tracked keypoints," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 104–111.

[43] I. Oikonomidis, N. Kyriazis, and A. A. Argyros, "Efficient model-based 3D tracking of hand articulations using Kinect," in *Proc. 22nd Brit. Mach. Vis. Conf.*, 2011, pp. 1–11.

[44] K. Lai, L. Bo, X. Ren, and D. Fox, "A scalable tree-based approach for joint object and pose recognition," in *Proc. 25th Conf. Artif. Intell.*, 2011, pp. 1474–1480.

[45] J. Lei, X. Ren, and D. Fox, "Fine-grained kitchen activity recognition using RGB-D," in *Proc. ACM Conf. Ubiquitous Comput.*, 2012, pp. 208–211.

[46] P. Rashidi and D. J. Cook, "Mining sensor streams for discovering human activity patterns over time," in *Proc. IEEE Int. Conf. Data Mining*, Dec. 2010, pp. 431–440.

[47] V. R. L. Shen, H.-Y. Lai, and A.-F. Lai, "The implementation of a smartphone-based fall detection system using a high-level fuzzy Petri net," *Appl. Soft Comput. J.*, vol. 26, pp. 390–400, Jan. 2015.

[48] Y. Zigel, D. Litvak, and I. Gannot, "A Method for automatic fall detection of elderly people using floor vibrations and sound—Proof of concept on human mimicking doll falls," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 12, pp. 2858–2867, Dec. 2009.

[49] J. Cheng, X. Chen, and M. Shen, "A framework for daily activity monitoring and fall detection based on surface electromyography and accelerometer signals," *IEEE J. Biomed. Health Informat.*, vol. 17, no. 1, pp. 38–45, Jan. 2013.

[50] N. El-Bendary, Q. Tan, F. C. Pivot, and A. Lam, "Fall detection and prevention for the elderly: A review of trends and challenges," *Int. J. Smart Sens. Intell. Syst.* , vol. 6, no. 3, pp. 1230–1266, 2013.

[51] W. Feng, R. Liu, and M. Zhu, "Fall detection for elderly person care in a vision-based home surveillance environment using a monocular camera," *Signal Image Video Process.*, vol. 8, no. 6, pp. 1129–1138, 2014.

[52] Y. T. Liao, C.-L. Huang, and S.-C. Hsu, "Slip and fall event detection using Bayesian Belief Network," *Pattern Recognit.*, vol. 45, no. 1, pp. 24–32, 2012.

[53] L. Yang, Y. Ren, H. Hu, and B. Tian, "New fast fall detection method based on spatio-temporal context tracking of head by using depth images," *Sensors*, vol. 15, no. 9, pp. 23004–23019, 2015.

[54] M. A. Oskoei and H. Hu, "Myoelectric control systems—A survey," *Biomed. Signal Process. Control*, vol. 2, no. 4, pp. 275–294, 2007.

[55] C. Greenhalgh, S. Izadi, J. Mathrick, J. Humble, and I. Taylor, "ECT: A toolkit to support rapid construction of ubicomp environments," in *Proc. Ubicomp*, Nottingham, U.K., 2004, pp. 207–234.

[56] R. E. Grinter, W. K. Edwards, M. W. Newman, and N. Ducheneaut, "The work to make a home network work," in *Proc. ECSCW*, 2005, pp. 97–119.

[57] H. Hutchinson *et al.*, "Technology probes: Inspiring design for and with families," in *Proc. CHI*, Ft. Lauderdale, FL, USA, 2003, pp. 17–24.

[58] S. S. Intille, K. Larson, J. S. Beaudin, J. Nawyn, E. M. Tapia, and P. Kaushik, "A living laboratory for the design and evaluation of ubiquitous computing technologies," in *Proc. CHI*, Portland, OR, USA, 2005, pp. 1941–1944.

[59] A. LaMarca *et al.*, "Place lab: Device positioning using radio beacons in the wild," in *Pervasive Computing* (Lecture Notes in Computer Science), vol. 3468. Munich, Germany: Springer-Verlag, 2005, pp. 421–486.

[60] D. A. Norman, "Home theater: Not ready for prime time," *IEEE Comput.*, vol. 35, no. 6, pp. 100–102, Jun. 2002.

[61] J. O'Brien, T. Rodden, M. Rouncefield, and J. Hughes, "At home with the technology: An ethnographic study of a set-top-box trial," *ACM Trans Comput. Human Interactions*, vol. 6, no. 3, pp. 282–308, 1999.

[62] S. Brand, *How Buildings Learn*. New York, NY, USA: Viking Penguin, 1994.

[63] Z. Pang, L. Zheng, J. Tian, S. Kao-Walter, E. Dubrova, and Q. Chen, "Design of a terminal solution for integration of in-home health care devices and services towards the Internet-of-Things," *Enterprise Inf. Syst.*, vol. 9, no. 1, pp. 86–116, Jan. 2015.

[64] *WHO Global Report on Falls Prevention in Older Age*, World Health Organization, Geneva, Switzerland, 2008.

[65] M. Terroso, N. Rosa, A. T. Marques, and R. Simoes, "Physical consequences of falls in the elderly: A literature review from 1995 to 2010," *Eur. Rev. Aging Phys. Activity*, vol. 11, no. 1, pp. 51–59, 2014.

[66] N. Noury *et al.*, "Fall detection—Principles and methods," in *Proc. IEEE 29th Ann. Int. Conf. Eng. Med. Biol. Soc.*, Aug. 2007, pp. 1663–1666.

[67] A. K. Bourke, P. W.J. van de Ven, A. E. Chaya, G. M. OLaighin, and J. Nelson, "Testing of a long-term fall detection system incorporated into a custom vest for the elderly," in *Proc. IEEE 30th Ann. Int. Eng. Med. Biol. Soc. Conf.*, Aug. 2008, pp. 2844–2847.

[68] G. Wu and S. Xue, "Portable preimpact fall detector with inertial sensors," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 2, pp. 178–183, Apr. 2008.

[69] M. Prado-Velasco, M. G. del Rio-Cidoncha, and R. Ortiz-Marin, "The inescapable smart impact detection system (ISIS): An ubiquitous and personalized fall detector based on a distributed 'divide and conquer strategy,'" in *Proc. IEEE 30th Annu. Int. Eng. Med. Biol. Soc. Conf.*, Aug. 2008, pp. 3332–3335.

[70] G. Demiris *et al.*, "Older adults' attitudes towards and perceptions of 'smart home' technologies: A pilot study," *Med. Inform. Internet Med.*, vol. 29, no. 2, pp. 87–94, 2004.

[71] A. König *et al.*, "Validation of an automatic video monitoring system for the detection of instrumental activities of daily living in dementia patients," *J. Alzheimer's Dis.*, vol. 44, no. 2, pp. 675–685, 2015.

[72] T. R. Bennett, J. Wu, N. Kehtarnavaz, and R. Jafari, "Inertial measurement unit-based wearable computers for assisted living applications: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 28–35, Mar. 2016.

[73] M. Mubashir, L. Shao, and L. Seed, "A survey on fall detection: Principles and approaches," *Neurocomputing*, vol. 100, pp. 144–152, Jan. 2013.

[74] Z. Zhang, C. Conly, and V. Athitsos, "A survey on vision-based fall detection," in *Proc. 8th ACM Int. Conf. Pervasive Technol. Rel. Assistive Environ.*, 2015, pp. 46:1–46:7.

[75] G. Koshmak, A. Loutfi, and M. Linden, "Challenges and issues in multisensor fusion approach for fall detection: Review paper," *Hindawi J. Sensors*, vol. 2016, Aug. 2016, Art. no. 6931789.

[76] K. Chaccour, R. Darazi, A. H. El Hassani, and E. Andrès, "From fall detection to fall prevention: A generic classification of fall-related systems," *IEEE Sensors J.*, vol. 17, no. 3, pp. 812–822, Feb. 2017.

[77] B. Y. Su, K. C. Ho, M. J. Rantz, and M. Skubic, "Doppler radar fall activity detection using the wavelet transform," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 3, pp. 865–875, Mar. 2015.

[78] I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, and Y. Bengio. (2016). "Quantized neural networks: Training neural networks with low precision weights and activations." [Online]. Available: https://arxiv.org/abs/1609.07061

[79] M. Kepski and B. Kwolek, "Unobtrusive fall detection at home using Kinect sensor," in *Computer Analysis of Images and Patterns*, R. Wilson, E. Hancock, A. Bors, W. Smith, Eds. Berlin, Germany: Springer, 2013, pp. 457–464.

[80] S.-W. Yang and S.-K. Lin, "Fall detection for multiple pedestrians using depth image processing technique," *Comput. Methos Programs Biomed.*, vol. 114, no. 2, pp. 172–182, 2014.
[81] A. Amini, K. Banitsas, and J. Cosmas, "A comparison between heuristic and machine learning techniques in fall detection using Kinect v2," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, May 2016, pp. 1–6.
[82] A. Davari, T. Aydin, and T. Erdem, "Automatic fall detection for elderly by using features extracted from skeletal data," in *Proc. Int. Conf. Electron., Comput. Comput. (ICECCO)*, Nov. 2013, pp. 127–130.
[83] C. Zhang, Y. Tian, and E. Capezuti, "Privacy preserving automatic fall detection for elderly using RGBD cameras," in *Proc. Int. Conf. Comput. Helping People Special Needs*, 2012, pp. 625–633.
[84] X. Dai, M. Wu, B. Davidson, M. Mahoor, and J. Zhang, "Image-based fall detection with human posture sequence modeling," in *Proc. IEEE Int. Conf. Healthcare Informat. (ICHI)*, Sep. 2013, pp. 376–381.
[85] R. Alazrai, A. Zmily, and Y. Mowafi, "Fall detection for elderly using anatomical-plane-based representation," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 5916–5919.
[86] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders, "Selective search for object recognition," *Int. J. Comput. Vis.*, vol. 104, no. 2, pp. 154–171, Apr. 2013.
[87] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587.
[88] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in *Proc. 13th Eur. Conf. Comput. Vis.*, 2014, pp. 346–361.
[89] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 1440–1448.
[90] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent.*, 2015, pp. 1–29.
[91] C. L. Zitnick and P. Dollár, "Edge boxes: Locating object proposals from edges," in *Proc. 13th Eur. Conf. Comput. Vis.*, 2014, pp. 391–405.
[92] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3431–3440.
[93] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.
[94] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "OverFeat: Integrated recognition, localization and detection using convolutional networks," in *Proc. Int. Conf. Learn. Represent.*, 2014, pp. 1–16.
[95] H. Jhuang, T. Serre, L. Wolf, and T. Poggio, "A biologically inspired system for action recognition," in *Proc. 11th IEEE Int.Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.
[96] C. Schuldt, I. Laptev, and B. Caputo, "Recognizing human actions: A local SVM approach," in *Proc. 17th Int. Conf. Pattern Recognit.*, Aug. 2004, pp. 32–36.
[97] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, Jan. 2013.
[98] E. Cippitelli, F. Fioranelli, E. Gambi, and S. Spinsante, "Radar and RGB-depth sensors for fall detection: A review," *IEEE Sensors J.*, vol. 17, no. 12, pp. 3585–3604, Jun. 2017.
[99] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
[100] P. Dollár, V. Rabaud, G. Cottrell, and S. Belongie, "Behavior recognition via sparse spatio-temporal features," in *Proc. IEEE Int. Workshop Vis. Surveill. Perform. Eval. Tracking Surveill.*, 2005, pp. 65–72.
[101] J. C. Niebles, H. Wang, and L. Fei-Fei, "Unsupervised learning of human action categories using spatial-temporal words," *Int. J. Comput. Vis.*, vol. 79, no. 3, pp. 299–318, 2008.
[102] K. Schindler and L. van Gool, "Action snippets: How many frames does human action recognition require?" in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.


**YI-ZENG HSIEH** received the B.S., M.S., and Ph.D. degrees in computer science and information engineering from National Central University, Taoyuan, Taiwan, in 2004, 2006, and 2012, respectively. He is currently an Assistant Professor of electrical engineering with National Taiwan Ocean University. His current research interests include neural networks, pattern recognition, image processing, and machine learning.

**YU-LIN JENG** is currently an Assistant Professor at the Department of Information Management, Southern Taiwan University of Science and Technology. He is actively involved in industrial projects and his research interests include e-learning system, cloud software applications, and mobile software development.