ACTIVITY RECOGNITION AND FALL DETECTION IN ELDERLY PEOPLE

Riya Dichwalkar [1], Shreya Oak [1], Tania Rajabally [1], and Dhananjay Kalbande [1]

[1]Sardar Patel Institute of Technology, Mumbai

***Abstract*—With the advancement in technology, personalised monitoring is emerging. During the COVID-19 pandemic, the old age homes have the most susceptible population. We have used activity recognition to detect abnormal activities which also include symptoms of COVID-19 like sneezing, fatigue and headache.**

Elderly people are prone to falls which can be fatal or cause injuries. Activity recognition and fall detection can be used to keep a tab on them. Generally senior citizens live alone or in care homes. Hence if they are constantly monitored, adequate help can be provided.

In this paper, we propose a system with activity recognition and fall detection. OpenCV is used for activity recognition and Markov model for the sequence of activities. The probability of that sequence of activities is predicted. If it is below a threshold, it is considered as an abnormal activity. A Single Shot Detector model is implemented for fall detection.

***Keywords*—Elderly People, Activity Detection, Fall Detection, Markov Model**

I. Introduction

Nowadays most of the elderly people live alone, hence monitoring their daily activities is imperative for their family. With a busy schedule, it is not always possible to be with the elderly people at all times. Hence remote-monitoring is preferred by families. Using technology they can track the activity of elderly people, and be alerted incase of abnormal activities such as falling down. Such kind of monitoring helps elderly people to stay safe with little care. It also reduces the hospital admissions due to a fall or any deleterious accident.

In the current day scenario, with the outbreak of the deadly Coronavirus, many elderly people who are living alone amidst this lockdown have limited access to facilities such as medical help. In this situation it is of utmost importance that we monitor their activities, and check for any abnormalities in it. The elderly people generally have a lower immunity and are hence easy targets of the novel coronavirus. Reaching out to each elderly person and checking if they have any symptoms will require huge skilled manpower which is a tough job. Automating this process as much as possible with least interaction among humans can be done using technology. Certain symptoms which can be recognized by a video can be used to monitor if the elderly are showing any symptoms repeatedly and then a skilled individual can be sent to only those places for testing or medical assistance. This will help testing only those who are showing symptoms and thus drastically reduce human intervention, and making it possible to quickly track down all the cases. Hence, technology plays a role to monitor their activities.

A surveillance camera will capture the video feed of the activities of the elderly people. Then these videos will be fed in our two models. The first model is for activity recognition. The activity sequences recognized are sent as input to the Markov Model(HMM). Here the probability of the activities occurring is compared with our threshold. This comparison will then reveal whether abnormal activity was observed, and an alert will be sent to the concerned family members or authorities. Our second model detects falls in elderly people.

The second leading cause of accidental deaths are falls. Above the age of 65, one is more prone to falling [1]. In dotage, the risk for fatal accidents increases. It is due to several factors. They can have chronic health conditions which may cause dizziness and loss of consciousness. Weaker muscles and poor vision may lead to the person collapsing. Some illnesses that affect the cerebellum, also affect their balance. External factors may also affect the balance by displacing the centre of mass of the body. Falls are unintentional, albeit deleterious. Not only do they cause lacerations, contusions, bone fractures, trauma, tissue tears, but in some cases lead to death. It is better to alert the concerned family about the fall, in time.

To address this, we prepared a separate model. Since a fall was not being detected with satisfiable accuracy, to cater to this we pass it to another model. This will help us to estimate a fall quickly and efficiently. We use Single Shot Detector for fall detection. If a fall is detected, an alert is sent out.

In section II of this paper, we have discussed the existing solutions and the new improvisations. Our solution is explained in section III. Based on the test cases, the result of our proposed system is elaborated in the IV section of this paper.

II. Literature Review

With the increase in busy schedules, it becomes difficult for humans to always be present to monitor the elderly people and provide assistance when required. The paper [2] analyses the possibility of a supervision system. The raw data is collected and processed. Then the actions were observed and the chain of activities were recognised. Sensors are placed at strategic places in the house. The data collected from various sensors tells us about the activities performed by the individuals. Hidden markov model was then used to recognise the pattern

Authorized licensed use limited to: Macquarie University. Downloaded on October 16,2020 at 15:22:31 UTC from IEEE Xplore. Restrictions apply.

of activities. Viterbi algorithm was used to determine the activity. Based on the observations, the algorithm gave the most likely sequence of activities. This helps in predicting the activities of elderly people throughout the day and at a particular location.

The need of video surveillance for better security is increasing exponentially. The paper [3] proposed a method in which neural networks were used for activity detection. The input video was first broken down into frames. It was pre-processed and segmented. Background and noise subtraction was done. A series of image processing operations were done for feature extraction. Once the object was detected, it was classified as human or non human. The neural network was then trained with KTH and Weizmann video dataset. The frames obtained from the input video were tested and the activity was recognised. The confusion matrix was used to analyse the activity recognition rate. The implemented system could recognise five activities. A graphical user interface was provided to make this entire process seamless. In this way, humans could be monitored and security could be maintained.

For fall detection the paper [4] does a background subtraction from the input video frame. Thus, the Region of Interest is obtained and subsequently the threshold is calculated, which is the total height of the region of interest divided by three. If the region of interest falls below the threshold, a fall is detected.

The authors of the paper [5] have used Computer vision and the internet of things. They have employed a Raspberry Pi3 Model B camera for video surveillance. The video undergoes foreground extraction, to remove the static background. For foreground extraction they have used OpenCV's gaussian mixture based segmentation algorithm. From this extraction, they detect four movement features. These are aspect ratio, centre of mass, Hu moment variant and the angle of orientation. Certain inferences can be made from observing these four. Aspect ratio is the width of the extraction divided by its height. A high ratio signifies that the subject is lying down. Changes in the aspect ratio are perused to set a threshold value. A change in the centre of mass of the subject is noted. When the angle of orientation is nearly 90, the subject has fallen down. After using OpenCV's Hu moments, they obtain a feature vector whose values keep altering. It was observed that if the change is more than 50%, a fall is detected. If a fall is detected, an alert is sent out via email using SMTP in python.

The paper [6] used KTH and Weizmann video dataset and worked on total five activities. Out of which one is abnormal and the rest all are normal activities. In this paper the background is extracted from the raw video and the region of interest is extracted. The features obtained by this process are given as input to the Kmeans algorithm to generate index sequence which is then given as input to the hidden markov model. The authors have used Discrete Fourier Transform (DFT) and PCA for shape features extraction. The optical flow features extraction used Lucas Kanade optical flow method and used the terms of velocity, flow direction, vorticity, divergence and gradient tensor features. For classification of the activities First Order Hidden Markov Model is used. Baum Welch Algorithm is used to update the transition and emission matrices of HMM. The paper was limited to the activity recognition of a single person.

The authors of the paper [7] have built an activity recognition framework based on Markov Decision Process (MDP) using fish-eye camera streams. They collected data with a Raspberry Pi3 microcomputer. The images then are calibrated to obtain the perspective images of 1000 * 600 dimensions using OpenCV. They considered four states for activity sequence. New Markov Decision Process (MDP) based inference algorithm is developed for activity recognition. The MDP selects the next optimal state for each action based on transition probability T (probability of transitions of state based on the actions) and reward R where reward is the immediate points received after performing the action to that state. They also tried rule based inference instead of MDP. In rule based inference methods they used conditional statements to track the transitions of state based on the actions. They compared the results and they found that frameworks based on MDP perform better than the rule based methods and MDP-RCNN framework gave better results than MDP-YOLO.

In previous papers, we see that they have used one method to detect the abnormal activity or fall of a person. We have improved upon this by combining Activity recognition and Markov model.After an extensive research about the daily activities carried out by elderly people, we decided a threshold to compare the probabilities of the activities we got from our Activity recognition model. By a comparison with this threshold, it is decided whether the activity is alarming or not. To gain better results for detecting the fall in an elderly person, we have created a separate model, which uses object detection to first identify the human. Then subsequently, it checks for a fall by using the bounding box method.

III. Proposed Methodology

The input video will be broken down into frames which are then resized accordingly, converted into blobs and sent to two different models- Activity Recognition Model and Single Shot Detector Model. Refer to Fig. 1. for an outline of our methodology.

The activity recognition model recognises the activities in the frames. This sequence of activities is then sent to the markov model where the probability of the occurrence of that sequence is predicted. If this probability is below the threshold probability. then that sequence of activities is classified as abnormal. Since falls require immediate attention as compared to other abnormal activities, the videos will also be separately tested for fall detection so that better accuracy is obtained and immediate help can be provided. For this, the blob of images is sent to a SSD model where the object is detected. For every object that is detected as a person, the dimensions of the bounding box are obtained to determine the coordinates. If the difference between the X coordinates is greater than the difference between the Y coordinates, this indicates that the

Authorized licensed use limited to: Macquarie University. Downloaded on October 16,2020 at 15:22:31 UTC from IEEE Xplore. Restrictions apply.

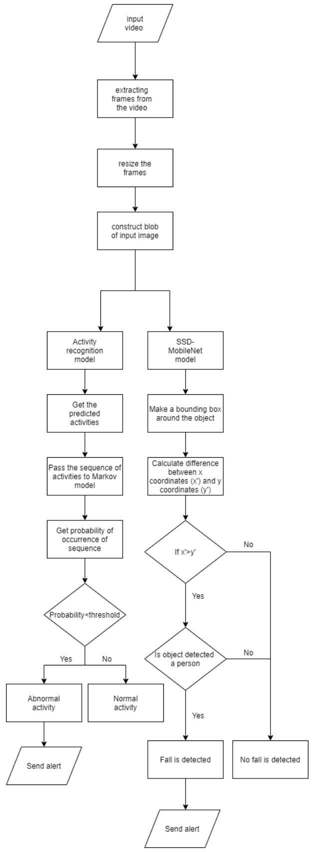

Fig. 1. Workflow diagram

person is in a sleeping position and hence might have fallen down. Hence, the difference in the dimensions are checked and if the object detected is a person, then it is predicted that the person has fallen down.

In this way, one can be informed when any abnormal activity takes place and immediate help can be provided to the elderly people.

A. Preprocessing

- Conversion of videos into frames
  The sample duration (i.e. the number of frames for classification) and sample size (i.e. the spatial dimensions of the frame) is decided and accordingly each video is converted into that. For the activity recognition model the frames are resized to a width of 400 pixels. For the fall detection model, the frames are resized to 300 * 300 pixels.
- Frames converted into blobs
  Open CV's dnn module is used to convert the frames into a blob of images which then serves as an input to our models. A blob is used to represent a group of pixels which have similar intensity but different from the ones surrounding it.

B. Activity Recognition Model

For activity recognition we used the The Kinetics Human Action Video Dataset [8]. We trained the model for 26 activities which includes cooking, extinguishing fire, reading,sneezing,exercising,bending back and many more.

We used the ResNet3D model to classify the frame into the activities. 3D kernels are used instead of the standard 2D filters. Hence we can include the temporal component for activity detection.

We have used 16 frames for classification and the spatial dimensions of the frames is 112. For classification,the frames are resized to a width of 400 pixels and converted into a tensor and sent as input to the resnet model. The blob is passed through the network. We then obtain a list of the predicted outputs. For each blob, the highest prediction is considered as the label. In this way each and every frame from the frame list is predicted and the activity is displayed. Fig. 2 to Fig. 4 show some of the activities detected.

Fig. 2. Activity detected- Knitting [9]

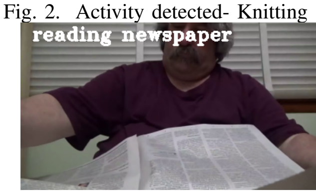

Fig. 3. Activity detected- reading [10]

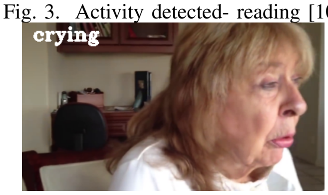

Fig. 4. Activity detected- crying [11]

C. Markov Model

According to Markov property, the transition from one state to another is solely dependent on the time taken and the current state. It does not depend on states that occured before the current state. This property can be used to the benefit of the model created as Markov processes are memoryless.

A discrete time Markov chain is used while creating the model where states are changed randomly between steps.

$$P(A_{n+1} = x | A_1 = a_1, A_2 = a_2, ..., A_n = a_n) = P(A_{n+1} = a | A_n = a_n)$$

In a Markov chain, we only need the probability of the previous state to find out the probability distribution of the

Authorized licensed use limited to: Macquarie University. Downloaded on October 16,2020 at 15:22:31 UTC from IEEE Xplore. Restrictions apply.

current state. The transition matrix shows the probability of the transition of an activity from the activity in the row to the activity in the column. Based on the activities recognised by the activity recognition model, the transition matrix was made. For old people, the matrix will vary as their activities are different from those of an active young adult. The probabilities of transitioning from the activity in a particular row to the activity in the column was used to create the transition matrix. To find out the probability of occurrence of that transition, the probability is multiplied by the corresponding value from the transition matrix.

The final probability is found in this manner by multiplying the individual probabilities of the corresponding transition from the transition matrix to the probability calculated of the previous state.

We obtain the sequence of activities from the activity recognition model which are then given to the markov model. Here for each transition, the following action is performed.

$$P(current\, state) = P(previously\, obtained) * \\ corresponding\, value\, from\, the\, transition\, matrix \\ for\, the\, current\, transition.$$

Based on the transition matrix, the probability of the occurrence of that particular sequence is predicted. The final output of the code gives you how likely it is that the activities of a normal person would occur in that order.This probability is then compared with the threshold. Since the probability found is the likeliness of the occurrence of that sequence, higher the probability the more likely that an old person would follow that sequence. The threshold is the value which is used to classify whether the sequence is normal or abnormal. The decided threshold was 0.2. If the probability is below the threshold, it is implied that the chances of the occurrence of activities in this sequence is very unlikely. This means that this is an abnormal activity.

D. Fall Detection

The blob of images is also sent to an Single Shot Detector-MobileNet model where object detection takes place. We have taken sample videos [12] to test our model, see Fig. 5 to Fig. 7 for the results.

First we will look at the SSD and MobileNet architectures and why we used a combination of both.

- Single Shot Detector
  Single shot detector, also known as SSD is a feedforward convolutional technique which is used for object detection in real time [13]. It is faster as compared to faster Region based- RCNN as it has eliminated the region proposal network,that is employed by a faster Region based-CNN. To compensate for the lower accuracy as compared to faster R-CNN, SSD makes use of default boxes and multi-scale features. SSD makes a fixed size group of bounding boxes and scores to determine the presence of object class instances. This is then followed by non maximum suppression, which means that each object is detected only once. The final object detection is then produced. SSD object detection uses VGG-16 to extract feature maps. After this, it uses 3 * 3 convolutional filters to make predictions. It has several extra convolutional layers, which makes it better than YOLO.
  Classification is performed by comparing the value of the default boundary box's intersection of union with 0.5, which is the threshold. If it exceeds 0.5, then, it is termed as a positive match. Only positive matches are used to calculate the cost. It has proved to be significantly faster than faster R-CNN and gives better accuracy than YOLO.
- MobileNet
  MobileNet depends on the depth wise separable convolutional neural networks [14]. This reduces overfitting, due to lesser parameters and is significantly cheaper computationally.
  It is dual layered, with depthwise convolutions as one layer and pointwise convolutions as another. The former makes use of one filter for each input channel, whereas the latter, makes a linear coalescence of the depthwise layer's outputs. The pointwise layers are 1 * 1 layers, where depthwise layers are 3 * 3. All, but the last layer, of a total 28 layers ,are followed by batch norm and Rectified Linear Unit,known as RELU. The last layer doesn't conform to non-linearity, and feeds it into a softmax classifier. To further reduce the cost of computation and parameters, the hyper-parameters α and ρ are introduced. α is the width multiplier and ρ is the resolution multiplier. α does so by reducing the thickness of the layers uniformly. ρ reduces internal representation of the layers when applied to the input.

In our paper we have used both Mobilenet and SSD for object detection. Using a combination of these both, increases the speed and efficacy of our model.

The combination of MobileNet and SSD was trained on the COCO dataset [15]. It was then adjusted and improved on PASCAL VOC to get 72.7% mAP. We have used the caffe framework [16].

The video is fed in the Single shot detector- MobileNet model. Various objects in the blob of images are detected and can be classified.

Different layers of the network use different sizes of the default box. The dimensions of the default boxes are chosen manually. The width is denoted by w and height by h.

$$w = scale * aspect\, ratio \tag{1}$$

$$h = scale \div \sqrt{aspect\, ratio} \tag{2}$$

Here scale varies from 0.2 to 0.9 as the layers progress. The aspect ratio in SSD are ,½,1,2,3. From the bounding box obtained by the SSD model, we can find out the difference

Authorized licensed use limited to: Macquarie University. Downloaded on October 16,2020 at 15:22:31 UTC from IEEE Xplore. Restrictions apply.

between the dimensions of the X coordinate and the difference between the dimensions of the Y coordinate.

Fig. 5. Original position [12]

Fig. 6. About to fall

Fig. 7. Fall

If the difference between X coordinates is lesser, the bounding box is vertical, which implies the person is standing. (See Fig 5. and Fig 6.) If the difference between X coordinates is more, the box is said to be in a horizontal position. This means that the object is in sleeping position which is likely due to a fall.(See Fig 7.) We then check if the object detected via the SSD model which is in sleeping position is a person. If both these criterias are satisfied, we detect this as a fall and alert the concerned people.

IV. Result

We used Kinetics Human Action Video dataset for Activity recognition. We trained our model on 26 activities. Out of them 18 were showing proper results for the action they were performing.

Fig. 8. Activity detected- massaging person's head [17]

Fig. 9. Activity detected- sneezing [11]

Among the various symptoms of coronavirus, some of the symptoms that can be easily identified via camera are coughing, sneezing(see Fig.9), blowing nose. Some other symptoms can be interpreted via some activities like regularly bending of back by old people can be a sign that they have back pain. Regular massaging(see Fig.8) of the head can indicate that they have severe head pains which is one of the major symptoms that people experience when they are infected by corona virus. The repeated observation of these activities can act as symptoms for this pandemic disease. If a person is diagnosed with the disease, we can use the video captured by the surveillance camera to find out all the other individuals they came in contact with like shaking hands, hugging, and many other activities which have direct contact with the infected individual. These activities are detected by our model and we can directly skip to the timestamp of these activities for early detection of all the contacted individuals who can immediately be quarantined.

We also detected other activities like knitting, reading newspaper,crying (see Fig 2 to Fig.4) which were sent to the Markov Model to check if they are abnormal.The activities which were below the threshold of 0.2 were detected as abnormal. The markov model gave the results as shown in Table 1 for normal and abnormal activities that were detected. Videos in which multiple people were present didn't give us accurate results.

TABLE I
RESULT OF MARKOV MODEL

Activity Sequence	***Probability***	***Less than threshold(abnormal)***
Reading	0.4	no
Brushing	0.5	no
Scambling eggs	0.24	no
Stretching Arms	0.45	no

Activity Sequence	***Probability***	***Less than threshold(abnormal)***
Cooking	0.3	no
Crying	0.01	yes
Stretching arms	0.4	no
Knitting	0.5	no

For fall detection we used the COCO dataset and tested our model on [12] videos. Out of our 5 test videos, 3 correctly predicted the fall. The image resolution mattered here, in the blur images our SSD-MobileNet model was unable to recognize faces accurately.

V. Conclusion

Elderly people generally follow a regular pattern of activities in their day-to-day life and we used this fact to detect whenever abnormal activities are performed. This paper proposed a method to detect these abnormalities via videos. For fall detection, Single shot detector - MobileNet model is used. If a fall is detected, the system will immediately notify their relatives. The system is evaluated obtaining an accuracy of 60% of fall detection. Results may vary with

Authorized licensed use limited to: Macquarie University. Downloaded on October 16,2020 at 15:22:31 UTC from IEEE Xplore. Restrictions apply.

different resolutions of cameras, in 2 of the test videos our model got confused whether the object detected is that of the person, hence giving fallacious predictions. The activity recognition model followed by the Markov model is used to find other abnormal activities which can also be notified.

We have noted that our approach is not helpful in the medical emergency situations of elderly people such as heart attack which cannot be detected on camera. Our approach is mainly focused on the significant activities seen in the videos. For future work, our system can be improved to recognise and identify multiple people in the videos, keep track of more number of activities and have a wider range of abnormal activities from real data collected by the system.

We trained our model on the common COVID-19 symptoms which can be observed, we could determine whether a person was sneezing or had a headache.

We plan to gauge if the person is coughing or sneezing more accurately by monitoring their audio signals. This will let us ascertain if the person needs medical assistance for COVID-19. We also plan to research more on the other uncommon symptoms which occur in COVID-19 patients.

REFERENCES

[1] "Falls fact sheet by WHO," World Health Organization. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/falls

[2] G. Sebestyen, I. Stoica, and A. Hangan, "Human activity recognition and monitoring for elderly people," IEEE 12th international conference on intelligent computer communication and processing(ICCP), pp. 341–347, 2016.

[3] M. Babiker, O. O. Khalifa, K. K. Htike, A. Hassan, and M. Zaharadeen, "Automated daily human activity recognition for video surveillance using neural network," IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA), pp. 1–5, 2017.

[4] P. S. Sase and S. H. Bhandari, "Human fall detection using depth videos," IEEE 5th International Conference on Signal Processing and Integrated Networks (SPIN), pp. 546–549, 2018.

[5] N. B. Joshi and S. Nalbalwar, "A fall detection and alert system for an elderly using computer vision and internet of things," 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), pp. 1276–1281, 2017.

[6] M. H. Kolekar and D. P. Dash, "Hidden markov model based human activity recognition using shape and optical flow based features," IEEE Region 10 Conference (TENCON), pp. 393–397, 2016.

[7] A. Oner, M. S. Albayrak, F. Guner, and I. M. Atakli, "An activity recognition application based on markov decision process through fish eye camera," IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), pp. 251–258, 2019.

[8] W. Kay, J. Carreira, K. Simonyan, B. Zhang, C. Hillier, S. Vijayanarasimhan, F. Viola, T. Green, T. Back, P. Natsevet al., "The kinetics human action video dataset," arXiv preprint arXiv:1705.06950, 2017.

[9] seattletimesdotcom,"100-year-old knitter shares her talents," Youtube, Jul2017. [Online]. Available: https://www.youtube.com/watch?v=T5aAGJ29u2s

[10] ManBehindACameraASMR,"Asmr newspaper page turning (notalking)," Youtube, May2018. [Online]. Available: https://www.youtube.com/watch?v=nWo--dSu9bs

[11] Mr Satisfying, "Grandma sneezes dramatically," Youtube, Oct 2016.[Online]. Available: https://www.youtube.com/watch?v=2mfuHazBzsw

[12] The Lancet, "Falls in elderly people 1/5," Youtube, Dec 2013. [Online].Available: https://www.youtube.com/watch?v=p5i4z3sNaKM

[13] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu,and A. C. Berg, "Ssd: Single shot multibox detector," Springer European conference on computer vision, pp. 21–37, 2016.

[14] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang,T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.

[15] "Coco dataset," Creative Commons. [Online]. Available: http://cocodataset.org/

[16] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick,S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," Proceedings of the 22nd ACM international conference on Multimedia, pp. 675–678, 2014.

[17] Howcast,"How to self-massage your scalp — head massage," Youtube, Jan2014. [Online]. Available: https://www.youtube.com/watch?v=98bJCul6QgE

[18] Q. Zhang and S. Zhu, "Real-time activity and fall risk detection for aging population using deep learning," 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference(UEMCON), pp. 1055–1059, 2018.

[19] S. Neili, S. Gazzah, M. A. El Yacoubi, and N. E. B. Amara, "Human posture recognition approach based on convnets and svm classifier," IEEE International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), pp. 1–6, 2017.

[20] A. Subasi, M. Radhwan, R. Kurdi, and K. Khateeb, "Iot based mobile healthcare system for human activity recognition," IEEE 15th Learning and Technology Conference (L&T), pp. 29–34, 2018.

[21] C.-Y. Lin, S.-M. Wang, J.-W. Hong, L.-W. Kang, and C.-L. Huang,"Vision-based fall detection through shape features," IEEE Second International Conference on Multimedia Big Data (BigMM), pp. 237–240, 2016.

[22] A. Y. Alaoui, A. El Hassouny, R. O. H. Thami, and H. Tairi, "Video based human fall detection using von mises distribution of motion vectors," IEEE Intelligent Systems and Computer Vision (ISCV), pp. 1–5, 2017.

[23] S. C. Agrawal, R. K. Tripathi, and A. S. Jalal, "Human-fall detection from an indoor video surveillance," IEEE 8th International Conference on Computing, Communication and Networking Technologies(ICCCNT), pp. 1–5, 2017.

[24] S. Kalita, A. Karmakar, and S. M. Hazarika, "Human fall detection during activities of daily living using extended core9," IEEE Second International Conference on Advanced Computational and Communication Paradigms (ICACCP), pp. 1–6, 2019.

[25] S. W. Pienaar and R. Malekian, "Human activity recognition using visual object detection," IEEE 2nd Wireless Africa Conference(WAC), pp. 1–5, 2019.

[26] A. Ahir, P. K. Pateriya, D. Kaur, and M. K. Rai, "A review on abnormal activity detection methods," IEEE International Conference on Computational Techniques, Electronics and Mechanical Systems(CTEMS), pp. 521–526, 2018.

[27] C. Liu, J. Ying, F. Han, and M. Ruan, "Abnormal human activity recognition using bayes classifier and convolutional neural network," IEEE 3rd International Conference on Signal and Image Processing (ICSIP), pp. 33–37, 2018.

[28] C. N. Phyo, T. T. Zin, and P. Tin, "Deep learning for recognizing human activities using motions of skeletal joints," IEEE Transactions on Consumer Electronics, vol. 65, no. 2, pp. 243–252, 2019.

[29] Mop King, "How a man brushes his teeth," Youtube, Oct 2013. [Online]. Available: https://www.youtube.com/watch?v=_Yv5mPo8_wA

[30] Great Depression Cooking, "Great Depression Cooking - Depression Breakfast," Youtube, Jan2009. [Online]. Available: https://www.youtube.com/watch?v=9zXqkHvs0po

[31] OUresearch on YouTube, "Developing age-friendly kitchens," Youtube, Feb2014. [Online]. Available: https://www.youtube.com/watch?v=YDutKpHHCYg

[32] The Lancet , "Falls in elderly people 2/5," Youtube, Dec 2013. [Online]. Available: https://www.youtube.com/watch?v=1IsM08Sh_wg

Authorized licensed use limited to: Macquarie University. Downloaded on October 16,2020 at 15:22:31 UTC from IEEE Xplore. Restrictions apply.