Video Surveillance Using Machine Learning for Checking Emergency Conditions

Sharvee Kawade
Computer Science & Information Technology
G H Raisoni Institute of Engineering and Technology
(Savitribai Phule Pune University)
Pune, India
shrvi01@gmail.com

Ashwini Dewade
Computer Science & Information Technology
G H Raisoni Institute of Engineering and Technology
(Savitribai Phule Pune University)
Pune, India
dewadeashwini11@gmail.com

Aishwarya Thepade
Computer Science & Information Technology
G H Raisoni Institute of Engineering and Technology
(Savitribai Phule Pune University)
Pune, India
aishwaryathepade03@gmail.com

Sarika Sangle
Computer Science & Information Technology
G H Raisoni Institute of Engineering and Technology
(Savitribai Phule Pune University)
Pune, India
sarikasangle123@gmail.com

Under the Guidance of,
Mrs. Shubhangi Suryawanshi
Computer science & Information Technology, GHRIET
(Savitribai Phule Pune University)
Pune, India
Subhangi.suryawanshi@raisoni.net

Abstract— The present paper considers the uses of deep learning and transfers learning techniques in Fall Detection by means of surveillance camera data processing. It becomes very necessary to take care of aged at home or at least being aware of their state while being away. Keeping this in mind we present an approach of emergency condition like fall detection due to various reasons –body imbalance due to lack of cerebral blood supply, muscle weakness, stroke etc. such health hazardous situations need immediate aid which is our primary effort. Our focus is live video streaming and grabbing frames of ROI further analyzing fall or not fall.

Keywords—*deep learning, machine learning, live video streaming, intelligent video analysis, fall detection.*

*****

I. INTRODUCTION

At present, surveillance cameras are widely used. They are utilized to identify suspicious persons, as well as persons in a state of alcohol and drug intoxication, which may be dangerous to a life and health of others. The developers of the closed-circuit television system (CCTV systems) offer complexes consisting of IP cameras for the automated detection of suspicious persons. These systems can be based on biometric identification (for example, the NeoFace system, smart glasses R7) [1], [2], as well as on the recognition of emotions or facial expressions (for example, DeepFace) [3]– [6]. The main disadvantage of this type of systems is that information about the possible trespasser may be absent in the database used for biometric identification.

It should be noted that in the majority of cases, data from CCTV cameras are used only for the creating archives of video records, and the possibilities of intelligent video analysis of data from remote objects are practically not used. Such situation is determined in particular by technical difficulties associated with the use of intelligent video surveillance systems in practice. For example, classification algorithms are usually extremely sensitive to lighting conditions [7]. In the present work, the problem of developing such a system for recognizing human movements through the records of CCTV cameras and detecting falls has been considered. The fall of a person can be caused by many factors: loss of balance due to lack of cerebral blood supply, muscle weakness, etc. In any case, a situation when a person has fallen and cannot get up without assistance is dangerous and requires an immediate response.

The problem of detecting falls using video analysis has been studied in a large number of works [8], [9], which were based on analyzing the shape and position of the person in the frame, gradients in the vertical and horizontal directions, and changes in images in the time domain.

II. LITERATURE SURVEY

Presents a novel nonlinear subspace learning technique for class-specific data representation is proposed. Novel data representation is obtained by applying nonlinear class-specific data projection to a Discriminant feature space,[1] where the data belonging to the class under consideration are enforced to be close to their class representation, while the data belonging to the remaining classes are enforced to be as far as possible from it. A class is represented by an optimized class vector, enhancing class discrimination in the resulting feature space. An iterative optimization scheme is proposed to this end, where both the optimal nonlinear data projection and the optimal class representation are determined in each optimization step. The proposed approach is tested three problems relating to human behavior analysis: Face recognition, facial expression recognition, and human action recognition. Experimental results denote the effectiveness of the proposed approach, since the proposed class-specific reference Discriminant analysis outperforms kernel Discriminant analysis, kernel spectral regression, and class-specific kernel Discriminant analysis, as well as support vector machine-based classification, in most cases.

Proposed a new method to detect elderly person falls is proposed. The combination of motion and change in the human shape gives crucial information on human activities. Their fall detection system has proven its robustness on realistic image sequences of simulated falls and daily activities. In this work, they make the assumption that the person is on the ground with no or little motion after a fall. This can be argued in some circumstances, for example in the case of injury,[2] where the person could move rapidly because of the pain. A solution to increase the robustness of their system could be the addition of 3D information to check if the head is near the door for Instance. Another way to improve the system could be the use of the audio information from the microphone of the webcam. A speech recognition algorithm could also be used to identify distress cries like "Help!". Error detections typically occur when the person brutally sits down because of large motion and variation in orientation. To limit this type of false detection, they could de ne normal inactivity zones. The thresholds were chosen manually by logical reasoning on what is a fall and observation of their video sequences. However, an automatic method could be implemented to de ne the thresholds using a training dataset.

Designed an intelligent emergency response system to detect falls in the home. It uses image-based sensors. A pilot study was conducted using 21 subjects to evaluate the efficacy and performance of the fall detection component of the system.[3] Trials were conducted in a mock-up bedroom setting, with a bed, a chair and other typical bedroom furnishings. A small digital video camera was installed in the ceiling at a height of approximately 2.6 m. The digital camera covered an area of approximately 5.0 m-3.8 m. The subjects were asked to assume a series of postures, namely walking/standing, sitting/lying down in an inactive zone, stooping, lying down in a `stretched' position, and lying down in a `tucked' position. These very scenarios were repeated three times by each subject in a random order. These test positions totaled 315 tasks with 126 fall-simulated tasks and 189 non-fall-simulated tasks. The system detected a fall on 77.

Recent research into fall detection has included automated fall detection, i.e. the use of technology to detect when a fall occurs. [4] Worn fall detectors are mechanical sensors that can be worn on either the hip or torso, which trigger an alarm when both the orientation and acceleration forces of the person reach a pre-set threshold.[16] A primary limitation of existing devices is that they require effort from the user in order to be effective. For example, the user must remember to wear the device, which cannot be reliably assumed. Other developments have focused on devices that are embedded in the user’s environment. There are several consequences of a fall in the elderly, including hip fractures, fear of falling and

‘long lies’, all of which could detrimentally affect both the psychological as well as physical wellbeing of the individual. A long lie has been defined in a study conducted by Nevitt et al. as a fall in which the person reported remaining on the ground for 5 min or more before being able to get up without assistance, or help arriving.[19] A ‘long lie’ can lead to several physiological complications, such as hypothermia, dehydration, bronchopneumonia and pressure sores. Furthermore, elderly subjects tended to be less tolerant of prolonged intervals from major injury until intervention. Studies have also shown that long lies are more likely to lead to a greater fear of falling, and a damaging consequence of both functional decline and social isolation[16] Finally, reducing the occurrence of falling is very difficult due to the complex etiology of falls.

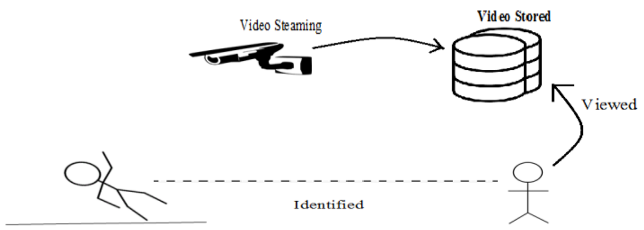
The VACE (Video Analysis and Content Extraction). Program will embark on a second two-year R&D Phase. The focus of this phase of VACE is on moving beyond the detection, recognition and tracking of objects in video streams to the detection, [20] recognition and understanding of the activities that the objects are engaged in. The VACE Program is interested in video events in all types of video: these types include: News Broadcast video, Meeting/Conference Video, UAV Motion Imagery and Ground Reconnaissance Video as well as Surveillance Video.

Fall is a major health issue among elderly person. Fall may cause serious injury to the elderly person such as heap fractures and head traumas. A method for detecting falls among elderly person is presented which uses the Microsoft Kinect sensor. The system consists of two parts. [21] The first one is to detect fall using Kinect sensor and second part is to generate control signal to send message. To detect fall condition, motion related features are calculated. For this sensor's color and output is used. Ground event segmentation is done to detect moving objects from the video stream followed by feature extraction and event classification. After fall detection, help message will send to the registered number so that immediate medical assistance will get. The system works on real time data.

In this paper we present a method for real-time detection of human fall from video for support of elderly people living alone in their homes. The detection algorithm has four steps: background estimation, [22] extraction of moving objects, motion feature extraction, and fall detection. The detection is based on features that quantify dynamics of human motion and body orientation.

III. METHODOLOGY

Until now we have come across projects where manually checking of the robbery, fall etc. was checked by the security in charge. And also recently in our survey we got to know that at most current work is being done by taking a database of videos and by using Machine Learning domain it has become easier and efficient to check for detecting falls, robbery, security of homes, work places etc. But the drawback here can be that previous work checked database which was collected and then checking was done on stored video.



Further we also surveyed and found out that this also can be improvised by recognizing facial expression of the victim so that we can differentiate among whether he/She have fallen because of fainting, heart attack, or whether just kids are playing and hopping around.

During the fall event, the motion quantity of the object is larger than other normal activities. Here in our implementation we take Region of Interest (ROI) as the height variance of the human in every frame. The ratios of width to height during the fall activities such as forward falls, backward falls, and falling aside also are obviously very higher than the ratios of normal activities using temporal templates. If the ratio is above a threshold value, then the system decides whether a fall activity occurs or not. On the basis of survey and our understanding we wish to propose a safety system that would enable us to keep a check for safety conditions.

As stated before we came up with an idea of live video surveying and using AI checking immediately for fall and notifying the user, we can surround this concept by more variations, and simplicity in implementation. We also propose use of CNN AlexNet (2012) for identification of images and objects in an image successfully, it has 11x11, 5x5, 3x3 convolutions and more filters per layer.

Below we suggest a method and describe corresponding modules.

A. Video Frame Collection

The Input we would be providing is videos captured through JMF i.e. Java Media Files and further grabbing frames using frame grabbing technique for analysis. We can set time interval for grabbing frames.

B. ROI – (Region of Interest)

The video frames which we have collected during first module is taken as input in this stage. Also we carry out the process of gray scale conversion of grabbed frames and image thresholding. Where we consider threshold value as 125. Pixels having value greater than 125 are made Black and greater than 125 are made White.

C. Temporal Effect

Input provided is taken as ROI frames. The process to be carried out mainly in this module is Axial covariance process. And the corresponding output we get is shape identification and differentiation through covariance ratio list.

D. CNN

Input here is a morphological vector. Here the main task is to carry out frame intersection technique and learning. And corresponding output is fall vector.

E. Decision Tree

Input is the fall vector obtained in previous module. Further we use If –Then rules for classification of images whether fall has occurred or not. And finally we are able to identify fall without human intervention and engaging of human force.

IV. ACKNOWLEDGEMENT

I would like to thank all the staff of Computer department, GHRIET, Wagholi for their unconditional support and encouragement throughout the academic period.

V. REFERENCES

[1]. www.necam.com/docs/?id=6c812b4d-2a12-40ed-9fea-fae81550c7aa
[2]. www.osterhoutgroup.com/pub/static/ version1515417478/ frontend/ In fortis/ ultimo/ en US/ pdf/ R-7-TechSheet.pdf
[3]. Yaniv TaigmanMing YangMarc'Aurelio RanzatoLior Wolf, DeepFace: closing the gap to human-level

performance in face verification, Conference on Computer Vision and Pattern Recognition (CVPR), June 24, 2014.

[4]. www.robots.ox.ac.uk/ vgg/publications/2015/Parkhi15/parkhi15.pdf.

[5]. Mohammadian, A., Aghaeinia, H., Towhidkhah, F. Video-based facial expression recognition by removing the style variations in Image Processing, IET, 2015, vol. 9, no. 7, pp. 596–603.

[6]. Iosifidis A., Tefas A., Pitas, I. Class-specific reference discriminant analysis with application in human behavior analysis, IEEE Transactions on Human- Machine Systems, 2015, vol. 45, no. 3, pp. 315–326.

[7]. Rice, D., Evaluating camera performance in challenging lighting situations, 2014. www.sdmmag.com/articles/90525-evaluating-camera-performance-in-challenging-lighting-situations

[8]. Rougier, C., Meunier, J., St-Arnaud, A., Rousseau, J. Fall detection from human shape and motion history using video surveillance, Proc. 21st Int. Conf. AINAW, 2007, vol. 2, pp. 875–880.

[9]. Lee, T., Mihailidis, A. An intelligent emergency response system: Preliminary development and testing of automated fall detection, J. Telemed. Telecare, 2005, vol. 11, no. 4, pp. 194–198.

[10]. https://en.wikipedia.org/wiki/Convolutional neural network

[11]. https://en.wikipedia.org/wiki/Transfer learning

[12]. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. ”ImageNet classification with deep convolutional neural networks.” Advances in neural information processing systems. 2012.

[13]. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, november 1998.

[14]. Charfi, I., Miteran, J., Dubois, J., Atri, M., Tourki, R., Optimised spatio- temporal descriptors for real-time fall detection: comparison of SVM and Adaboost based classification, Journal of Electronic Imaging (JEI), Vol.22. Issue.4, pp.17, October 2013.

[15]. Cohen, J., A coefficient of agreement for nominal scales, Educational and Psychological Measurement, 1960, 20 (1), pp. 37–46.

[16]. Maki BE. Physical consequences of falls part II: an aging population will lead to mounting fall-related health-care costs. Geriatrics and Aging 2000;3:23

[17]. Shaw FE. Falls in older people with dementia. Geriatrics and Aging 2003;6:37–40

[18]. Johnson M, Cusick A, Chang S. Home-Screen: a short scale to measure fall risk in the home. Public Health Nurs 2001;18:169–77

[19]. Cryer C, Knox A, Martin D, Barlow J. Cantebury hip protector project team. Hip protector compliance among older people living in residential care homes. Inj Prev 2002;8:202–6

[20]. Detecting, Recognizing and Understanding Video Events in Surveillance Video. John D. Prange 2003 IEEE

[21]. Fall Detection System for Older Adults, Yogesh Angal, Arti Jagtap Dec 2-3, 2016

[22]. A method for real-time detection of human fall from video,M. Kreković, P. Čerić, T. Dominko, M. Ilijaš, K. Ivančić, V. Skolan i J. Šarlija May 21- 25,2012