Using 3D Convolutional Neural Network in Surveillance Videos for Recognizing Human Actions

Sathyashrisharmilha Pushparaj[1] and Sakthivel Arumugam[2]
[1]Department of Computer Science and Engineering, Adithya Institute of Technology, India
[2]Department of Information Technology, Woldia University, Ethiopia

Abstract: *Human action recognition is a very important component of visual surveillance systems. The demand for automatic surveillance systems play a crucial role in the circumstances where continuous patrolling by human guards are not possible. The analysis in surveillance scenarios often requires the detection of certain specific human actions. The automated recognition of human actions in detecting certain human actions are considered here. The main aim is to develop a novel 3D Convolutional Neural Network (CNN) model for human action recognition in realistic environment. The features are extracted from both the spatial and the temporal dimensions by performing 3D convolutions, by which, capturing the motion information encoded in multiple adjacent frames. The evolved model generates multiple information from the input frames, and the information from all the channels are combined and that is to be the final feature. The developed model automatically tends to recognize specific human actions which needs attention in the real world environment like in pathways or in corridors of any organization. This proposed work is well suitable for the situations like where continuous patrolling of humans are not possible, to prevent certain human actions which are not allowed inside the organisation premises.*

Keywords: *Security surveillance, convolutional neural networks, 3D convolution, feature extraction, image analysis and action recognition*.

Received April 29, 2014; accepted January 27, 2015

1. Introduction

Computer Vision and Pattern Recognition have received much attention in recent years. Recognizing human actions from the realistic surveillance videos has extensively gained its importance in a wide range of applications. Such applications are as follows: advanced intelligent visual surveillance [6, 7, 14] e.g., detecting any suspicious or certain specific actions done by the human which are captured in the surveillance videos and intelligent video content analysis [2, 15] e.g., recognizing the suspicious actions by which it helps in preventing the abnormal actions done by humans in public places, corridors, verandas of the organization in order to enhance the effectiveness of patrolling for every seconds [7]. In that case, this recognition of human actions can be much better than compared to object recognition.

In order to proceed the work, everyone must have a prior and clear knowledge about action and activity. The term ‘action’ is different from the term ‘activity’. An action is referred to as a simple motion pattern. The activities consist of a series of actions. Therefore, each and every activity can be split into its elementary motion patterns called actions.

The Human action recognition [1] provides a remarkable challenge in providing an accurate recognition of actions. The accuracy of recognizing the actions are not appropriate due to cluttered backgrounds, occlusions, and variations in the viewpoint etc., [13]. Action recognition methods can be divided based either on the visual information employed for action description, or based on the number of cameras employed in order to obtain the available visual information [12]. The Figure 1 shows the general mind-map for current trends in image processing.

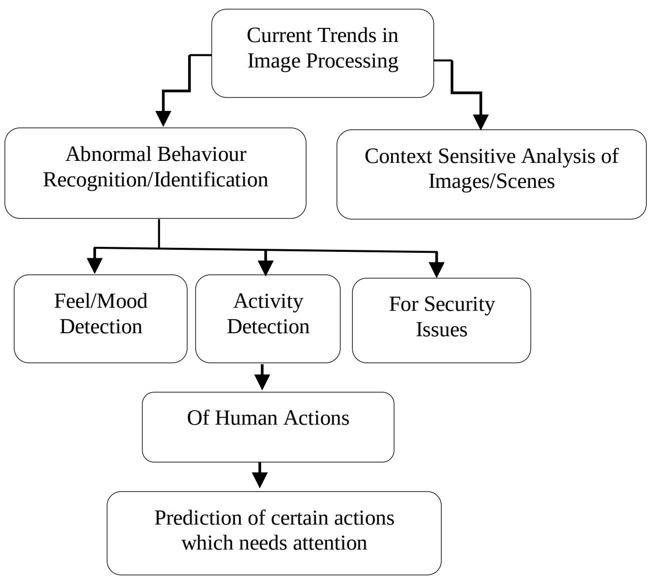

Figure 1. Mind-Map for current trends in image processing.

The purpose of this system is to train a system which could automatically recognize certain different actions done by the humans not abide by the rules in common places. Next, this system helps the management to organize well and everything goes well in its premises. To do so, the first thing to be done is segmenting the video into a separate image frames. By using the 3D Convolutional Neural Network, the features are to be extracted from the continuous frames, as the actions prevails to be same in the adjacent frames. Using the Gabor filter [4] to increase the performance. Since, neural network is the core idea hidden here, based on the template actions, the actions can be distinguished as normal or suspicious actions. This work focuses on online error back- propagation algorithm [3] for training purposes and the convolutional neural network are the best techniques for recognizing non-rigid objects like humans.

The rest of this paper is organized as follows: The description and comparison of reference work through surveying the papers are given in section 2. Some related work for the human action recognition is in section 3. The proposed work is in section 4. The forthcoming experimental results are discussed in section 5. The analysis of performance is in section 6. Finally, the work is concluded in section 7 with discussions.

2. Survey of Literature

Yao *et al*. [16], described about the Animated Pose Template (APT) for detecting the short-term, long-term and contextual actions. These actions are identified through cluttered scenes from the video. The whole work revolved about the pose template [10, 11]. The pose templates can be classified as Shape and Motion templates represented in AND-node, Histogram of Oriented Gradient (HOG) and in Histogram of Optical Flow (HOF) respectively. The Pose templates are suitable for detecting the short term action snippets in 2-5 frames. The training of the model is happened by manually or also by dynamic programming. It is emphasized on Semi-supervised structural Support Vector Machine (SVM) algorithm. Though this model suffers from 2D view dependency still it suits for most of the applications.

Fan *et al*. [5], proposed a novel learning method for tracking the humans using convolutional neural networks. Human Tracking [8] is a special case of object tracking. The idea behind the Convolutional Neural Network (CNN) tells that from the input video sequences having the individual adjacent frames learns both the spatial and temporal features jointly from the image pairs of two adjacent frames. A Creative Shift-variant CNN architecture was developed to alleviate the drift problem, also well known for tracking the objects. The actual capability of the tracker can be exactly demonstrated in testing the sequences. Since this whole work depends on the scale of an object, there occurs the scaling problem and it is solved by tracking key points.

Paul *et al*. [15] presented a survey undertaken in the detection of humans from videos. The main idea was about the detection process and it can be applicable to most of the applications. It begins with the input followed by object detection later on followed by object classification and finally the detection of humans happened so. The most commonly used techniques for detection of objects are Background subtraction, Optical flow and Spatio-temporal filter. This work was to help people in having an idea to choose an appropriate techniques for detecting humans.

Oluwatoyin and Wang [14], described about the review of modeling human behaviors and activity patterns for recognition or detection of special events. First of all, finding the abnormality for building Intelligent Vision System (IVS). The IVS mainly aimed at scene understanding and correct semantic inferences. It exactly focuses on contextual abnormal human behavior detection. While analyzing the whole work, it is able to find that the abnormal activity happened. The learning framework could be supervised, unsupervised or semi-supervised. This was the whole summarization of the work.

Iosifidis *et al*. [12], proposed a novel method aims at view-independent human action recognition. Actions are of Representation and Classification. Action representation involves Fuzzy Vector Quantization. Action Classification done by Feed Forward Neural Network. A novel classification algorithm known as Minimum Class Variance Extreme Learning Machine (MCVELM) which enhances the classification performance. It successfully operate on real time scenarios, as it doesn't take assumptions. It uses five publicly available datasets. To the whole, it aimed at the effectiveness of both adopted action recognition approach and the MCVELM algorithm.

By making the discussions among various papers that yields an idea through which the sort of methods can be used for the appropriate purposes. Here, many methods have been proposed in order to increase the performance, improve the recognition rates, robust, can be applicable to the variations in the viewpoints.

3. Related Work

A lot of work has been done in recognizing actions from video sequences. CNNs shows an excellent performance for visual recognition. The CNNs is closely related with neural networks. For the task of visual recognition, the CNNs first extract and combine the local features from the sequence of input images of real time video. Next, these features are then combined by the subsequent layers in order to obtain higher order features. Once the input video is fed into the system, the immediate work done is the segmentation of video in frames. It is worth noting that feature extraction is a nontrivial problem. This may lead to the variations in the position of distinctive features in the input objects. To overcome these variations, CNNs combine the architectural ideas to ensure some degree of shift, scale and distortion invariances are local receptive fields, shared weights and downsampling.

Once a feature has been detected, its approximate position relative to other feature is relevant rather than its exact location. The CNN model consists of alternating layers of convolutional features detectors (C layers) and downsampling layers (S layers) [13]. In particular, at the S layer a still input image is first analysed by an array of Gabor filters. The main purpose for introducing the downsampling layers is to reduce the resolution of the feature map and also the sensitivity of the output to shifts and distortions.

Especially for the recognition or detection tasks, the primitive feature detectors that are useful on one part of the image are likely to be useful across the entire image. Units in a layer are organized in planes within which all the units share the same set of weights. By sharing the same set of weights, CNNs have the shift-invariant property so that it has been able to achieve excellent performance in various recognition or detection tasks.

The various methods and techniques used in developing this idea are as follows: The 2D and 3D convolution using sigmoid function for evaluating the values obtained from different frames at the target frame. Background subtraction done as the preprocessing work, Gabor filter for extracting features, Backpropagation for training the system through learning methods, Gauss-Newton method approximation to Hessian matrix and stochastic diagonal Levenberg-Marquardt method.

Before extracting the features, the preprocessing is done by means of background subtraction. This is achieved by setting a threshold value and it is given by Equation (1),

$$g(x,y) = \begin{cases} f(x,y), if\,|(x,y)-b(x,y)| > threshold \\ 0, ifotherwise \end{cases} \tag{1}$$

Where, *g(x,y)* is the result of pixel's value, *f(x,y)* is the foreground pixel's value and *b(x,y)* is the background pixel's value.

After preprocessing completed, feature extraction begins. In a 2D image frame, the 2D convolution is performed at the convolutional layers to extract features from local neighborhood on feature maps in the previous layer. Next, the additive bias is applied and the result is passed through the sigmoid function. The value of an unit at position *(x,y)* in the j^{th} feature map in the i^{th} layer, denoted as Val_{ij}^{xy}, is given by Equation (2),

$$Val_{ij}^{xy} = \tanh\left(B_{ij} + \sum_{n}\sum_{a=0}^{A-1}\sum_{b=0}^{B_1-1} W_{ijn}^{ab} V_{(i-1)n}^{(x+a)(y+b)}\right) \tag{2}$$

where, tanh (.) is the hyperbolic tangent function, B_{ij} is the bias for this feature map, n is the indexes over the set of features maps in the $(i-1)^{th}$ layer connected to the current feature map, W_{ijn}^{ab} is the value at the position (a,b) of the kernel connected to the kth feature map, A_i is the height of the kernel and B_i is the width of the kernel.

By extending this 2D convolution, 3D convolution is applied on 3D images. This is achieved by considering the depth axis in Equation (2), is given by Equation (3),

$$Val_{ij}^{xyz} = \tanh\left(B_{ij} + \sum_{n}\sum_{a=0}^{A-1}\sum_{b=0}^{B_1-1}\sum_{c=0}^{C-1} W_{ijn}^{abc} V_{(i-1)n}^{(x+a)(y+b)(z+c)}\right) \tag{3}$$

Where, C_i is the size of the 3D kernel along the temporal dimension and W_{ijn}^{abc} is the *(a, b, c)*th value of the kernel connected to the mth feature map in the previous layer.

After extracting the features, initially, a technique or an algorithm is applied on the neural network known as Back propagation [9]. A neural network is a set of connected input/output units in which each connection associated along with weights. During the learning phase, the network learns by adjusting the weights so as to be able to predict the correct class label of the inputs. Well-suited for continuous-valued input and outputs. Its algorithms are inherently parallel used to speed up the computation process. The multilayer feed forward networks is a type of neural network on which the back-propagation algorithm performs. This network is given with enough hidden units and enough training samples, can closely approximate any function. The Online Error Backpropagation which exactly same as the backpropagation but it indulges itself with online. The Backpropagation follows the steps as below:

1. Initialize the weights to small random numbers (-1.0 to 1.0 or -0.5 to 0.5).
2. Propagate the input through bias. The bias acts as threshold in that it serves to vary the activity of the units.
3. Backpropagate the error by updating the weights and biases to reflect the error of the network's prediction.
4. Terminating condition occurs when the prespecified number of epochs has been expired.

While training the neural network system, the convolution and sub-sampling is applied on each alternate layers in the network. At the final output layer, it is necessary to minimize and regularize the trainable parameters. The parameters are initialized randomly and trained using stochastic diagonal

Levernberg-Marquardt method in which each parameter uses the diagonal terms of an estimate of the Gauss-Newton method.

Through seeking the knowledge from the related works, an innovative idea emerges and that is to be discussed in the following section.

4. Proposed Work

In this section, each step of the proposed method is described in detail through Figure 2. As per the Figure 2, the real world video is fed into the system. After that, the input video is split into the frames and pre-processing are done. During pre-processing, the frames are free from noise and also background subtraction is to be done here. Only the pre-processed frames are allowed into the system. The features are extracted by using the 3D Convolution.

Construct the 3D model for human action recognition to predict humans' specific or suspicious actions in outdoors. In this work, to treat the video frames as still images and apply CNNs to recognize actions at the individual frame level. To effectively incorporate the motion information in video analysis, propose to perform 3D convolution in the convolutional layers of CNNs so that discriminative features along with the spatial and the temporal dimensions are captured.

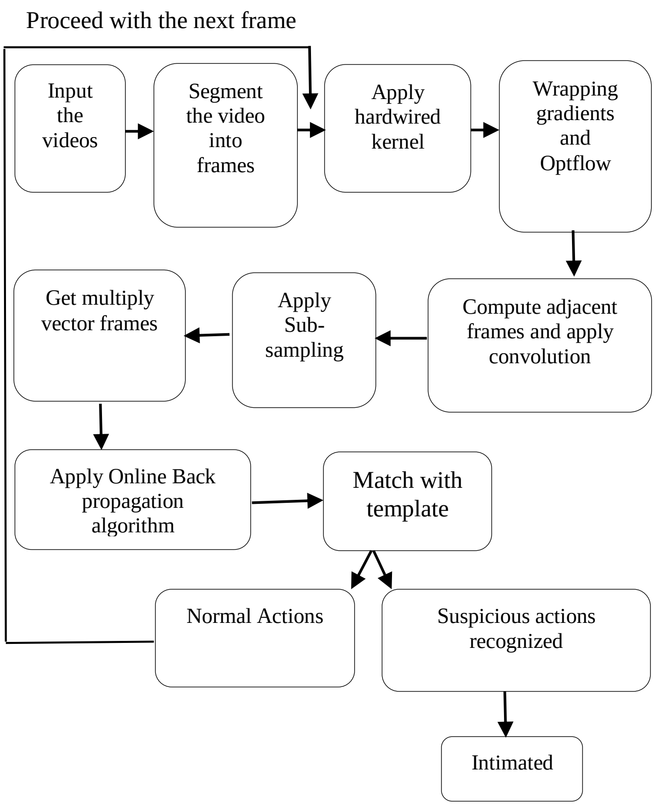

Figure 2. Work flow of the proposed model.

The developed model generates multiple channels of information from adjacent video frames and performs convolution and subsampling separately in each channel. Then, it construct features from both spatial and temporal dimensions by performing 3D convolutions. The developed model regularization and combination schemes to further improve the model performance. Its benefits to apply this model for huge amount of data and large size videos. By applying multiple distinct convolutional operations at the same location on the input, multiple types of features can be extracted.

- *Feature extraction using 3D Convolution*: is achieved by convolving a 3D kernel to the cube formed by stacking multiple contiguous frames together. By this construction, the feature maps in the convolution layer are connected to multiple contiguous frames in the previous layer, thereby capturing motion information. 3D convolutional kernel can only extract one type of features from the frame cube since the kernel weights are replicated across the entire cube. A general design principle of CNNs is that the number of feature maps should be increased in late layers by generating multiple types of features from the same set of lower level feature maps.
- *Action or Motion Regularization*: Computing motion features from a large number of frames and regularizing the 3D CNN models by using these motion features as auxiliary outputs. In particular, for each training action generate a feature vector encoding the long-term action information beyond the information contained in the input frame cube to the CNN. This is achieved by connecting a number of auxiliary output units to the last hidden layers of CNN and clamping the computed feature vectors on the auxiliary units during training. By regularizations, consistent performance is achieved. *Combination of outputs: C*onstructing multiple 3D CNN models with different architectures are to be developed, hence capturing potentially complementary information from the inputs. In the prediction phase, the input is given to each model and the outputs of these models are then combined. By this method the performance of CNN model improved.
- *Implementation Design*: is done in MATLAB. All the model parameters are randomly initialized and are trained using the stochastic diagonal Levenberg Marquardt method. In this method, a learning rate is computed for each parameter using the diagonal terms of an estimate of the Gauss-Newton approximation to the Hessian matrix on randomly sampled training instances.

After an idea emerges for developing the proposed model, several experiments have to be done for training and for testing the system. This is followed by experimental setup.

5. Experimental Setup and Results

The Proposed real time Suspicious Human Action Recognition System was experimented under certain hardware and software constraints or requirements. All the experiments have been conducted on an Intel(R) Core Trademark (TM) i5 Central Processing Unit (CPU) M 480 @ 2.67GHz processor and 4 GB of Random Access Memory (RAM) using MATLAB tool implementation on Microsoft Windows 7 Ultimate 32-bit operating system environment.

The main focus is on the real time video data to evaluate the developed 3D CNN models for action recognition in surveillance videos. Meantime, perform experiments on the Weizmann datasets. Testing the proposed method on these publicly available data sets that are widely used in action recognition. The performance on these benchmarks are saturating and achieves near-perfect results.

As the videos were recorded in real-world environments, and in each frame contains multiple humans, so there arrives a need for human detector and a detection-driven tracker to locate human heads. The well detailed procedure for tracking humans is described in here. Followed by the detection and tracking of humans, a bounding box for each human that performs an action was computed.

Firstly, break each video into its constituent frames and apply bounding box on each frame to reduce the input dimension size. The Weizmann human action dataset contains 83 videos which consists of 9 different people, each performing 10 different actions: jumping, jumping jack, jumping in space, bending down, skipping, walking, running, galloping sideways, waving one hand and waving both hands. There are 93 sequences with resolution 180X144. This resolution is to be reduced into the resolution of 64X48. A subsequence of 13(64X48X13) consecutive frames with 12 frames overlap is given as input to 3D Convolutional Neural Network. Until now, the test have been done on silhouette frame of videos.

The recognition rate is defined as the percentage of correctly recognized actions from the number of all samples, which is given by Equation (4),

$$\mathrm{Re}\,cog_{Rate} = \frac{N_R}{N} \times 100 \qquad (4)$$

To estimate the performance of the proposed system, it is necessary to calculate the percentage error which is given by Equation (5),

$$\%Error = \frac{Error}{Target} \times 100 \qquad (5)$$

Table 1. Shows the performance of human detection and tracking on the real input videos through precision and recall from four different cameras.

No. of cameras	C1 %	C2 %	C3 %	C4 %	AVG %
PRE	65.36	56.36	77.56	56.32	63.9
REC	60.58	30.39	59.45	68.36	54.69

Table 2. Shows precision and recall for head motion detection in ten different videos.

VIDEOS	PRECISION (%)	RECALL (%)
1	82	85
2	84	87
3	85	84
4	87	83
5	84	85
6	84	89
7	80	83
8	80	85
9	82	84
10	83	86

The work is mainly based on the real time dataset of an educational institution. On the available video data, the experiments are done for training and testing the system with various input videos. By this work, it is able to recognize the human actions which needs attention immediately. For instance, in Figure 3, *PhoneToEar* under the college premises is not allowed, so this action to be notified. The actions which are not abide by the rules are to be intimated for the immediate response.

Thus, the several experiments are conducted to train the system which yields the testing results. The analysis of the results are done in the following section.

Table 3. Shows precision and recall for contact detection from eight different videos.

VIDEOS	PRECISION (%)	RECALL (%)
1	90	94
2	89	85
3	93	87
4	92	89
5	87	86
6	84	91
7	81	94
8	86	87

Table 4. Shows the best case and worst case accuracy rates for detection of actions.

Actions to match with template	Best Case	Worst Case
Introduction to a new person	82	62
Standing on the hand resting wall	89	67
Face and Hand Detection	80	63
Multi-person Scenario	48	35

Figure 3. Sample human actions from camera numbers 1, 2, and 3 (left to right).

6. Performance Analysis

The performance of the Human Action Recognition system in real time is evaluated from the following aspects:

6.1. Human Motion Detection

To find the humans in the videos, features like any movements of the humans in each frame is to be identified. Using threshold, to detect the suspectable actions of the human set as a keyhole. Precision is the ratio of the number of humans identified correctly by the system to the total number of humans identified. Recall is the ratio of the number of humans identified by the system to total number of humans in the video data. Figure 4 shows the Precision and Recall for human motion detection for the input test video data. From the analysis of Figure 4, which have the highest precision and recall rate achieves in video number 4 and 6 respectively. Meantime, its lowest precision and recall rate at video number 7 at both cases.

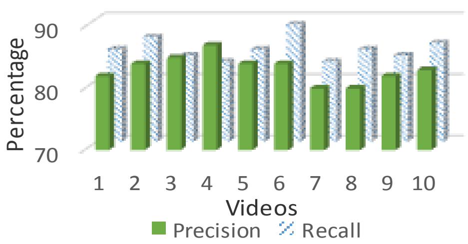

Figure 4. Shows the precision and recall for human motion detection.

6.2. Contact Detection Among Humans

Among the contact detection, the actions of the limbs are to be considered. So, there is a need of motion detection and edge detection. Here, Precision and Recall is done for hand and leg movements of humans. Figure 5 shows the Precision and Recall for contact detection among humans has been done for the sample video data. By analysing this Figure 5, both the precision and recall for the video1, is excellent with 90 and 94 percent. For the video 5, the recall met to its extreme downfall. The precision starts towards its down face from the video3 and gains to up face in video 7.

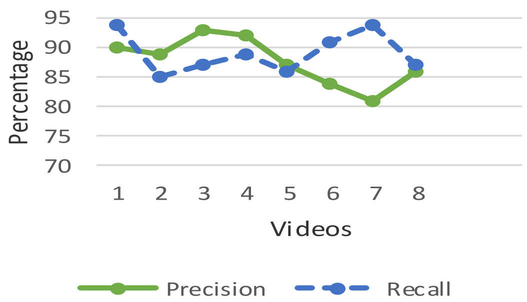

Figure 5. Shows the precision and recall for contact detection among humans.

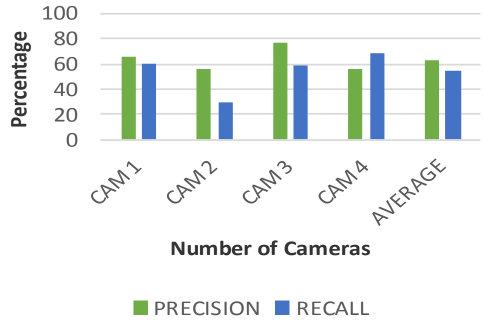

Figure 6. Shows the performance of human detection and tracking on the real time video.

By analysing the Figure 6, it is clear that the precision and recall rate varies for different videos from different cameras and its average are also taken.

The following Figure 7 represents the overall performance of the proposed system with respect to the detection and recognition of actions. For multi-people scenario, the reduction in accuracy rate is much higher and rest of the actions it sounds good.

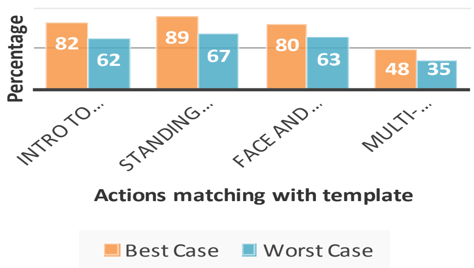

Figure 7. Shows the accuracy rate for detected actions.

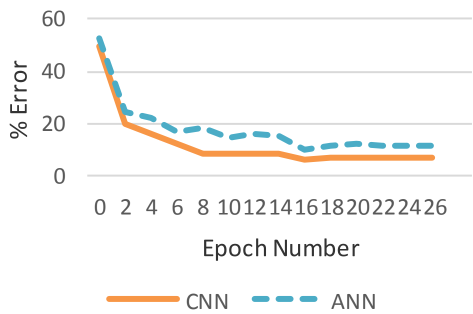

Figure 8. Shows the variation while training the system using CNN and ANN through epochs.

By analysing the performance for various human actions with different methods are shown in the above charts.

7. Conclusions and Discussions

Human action recognition is an active area of research in the field of Computer Science which has been motivated by the need for fast video indexing application and reliable automated video surveillance systems. Current methods for human action recognition mainly use texture descriptors and image processing. Despite their high performance, these methods are highly problem dependent. In this work, neural networks are considered which automatically build high level representation of raw input without any pre-processing.

Considering the Convolutional Neural Network model for action recognition which suits well in situation of open area under surveillance. There are also other deep architectures, such as the deep belief networks, which achieve promising performance on object recognition tasks. Thus, the developed 3D CNN model trained using backpropagation and supervised algorithm which is capable to improve the performance of about 88-90% in predicting and recognizing the suspicious human actions within a short period of time. Prior studies are to be done in training the system using unsupervised algorithm. This work is suitable for many surveillance environment with the intention to safeguard the public. A point to be noted that certain actions done by the persons are taken as templates which depends on the situations or surroundings they belong to.

Acknowledgement

The authors wish to express their gratitude to the Chairman, the Principal and the Director of Adithya Institute of Technology, Coimbatore, India for providing the datasets and all other facilities in completing the research work.

References

[1] Barnachon M., Bouakaz S., Boufama B., and Guillou E., "Ongoing Human Action Recognition with Motion Capture," *Pattern Recognition*, vol. 47, no. 1, pp. 238-247, 2014.

[2] Charara N., Jarkass I., Sokhn M., and Khaled O., "ADABeV: Automatic Detection of Abnormal Behavior in Video Surveillance," *in Proceedings of International Conference on Image, Signal and Vision Computing*, Oslo, pp.172-178, 2012.

[3] Duffner S. and Garcia C., "An Online Backpropagation Algorithm with Validation Error-Based Adaptive Learning Rate," *in Proceedings of the 17th international conference Artificial Neural Networks*, Berlin, pp. 249-258, 2007.

[4] El-Sisi A., "Design and Implementation Biometic Access Control System using Fingerprint for Restricted Area Based on Gabor Filter," *The International Arab Journal of Information Technology*, vol. 8, no. 4, pp. 355-363, 2011.

[5] Fan J., Xu W., Wu Y., and Gong Y., "Human Tracking using Convolutional Neural Networks," *IEEE Transactions on Neural Networks*, vol. 21, no. 10, pp. 1610-1623, 2010.

[6] Gowsikhaa D., Abirami S., and Baskaran R., "Automated Human Behaviour Analysis from Surveillance Videos: A Survey," *Artificial Intelligence Review*, vol. 42, no. 4, pp. 747-765, 2014.

[7] Gowsikhaa D., Manjunath M., and Abirami S., "Suspicious Human Activity Detection from Surveillance Videos," *International Journal on Internet and Distributed Computing Systems*, vol. 2, no. 2, pp. 141-148, 2012.

[8] Guo Y., Chen Y., Tang F., Li A., Luo W., and Liu M., "Object Tracking using Learned Feature Manifolds," *Computer Vision and Image Understanding*, vol. 118, pp. 128-139, 2014.

[9] Han J. and Kamber M., *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publisher, 2006.

[10] Huo F. and Hendriks E., "Multiple People Tracking and Pose Estimation with Occlusion Estimation," *Computer Vision and Image Understanding*, vol. 116, no. 5, pp. 634-647, 2012.

[11] Huo F. and Hendriks E., "Real Time Multiple People Tracking and Pose Estimation," *in Proceedings of the 1st ACM International Workshop on Multimodal Pervasive Video Analysis*, Firenze, pp. 5-10, 2010.

[12] Iosifidis A., Tefas A., and Pitas I., "Minimum Class Variance Extreme Learning Machine for Human Action Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 11, pp. 1968-1979, 2013.

[13] Ji S., Xu W., Yang M., and Yu K., "3D Convolutional Neural Networks for Human Action Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 221-231, 2013.

[14] Oluwatoyin P. and Wang K., "Video-Based Abnormal Human Behavior Recognition-A Review," *IEEE Transactions on Systems, Man and Cybernetics-Part C: Applications and Review*, vol. 42, no. 6, pp. 865-878, 2012.

[15] Paul M., Haque S., and Chakraborty S., "Human Detection in Surveillance Videos and its Applications-A Review," *EURASIP Journal on Advances in Signal Processing*, pp. 1-16, 2013.

[16] Yao B., Liu Z., Nie B., and Zhu S., "Animated Pose Templates for Modelling and Detecting Human Actions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 3, pp. 436-452, 2013.

Sathyashrisharmilha Pushparaj received bachelors degree and masters degree in Computer Science and Engineering from Anna University of Chennai, India, in 2012 and 2014 respectively. Her current research interests include computer vision, pattern recognition and machine learning.

Sakthivel Arumugam is currently working as a Professor in Department of Information Technology, Woldia University, Woldia, Ethiopia. He has 15 years of experience in research and teaching. He obtained his BE, ME and Ph.D. degrees in Computer Science and Engineering. His areas of interest are mobile computing, soft computing, Cloud Computing, Green Computing and Security models. He has published 10 papers in international journals and 5 papers in international/national conferences. He is reviewer of IAJIT and IEEE and got the best active reviewer award twice. He reviewed two text books published by Oxford University Press and TMGH in India. He is a life member of ISTE, ACCS and IAENG.