# A CNN Based Approach for Crowd Anomaly Detection

Kinjal V Joshi
and
Narendra M Patel
Gujarat Technological University , India

Automatic Anomaly detection in a crowd scene is very significant because of more apprehension with people's safety in a public place. Because of usefulness and complexity, currently, it is an open research area. In this work, a new Convolutional Neural Network (CNN) model is proposed to detect crowd anomaly. Experiments are carried out on two publicly available datasets. The performance is measured by Accuracy and Area Under the ROC Curve (AUC). The experimental results determine the efficacy of the proposed model.

Keywords: Anomaly Detection, Convolutional Neural Network, Crowd Analysis

## 1. INTRODUCTION

Nowadays, with the increasingly growing needs of protection of people and personal properties, video surveillance has become a big concern of everyday life. A consequence of these needs has led to the deployment of cameras almost everywhere, which produce a large quantity of video. Most existing video surveillance systems are fully supervised by humans. Video monitoring is a very cumbersome and time-consuming task. It is impossible for a human to handle and find out the abnormal events. However, even one small mistake could cause an unacceptable loss. Thus, it is important to construct an automatic video abnormal detector, dealing with a large quantity of video frames and alerting people for a punctual and functional response when an anomaly happens. So, Lots of research on automatic video surveillance is going on.

Abnormal event detection is considered as one module of Crowd Analysis Junior et al. [2010]. The applications of Crowd Analysis are Crowd Management, Public Space Design, Virtual Environments, Visual Surveillance, and Intelligent Environments. Crowd analysis can be categorized into four modules- People detection and tracking, People counting, Behavior analysis and Abnormality detection. Recent research Direkoglu et al. [2017] reports that 30% of the research work of complete crowd analysis is done in the Abnormality detection module.

Automatic Anomaly detection in a crowd scene is a very challenging task because the definition of anomaly is subjective or context-dependent. The event which occurs rarely can be considered as an abnormal event. Anomaly detection is the problem of complex sequential visual patterns recognition, as the crowd scene contains occlusion and clutter. The conventional techniques use traditional low-level features like optical flow and Histogram of Oriented Gradients for this task, which are ineffective at identifying such complex patterns and also hard to implement in real-time applications. Many researchers have used unsupervised learning approach like auto-encoder and supervised learning approaches like CNN or SVM classifiers to detect anomaly in a crowd scene. Convolutional Neural Network gives very good performance on various computer vision areas like video summarization, behaviour recognition, security, object tracking, etc. Inspired by the performance of CNNs in the mentioned domains, the authors have proposed a CNN based approach to detect anomaly in a crowd scene. Here CNN is used as a classifier. It is not possible to include all types of abnormal behaviours to train CNN, so general-purpose abnormality detection might not possible, but as per requirements at a particular place, it can give a better performance compared to other methods. For example, in some places only pedestrians are allowed, so vehicle entry is abnormal. At some paid entry points, entry without making a payment is considered as abnormal behaviour. In some places like banks or shops, robbery can be considered as an abnormal event. In the examination hall, talking or trying to copy the answer from the other students answer sheet is abnormal behaviour. For surveillance cameras that are positioned on a road, accident, crowd formation, suddenly running, fighting, etc. can be considered as abnormal behaviour. So, as per the case, it is possible to generate labels of abnormal events, and supervised learning can be used. Currently, many real-world datasets for anomaly detection in a crowd scene are publicly available, as it is an open research problem. In this paper two datasets namely UMN [] and Airport-WrongDir, Zaharescu and Wildes [2010] are used, in which the two abnormal events, running and moving in the wrong direction are recognized respectively. The CNN based techniques which already exist either use a pre-trained CNN model or apply lots of efforts for pre-processing. The key contributions of the proposed method are summarized as follows.

a) In this paper, authors have proposed a new CNN architecture to detect crowd anomaly.
b) The proposed method performs better than state-of -the-art methods. Without any pre-processing, the proposed model gives very good Accuracy percentage and AUC value on publicly available datasets.

The rest of the paper is ordered as follows. The recent work done by different researchers for crowd anomaly detection is presented in section 2. The proposed model for anomaly detection is described in section 3. Section 4 represents information of the used datasets and implementation results. The conclusion is discussed in section 5.

## 2. RELATED WORK

The definition of an abnormal event is subjective. Machine learning and deep learning algorithms can be applied to detect a particular event in a specific dataset. It is a very challenging task to detect all types of abnormality in video surveillance. So, lots of research is going on in this area. In this section, different approaches based on handcrafted features and deep learning are mentioned. In section 4, accuracy or area under the ROC curve (AUC) of each method is presented.

### 2.1 Handcrafted features based approaches

R et al. [2009] uses the social force model and classify frames as normal and abnormal by using a bag of words approach. Cem Direkoglu et al. [2017] considers the direction and magnitude difference computed between the optical flow vectors in the current frame and the former frame at each pixel location, that improves accuracy over pure optical flow features. Wu et al. [2014] have proposed a Bayesian model for crowd behavior recognition. Chen and Huang [2011] have used optical flow to cluster human crowds into groups in an unsupervised manner using the method of adjacency-matrix based clustering (AMC). The behaviors of cluster crowd with attributes-orientation, position, and crowd size, are characterized by a model of force field. Wu et al. [2010] have used a chaotic feature set to train the probabilistic model. Cong et al. [2013] have proposed the sparse reconstruction cost (SRC) over the standard dictionary to measure the normalness of the testing sample. M et al. [2017] use optical flow for computation of crowd motion energy. The crowd motion energy is further modified by crowd motion intensity (CMI). The peaks in the CMI characteristics are the indicators of abnormal activity. In this method, the proper threshold plays a significant role in abnormal activity detection. X et al. [2018] have used contextual gradients between two local regions and then have constructed a histogram of oriented contextual gradient (HOCG) descriptor for abnormal event detection based on the contextual gradients. In the method proposed by XuguangZhang et al. [2018], the crowd behaviour is analyzed according to the change of the consistency, entropy, and contrast-three descriptors for co-occurrence matrix. Here Threshold judgment of three descriptors is critical to get an accurate result. Yin et al. [2015] have utilized the Local Binary Pattern Co-Occurrence Matrix (LBPCM) for crowd density estimation.

They have adopted high accuracy optical flow histograms of the orientation of interaction force to extract the crowd dynamic information (HOIF). Then SVM is used to detect an abnormal event. Benabbas et al. [2011] have used the optical flow field with block clustering to identify an unusual event in a crowd.

### 2.2 Deep learning-based approaches

Mostafa et al. [2017] have used a motion heat map to find the region of interest. After identifying the motion structure, different classifiers are used in which the CNN classifier gives a good result to detect anomaly in a crowd scene. In M. and R [2017], the authors have presented the cubic-patch based deep cascade method, characterized by a cascade of classifiers. This method has two main stages. In the first stage, a 3D auto-encoder is used for the detection of normal cubic patches, and interest points are selected from the remaining pieces and fed into a deeper 3D CNN. The deep auto-encoder and CNN are further subdivided into multiple sub-stages, which acts as a cascaded classifier. The shallow layers detect simple normal patches and deeper layers detect more complex normal patches. N [2018] proposed that keeping track of the changes in the CNN feature across time can facilitate capturing the local abnormality. The paper proposes a novel method of utilizing the semantic information from CNN-based layers and the low-level optical flow. The authors employed a Fully Convolutional Network as a pre-trained model and plugged an active binary quantization layer as the final layer to the net. They have captured temporal CNN patterns to represent motion in a crowd. Sabokrou et al. [2018] presented an efficient method for detection and localization of anomalies in videos using a fully convolutional neural network (FCN) and temporal data. Authors have used pre-trained CNN, i.e. AlexNet to extract shape and motion features and a new convolutional layer is added where kernels are trained concerning the chosen training video. Y et al. [2017] proposed a PCANet-based anomaly detection method, in which video events are represented in unsupervised fashions. The proposed approach extracts appearance and motion features simultaneously using the PCANet from 3D gradients. Additionally, the normal and abnormal event patterns are modeled by constructing a deep Gaussian mixture model (GMM) with observed regular events.
Z [2016] presented a Spatio-temporal CNN for anomaly detection from video sequences. The proposed architecture uses features from both spatial and temporal dimensions to extract both the appearance and motion information encoded in continuous frames. B [2017] have used pre-trained CNN to extract deep features, and then one class support vector machine classifier is trained to learn a model for normal events in the video. In C [2019], the authors proposed Deep One-Class (DOC) model, which integrates the one-class Support Vector Machine (SVM) into Convolutional Neural Network (CNN) for abnormal event detection from surveillance videos. High-level features are extracted using CNN. One-class SVM layer is used for classification and is also used to optimize parameters of the CNN model. Hinami R.and Mei T.and Satoh [2017] undertook the problem of joint detection and recounting of abnormal events by integrating a generic CNN model and environment-dependent anomaly detectors. The proposed approach first learns CNN with multiple visual tasks to exploit semantic information that is useful for detecting and recounting abnormal events. Then the model is plugged into anomaly detectors. In Xu et al. [2019] Authors have used variational auto-encoder to learn the appearance and motion features of the receptive fields and multiple Gaussian models are used to predict the anomaly scores of the corresponding receptive fields. In Smeureanu et al. [2017], Authors have used a pre-trained CNN model to extract features and One-class SVM for classification.

## 3. THE PROPOSED METHOD

The figure-1 represents the proposed CNN model. The size of the input image is taken 128x128.

The used datasets contain colour video, so the input layer size is 128x128x3. Literature about CNN sys, to extract large objects, kernel size should be large, and to detect small objects kernel size should be small. Here three convolution layers are used. For each convolution layer, biases

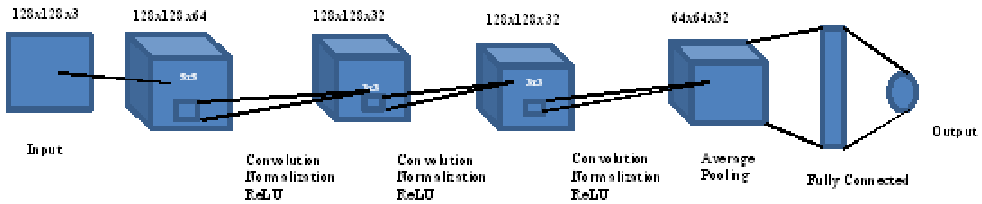


Figure 1.The Proposed Model

and weights are initialized with the Glorot initializer Glorot et al. [2010].Glorot et al. [2010] had done a study on how activations and gradients vary across layers during training and proposed a new initialization scheme that brings substantially faster convergence. So, As per Glorot et al. [2010] biases are initialized to be 0 and the weights Wij with the following commonly used heuristic:

$$Wij \sim U[-\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}]$$

where U[-a,a] is the uniform distribution in the interval between -a and a, and n is the size of the previous layer. Initially, Authors have taken large size kernels, so coarse details are extracted. Then kernel size is decreased to extract fine details. Initially, for the first convolution layer, kernel size is taken 5x5 with 64 channels. The activation function is used to decide the particular neuron would fire or not. Rectified Linear Unit is a widely used activation function in CNN because it does not activate all neurons at the same time. It is defined as y = max(0, x). It converts all negative inputs to zero and the neuron does not get activated. This makes it very computational effective as few neurons are activated per time. It does not saturate in the positive region. ReLU converges faster than other activation functions. It doesnt use complicated math. So here, ReLU is used as an activation function. A batch normalization layer normalizes each input channel across a mini-batch, so the learning process can be stabilized. Here batch size is taken 100. By the use of the batch normalization layer, learning rate value can be set high and the number of epochs can be reduced, so that model takes less time to be trained. So, to speed up the training of convolutional neural networks and reduce the sensitivity to network initialization, the batch normalization layer is used between the convolutional layer and nonlinearities, such as ReLU layers. The batch normalization layer also fulfils the goal of the dropout layer i.e. removing overfitting, so the dropout layer is not used in the proposed model. The same process is repeated twice with the kernel seize 3x3 and the number of channels 32 as shown in figure-1. Convolutional layers summarize the presence of features in an input image. The output feature maps are sensitive to the location of the features in the input. They can be more robust to changes in the position of the feature in the image by down sampling. Pooling layers provide an approach to down sampling feature maps by summarizing the presence of features in patches of the feature map. In the proposed model average pooling layer is used, which summarizes the average presence of a feature. The Average Pooling layer is used with pool size 2 and stride 2, So the new image size is 64x64x3. To generate a successful model, various parameters are tuned like Number of layers, Order of layers, Number of channels, Kernel size, Initial learning rate, Learning rate drop factor, Learning rate drop period, Epochs, and Batch Size. The learning rate is the most crucial parameter for model building. It specifies how fast the network learns. Generally, its value lies between 0 to 1. Here the initial learning rate value is taken 0.01.
Here three convolution layers are used with a different number of channels. For convolution layer-1, 2, and 3, there are 64, 32, and 32 activation channels for each input image respectively. The features learnt by the network can be seen by comparing areas of activation with the original image. The channels in earlier layers learn simple features like colour and edges, while channels

in the deeper layers learn complex features. The extracted features are shown in section 4. By identifying features in this way, It can be partially understood what the network has learned.

## 4. EXPERIMENTAL RESULTS

In this section, The authors have represented dataset description, intermediate and final results of experiments, and comparison between the proposed method and state-of-the-art methods using the same dataset.

### 4.1 Dataset Description

In this work, two publicly available datasets, namely UMN (University of Minnesota) and Airport-WrongDir are used. In both datasets ground truth and sufficiently abnormal test images are available. The UMN Dataset contains two outdoor and one indoor video samples (320 240 pixels image resolution). Each video consists of an initial part of walking as a normal state and ends with sequences of running as an abnormal state. The Airport -WrongDir dataset contains one video sample (300 300 pixels image resolution) in which people moving from right to left is considered as a normal event and anyone moving in the wrong direction i.e. left to right is considered as an abnormal event.

### 4.2 Results

The implementation is done with MATLAB2019a, 16 GB RAM, i9 Processor CPU machine. For both datasets, Authors have used 80% Training and 20% test data. The datasets are having enough video frames for a deep learning-based approach and are also easily handled with CPU. The images of both datasets are resized to 128 x 128. The same proposed model with the same tuning parameters gives excellent accuracy for both datasets. The figure-2 represents the most significant activation of each of the three convolution layers for an image of each dataset. In the activation channel image, white pixels represent strong positive activations, and black pixels represent strong negative activations. A channel that is mostly gray does not activate as strongly on the input image. The position of a pixel in the activation of a channel corresponds to the same position in the original image.

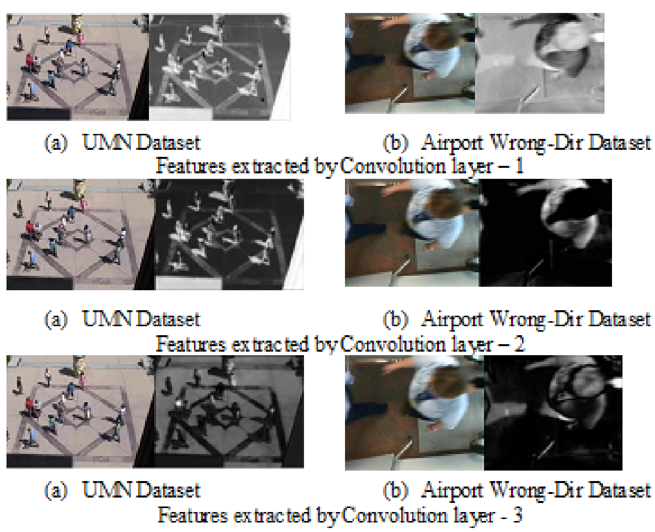


Figure 2. Largest Activation at each convolution layer

The figure-3 shows the training progress of the CNN model. This figure is useful to monitor the progress of the model training. By this figure it can be determined how quickly the network accuracy is improving, and whether the network is starting to overfit the training data. If training accuracy is much higher than test/validation accuracy then overfitting is considered. In these experiments training accuracy and test accuracy both are the same, so overfitting is not the case. The figure represents two graphs, Accuracy vs Iterations and Loss vs Iterations. The figure displays training metrics at every iteration. Each iteration is an estimation of the gradient and an update of the network parameters. In these experiments, the number of epochs is 20, so in the training progress graph, 20 white and gray columns are visible alternatively. In figure 3 (a) which shows the training progress of the model using the UMN dataset, it can be seen that after 2 epochs accuracy is stable. While figure 3(b), which shows the training progress of the model using the Airport-WrongDir dataset, represents that after 4 epochs accuracy is not going to be updated.

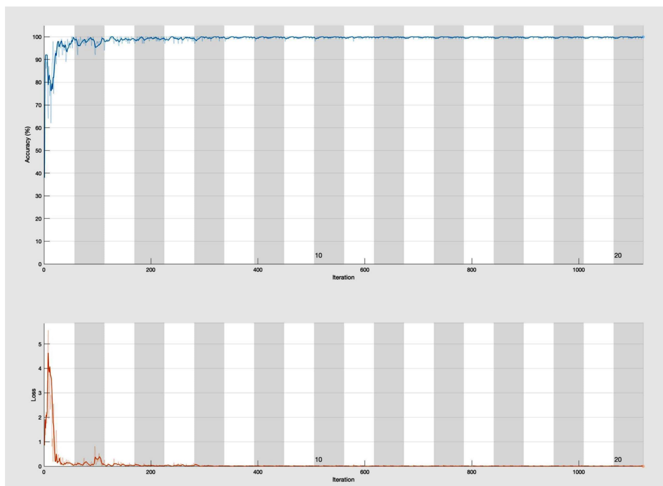


Figure-3 (a)Training Progress for UMN dataset

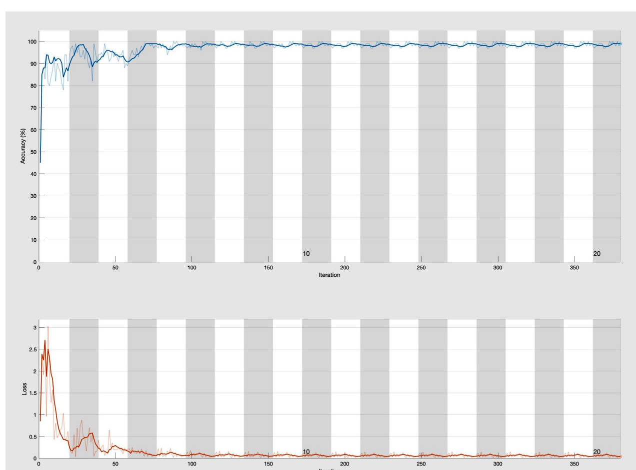


Figure-3 (b) Training Progress for Airport-WrongDir dataset

The performance of the proposed model is represented by accuracy and Area Under the ROC Curve (AUC). A receiver operating characteristic (ROC) curve, is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. The

ROC space is defined by False Positive Rate and True Positive Rate as x and y axes, respectively, which depicts relative trade-offs between true positive and false positive. The best possible prediction method would yield a point in the upper left corner or coordinate (0,1) of the ROC space, representing 100% sensitivity (no false negatives) and 100% specificity (no false positives). The (0,1) point is also called a perfect classification. The value of AUC lies between 0 to 1. The value near to1 represents good classification and the value below 0.5 represents bad classification. The table I represents experimental results.

Table I: Experimental Results

| **Dataset** | **Number of Images** | **Accuracy (%)** | **Area Under Curve** |
|---|---|---|---|
| UMN | 7003 | 99.64 | 0.9999 |
| Airport Wrong-Dir | 2392 | 97.28 | 0.9774 |

Figure-4 represents the ROC Curves for the classification results of both datasets. AUC value for the UMN dataset is near to 1 so, the red curve line overlaps the left and top side of the rectangle window in figure-4(a). AUC value for Airport-WrongDir is 0.9774, so the ROC curve can be seen in the window in figure-4(b).

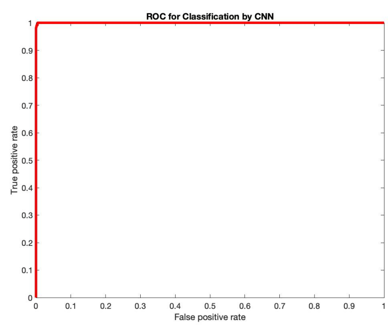


Figure-4 (a) ROC Curve for UMN Dataset

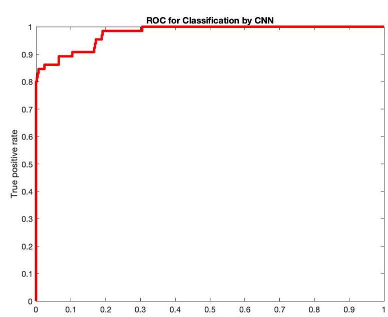


Figure-4 (b) ROC Curve for Airport-WrongDir dataset

The UMN is a widely used dataset for abnormal event detection. The table II represents comparisons between the proposed method and state-of-the-art methods using the UMN dataset. Very less research work is done using the Airport-WrongDir dataset, so the comparison is not represented.

Table II: Comparisons between the Proposed Method and State-of-the-art Methods using UMN Dataset

| **Approach** | **Accuracy (%)** | **Area Under the ROC Curve (AUC)** |
|---|---|---|
| Motion Structure + CNN Classifier Mostafa et al. [2017] | 96.74 | N/A |
| Social Force R et al. [2009] Direkoglu et al. [2017] | 85.09 | 0.96 |
| Pure Optical Flow R et al. [2009] | N/A | 0.84 |
| Novel Optical Flow Direkoglu et al. [2017] | 96.46 | N/A |
| Optical flow + Bayes Classification Li et al. [2015] | 91.57 | N/A |
| High-Frequency and Spatio-Temporal (HFST)Features + Hidden Markov Model Li et al. [2015] | N/A | 0.90 |
| Baysion Model Direkoglu et al. [2017] | 96.40 | N/A |
| Force Field Model Direkoglu et al. [2017] | 81.04 | N/A |
| Chaotic Invariants Direkoglu et al. [2017] Wu et al. [2010] | 87.91 | 0.994 |
| Sparse Reconstruction Cost Direkoglu et al. [2017] Li et al. [2013] | 85.09 | 0.96 |
| Energy Model M et al. [2017] | 91.66 | N/A |
| Histogram of Oriented Contextual Gradient X et al. [2018] | N/A | 0.993 |
| Local Binary Pattern Co-Occurrence Matrix and Histograms of the orientation of interaction force Yin et al. [2015] | N/A | 0.99 |
| Behavior Entropy Model Li et al. [2015] | N/A | 0.893 |
| Commotion Ravanbakhsh et al. [2017] | N/A | 0.98 |
| Plug and Play CNN Ravanbakhsh et al. [2017] | N/A | 0.98 |
| Generative Adversarial Nets (GAN) Ravanbakhsh et al. [2017] | N/A | 0.99 |
| The Proposed CNN Model | 99.64 | 0.9999 |

## 5. CONCLUSION AND FUTURE WORK

In this paper, the authors have proposed the CNN model for abnormal activity detection in a crowd scene. The experiments are done using UMN and Airpot-WrongDir datasets. The performance is measured by accuracy and area under the ROC curve (AUC). The same model gives a worthy performance on both datasets, although normal and abnormal events are defined differently. The used datasets contain enough number of abnormal events and are also easily handled with CPU. In the future, authors have planned to combine more datasets and to develop a generalized model using GPU for detecting all types of unusual events.

### References

Unusual crowd activity dataset of university of minnesota. *http://mha.cs.umn.edu*.

B, S. S. I. R. P. M. A. 2017. Deep appearance features for abnormal behavior detection in video. In *International Conference on Image Analysis and Processing (ICIAP), Lecture Notes in Computer Science, Springer.*

Benabbas, Y., Ihaddadene, N., and Djeraba, C. 2011. Motion pattern extraction and event detection for automatic visual surveillance. *Journal on Image and Video Processing 7*, pp.1–

15.

C, S. J. S. J. H. 2019. Abnormal event detection for video surveillance using deep one-class learning. *Multimedia Tools and Application 78*, pp.36333647.

Chen, D. and Huang, P. 2011. Motion-based unusual event detection in human crowds. *Visual Communication and Image Representation Vol.22,* No.2, pp.178–186.

Cong, Y., Yuan, J., and Liu, J. 2013. Abnormal event detection in crowded scenes using sparse representation. *Elsevier Journal of Pattern Recognition Vol.46,* No.7, pp.1851–1864.

Direkoglu, C., Sah, M., and O'Connor, N. E. 2017. Abnormal crowd behavior detection using novel optical flow-based features. In *IEEE International Conference on Advanced Video and Signal Based Surveillance*.

Glorot, Xavier, and Bengio, Y. 2010. Understanding the difficulty of training deep feedforward neural networks. In *the thirteenth international conference on artificial intelligence and statistics*. pp.249–256.

Hinami R.and Mei T.and Satoh, S. 2017. Joint detection and recounting of abnormal events by learning deep generic knowledge. *arXivpreprint arXiv:1709.09121*.

Junior, J. C. S. J., Musse, S. R., and CludioRositoJungn. 2010. Crowd analysis using computer vision techniques. *IEEE Signal Processing Magazine 27,* 5, pp.66–77.

Li, T., Chang, H., Wang, M., Ni, B., Hong, R., and Yan, S. 2015. Crowded scene analysis: a survey. *IEEE Transactions on Circuits and Systems for Video Technology*.

Li, W., Mahadevan, V., and NunoVasconcelos. 2013. Anomaly detection and localization in crowded scenes. *IEEE Transactions On Pattern Analysis and Machine Intelligence*.

M, H., V, V., and Y, V. 2017. Abnormal crowd behavior detection based on combined approach of energy model and threshold. In *Pattern Recognition and Machine Intelligence, Lecture Notes in Computer Science*. Springer.

M., S. and R, F. M. F. M. K. 2017. Deep-cascade: cascading 3d deep neural networks for fast anomaly detection and localization in crowded scenes. *IEEE Trans. Image Process 26,* 4, pp.19922004.

Mostafa, T. A., Uddin, J., and Ali, M. H. 2017. Abnormal event detection in crowded scenarios. In *3rd International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh*.

N, R. M. N. M. M. H. S. E. S. 2018. Plug-and-play cnn for crowd motion analysis: an application in abnormal event detection. In *the IEEE Winter Conference on Applications of Computer Vision (WACV)*.

R, M., A, O., and M., S. 2009. Abnormal crowd behavior detection using social force model. In *IEEE Conference on Computer Vision and Pattern Recognition*.

Ravanbakhsh, M., Nabi, M., Sangineto, E., Marcenaro, L., Regazzoni, C., and Sebe, N. 2017. Abnormal event detection in videos using generative adversarial nets. In *the IEEE International Conference on Image Processing*.

Sabokrou, M., Fayyaz, M., Fathy, M., Moayed, Z., and Kletted, R. 2018. Deep-anomaly: fully convolutional neural network for fast anomaly detection in crowded scenes. *Journal of Computer Vision and Image Understanding*.

Smeureanu, S., Ionescu, R. T., Popescu, M., and Alexe, B. 2017. Deep appearance features for abnormal behavior detection in video. In *International Conference on Image Analysis and Processing*.

Wu, S., Moore, B. E., and Shah, M. 2010. Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes. In *IEEE Conference on Computer Vision and Pattern Recognition*.

Wu, S., Wong, H. S., and Yu, Z. 2014. A bayesian model for crowd escape behavior detection. *IEEE Transactions on Circuits and Systems for Video Technology Vol.24,* No.1, pp.85–98.

X, H., Y, H., and Q, D. 2018. Abnormal event detection in crowded scenes using histogram of oriented contextual gradient descriptor. *EURASIP Journal on Advances in Signal Pro-*

*cessing Vol.54*.

Xu, M., Yu, X., Chen, D., Wu, C., and Jiang, Y. 2019. An efficient anomaly detection system for crowded scenes using variational autoencoders. *Applied Sciences 9,* 15.

XuguangZhang, Zhang, Q., ShuoHu, ChunshengGuo, and Yu, H. 2018. Energy level-based abnormal crowd behavior detection. *Sensors, MDPI*.

Y, F., Y, Y., and X, L. 2017. Learning deep event models for crowd anomaly detection. *Neurocomputing 219*, pp.548556.

Yin, Y., Liu, Q., and ShibiaoMao. 2015. Global anomaly crowd behavior detection using crowd behavior feature vector. *International Journal of Smart Home(SERSC) 9,* 12, pp.149–160.

Z, Z. S. S. W. Z. D. F. M. W. Y. Z. 2016. Spatial-temporal convolutional neural networks for anomaly detection and localization in crowded scenes. *Sig. Process. Image Commun 47*, pp.358368.

Zaharescu, A. and Wildes, R. 2010. Anomalous behaviour detection using spatiotemporal oriented energies subset inclusion histogram comparison and event-driven processing. In *The 11th European Conference on Computer Vision (ECCV).*

**Prof. Kinjal V. Joshi** is working as an Assistant Professor in Computer Engineering department at G H Patel College of Engineering and Technology, Gujarat, India since 2007. She is pursuing PhD in Computer Engineering from Gujart Technological University under the guidance of Dr. Narendra M Patel. Her research interests are Computer Vision, Artificial Intelligence and Machine Learning. She has published 15 research papers in international journals and conferences.

**Dr. Narendra M. Patel** is working as an Associate Professor in Computer Engineering department at Birla Vishwakarma Mahavidylaya, Gujarat, India. He has 26 years Teaching experience. His research interests are Computer Vision and Machine Learning. He has guided many projects of master and PhD level in this area. He has published more than 100 research papers in international journals and conferences.