Scientific Journal Impact Factor (SJIF) : 3.518
International Journal of Modern Trends in Engineering and Research
ISSN (ONLINE) : 2349 - 9745
ISSN (PRINT) : 2393 - 8161

Detection and Classification of Interesting Events Using ANN Algorithm

Jisha Shajan[1] and Prema Mani[2]
[1, 2]Department of Computer Science, IIET, Nellikuzhi

Abstract- The analysis of motion and behaviors in crowded environment using a surveillance system constitutes an issue for public security. The proposed system focuses on detecting and classifying interesting events in crowded video. A newly introduced method based on swarm theory, histogram of oriented swarms is used to capture the motion of crowded scenes, together with the histogram of oriented gradients are used to extract the appearance characteristics of a video sequence. Both motion and appearance combined to form final descriptor, which effectively characterize each scene and detect anomaly. Because of the infinite number of anomalies that is derived from each scene, it is difficult to provide examples of all possible abnormal classes, so an Artificial Neural Network (ANN) classifier is chosen.ANN aiming to identify any irregularities deviating from the normal pattern. This leads to achieves higher accuracy and classifier capable of detecting different kinds of anomalies at a low computational cost.

Keywords- Anomaly Detection, Crowd Analysis, Histogram of Oriented Gradients, Swarm Intelligence, ANN Classifier.

I. INTRODUCTION

The task of automatic event detection from long duration video has gained an increasing attention. In this work we propose a novel method for automatic anomaly detection and classification. We introduce two descriptors called, Histogram of Oriented Gradients and a Histogram of Oriented Swarms to capture the appearance and motion. After that an ANN classifier is used to classify the abnormal and normal events.

Swarm Intelligence is mainly used to build a novel motion descriptor. Swarm intelligence has been used in optimizing interaction force for global anomaly detection in the past only in the framework of Particle Swarm Optimization (PSO) in [1], where PSO optimizes a fitness function minimizing the interaction force derived from the Social Force Model. In our work, the main concept of swarm intelligence is to monitor the movements in crowded scenes by a swarm of agents flying over them, to capture the dynamics.

Histogram of Oriented Gradients proposed in [6], are used to capture the appearance characteristics of a video sequence even in small areas. HOG effectively captures the local edge and gradient structures.

Then, use the Artificial Neural Network (ANN) algorithm as a classifier to identify and classify any irregularities deviating from the normal pattern. The ANN was chosen, as it known to be best suited for detection and classification.

II. RELATED WORKS

The real world surveillance involve crowds of people, such information extraction is not so easy with traditionally used methods. Existing approaches are classified into two categories one that consider only motion information to detect anomaly and other that use both appearance and motion

Volume 03, Issue 02, [February – 2016] ISSN (Online):2349–9745; ISSN (Print):2393-8161

information to describe the scene dynamics. In the first category, Mehran et al. [2] introduce a Social Force Model (SFM) to describe a crowd's normal behavior based on motion characteristics and completely ignoring the appearance information. In the same category, Cong et al. [3] use a sparse reconstruction cost to measure the normality of the testing sample, considering the dictionary methods.

In the second category, Ito et al. [4] uses both motion and appearance for detecting interesting events by estimation density ratio to classify frames into normal and abnormal, it is suited for detecting events that occur over the entire frame and it misses local anomalies. Another interesting work by Vagia et al. [8] uses a one class classifier called Super Vector Data Description (SVDD) and is used for outlier detection. A common problem of all the methods is the inability to successfully detect and classify anomalies that move similarly to the normal pattern.

III. PROPOSED WORK

The purpose of ANN is to classify abnormal and normal events separately and it effectively identifies any irregularities deviating from normal pattern. In this work our algorithm uses data derived from automatically extracted region of interest (ROI) instead of entire video frames in order to capture anomalies appearing in a small part of the frame. By applying background subtraction, ROIs is extracted. Then the interesting points in them are tracked using the KLT tracker. Swarm intelligence together with the HOG descriptor, forms a new feature capable of successfully determine a region's normality.

Figure 1 shows the workflow of proposed system. First stage is the preprocessing of input frame. Background subtraction is performed in this stage. By using the background subtraction the region of interest is extracted from the frame. Only the region of interest is encoded and processed instead of entire frame, so as to only process pixels containing information relevant to the event taking place. Next is the patch conversion, histogram of oriented data are generated. Using these histograms non overlapping blocks are extracted and extracted blocks is placed in codebook as described in [10]. Next stage is the computation of histogram of oriented gradients and histogram of oriented swarms to extract the appearance and motion and they combined to form the final descriptor, for anomaly detection. Afterwards, an ANN classifier is used to classify the abnormal and normal events. ANN is used, as they generally exhibit good performance relatively to other classifier. Because of the infinite number of anomalies it is difficult to detect all possible anomaly classes, so ANN classifier is chosen.

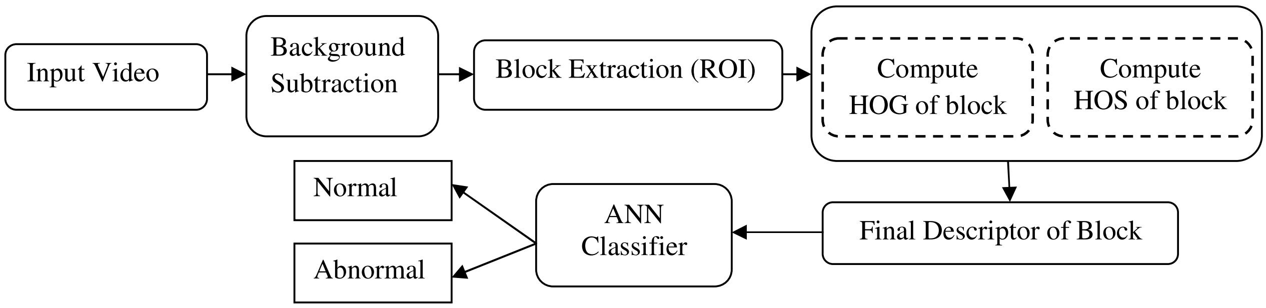

Figure 1: The workflow of anomaly detection and classification

In following subsections we will discuss in detail the HOG descriptor, HOS descriptor, swarm modeling for crowd's dynamics and anomaly detection and classification.

A. HOG Descriptor

To extract the appearance characteristics of a video sequence, the Histogram of Oriented Gradients (HOG) proposed in [6] is used. HOG effectively captures the gradient structure and local edge. The HOG descriptor is applied in ROI blocks and is extracted, so the final descriptor for each block contains temporal information. The computation procedure is as follows: each block k is divided into 4 cells and HOG histogram is calculated for each of them. The block is tracked over the time and the HOG descriptor is computed from the average of the consecutive frames so as to include information and achieve local noise reduction. The final descriptor is the concatenation of a 3 frame average for each cell c in block k:

$$HOG_{j,j+2}^{k}(c) = E\left[HOG_{j}^{k}(c), HOG_{j+1}^{k}(c), HOG_{j+2}^{k}(c)\right] \quad (1)$$

The HOG descriptor for block k, averaged over frames j to $j+2$ will be represented as $HOG_{j,j+2}$ including the average over all 4 cells. The $HOG_{j}^{k}(c)$ symbol represents the HOG histogram, calculated from the c^{th} cell of block k. The process of extracting HOG descriptor is shown in the figure 2.

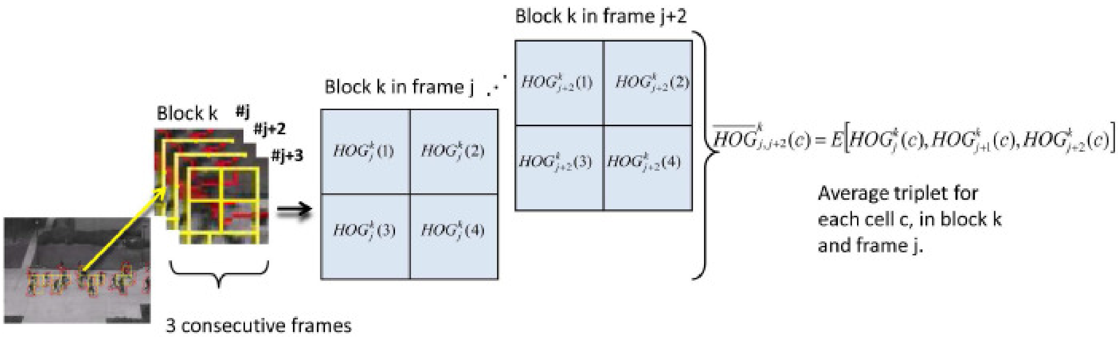

Figure 1: Extraction of HOG descriptor

B.HOS Descriptor

The Histogram of Oriented Swarms is used to capture the dynamics of crowded environments and it is based on the application of swarm intelligence, which is used to build a novel motion descriptor in [5]. The main concept of swarm intelligence is the monitoring of movements in crowded scenes by a swarm of agents flying over the prey, to capture the dynamics in a collective and efficient way.

In our implementation, we use physics-based modeling of crowded scenes as their behavior and characteristics are highly correlated with those of a swarm in nature. The swarm model that is used is based on the general theory described in [7], where the swarm modeling is used to filter noise in images and to better characterize the highly complex motion information from videos of crowds. In our implementation swarm comprise of agents and a prey: the core idea is to construct the prey based on the optical flow values and the agents track the prey, but also interact with each other, as they would in nature.

The first step is the prey generation. Here, we are interested in the extraction of motion using the swarm modeling, so optical flow (OF) value is used as the prey. The prey is tracked by the swarm comprises of optical flow magnitude values of pixels in the ROIs. The ROIs corresponds to rectangular areas containing a fixed number of n pixels. To generate the prey for a ROI of m frames, we consider the pixels of each ROI. In a particular ROI of frame j, each pixel at position i have OF magnitude equal toO_{ij}. The prey's position x_p is determined by the OF magnitude.

Next step is the extraction of agent's position. The agents are subjected to three types of forces: interaction force, friction force, external force. After the extraction of each prey, a swarm of agents is generated to characterize its motion. The agents are groups that we define to track the prey by flying over them. The agents are initially located in random positions, which change according to agent-prey forces, agent-to-agent forces and friction forces described in [8]. Accelerated motion is extracted from the result of these forces [8]. From this motion, the agent's position is extracted. These forces are inspired by the analysis of movements of individuals in crowd[9], matching real world characteristics of people in crowded environment.

By examining the evolution of the agent's position, determined by prey motion patterns and the forces affecting the agents, the final HOS descriptor is formed.

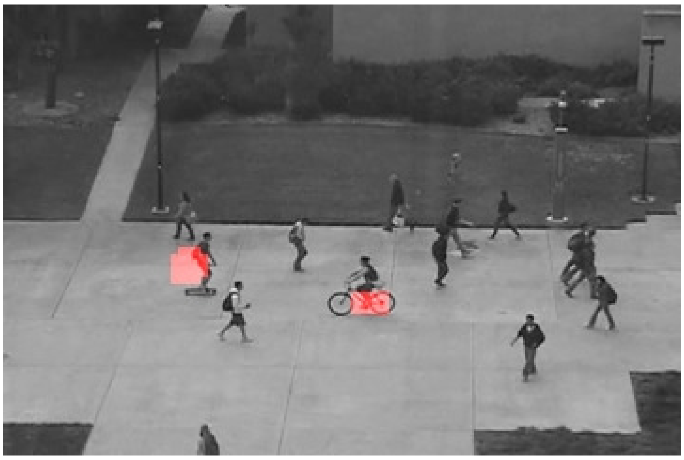

Figure 2: Detection of anomalies: Skater and Bicycle

C. Anomaly Detection and Classification

Appearance descriptor and motion descriptor are combined to form the final descriptor to detect the anomalies. For m frames, average triplets of HOG and HOS are concatenated, resulting in the feature vector of Eq.

$$f = \{ HOG_{1,3} , HOS_{1,3}, \ldots. , HOG_{m-2,m}, HOS_{m-2,m} \}$$

Here, $HOG_{m-2,m}$ is the average of HOG histograms for frames *m-2* to *m* for the corresponding blocks and the average of corresponding HOS histogram is the$HOS_{m-2,m}$. Afterwards, an Artificial Neural Network (ANN) is used for the classification. ANN is a type of network sees the node as 'artificial neurons' and are a form of computation inspired by the structure and function of the brain. An artificial neuron is an information processing unit, fundamental to the operation of a neural network and an ANN is composed of a number of interconnected units, where each unit has an input or output characteristics and implement a function.

An ANN is a data processing system consisting of a large number of simple highly interconnected processing units and the output of any unit is determined by its I/O characteristics, its interconnection with other units and possible external inputs. The neural network architecture consists of three layers: input layer, hidden layer, output layer. ANNs combine artificial neurons in order to process information and the processing element consists of two parts: the first part consists of inputs, which are multiplied with the weights and the second part consists of an activation function as shown in the figure 3.

International Journal of Modern Trends in Engineering and Research (IJMTER)
Volume 03, Issue 02, [February – 2016] ISSN (Online):2349–9745; ISSN (Print):2393-8161

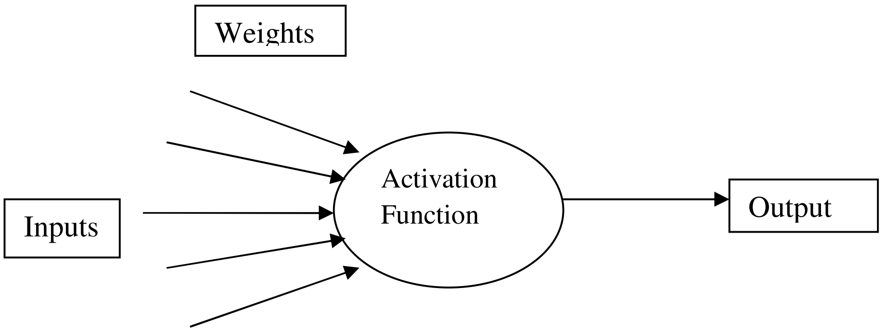

Figure 3: An artificial neuron

The ANN use Feed Forward Neural Networks, to best match input and target. The first step is to collect the data and create a feed forward network, which has input vectors, hidden layers and output vectors. The network receives inputs by neurons in the *input layer*, and the output of the network is given by the neurons on the *output layers*. There may be one or more *hidden layers* to intervene between the external input and the network output. The first layer has weights coming from the input layer and each layer has a weight coming from the previous layer. All layers have biases and the last layer is the network output. Then configure the network and initialize the weights and bias. Each input is weighted with an appropriate *w*. the sum of the weighted input and bias form input to transfer function *f*. neurons can use any differentiable transfer functions like *logsig, trainrp, purelin* to generate output from its net input. There are number of inputs but we get only one output either 1 or 0, the output value 0 indicate the abnormal or 1 indicates the normal object.

The ANN was chosen, as it is aiming to identify any irregularities deviating from the normal pattern. Because of the infinite number of anomalies, it is very difficult to provide examples of all possible anomaly classes, so ANN is chosen to classify the abnormal and normal objects. Our algorithm checks each frame's ROI independently, leads to more accurate and capable of detecting different kinds of anomalies, even when appearing for the first time in the dataset.

IV. RESULTS

The proposed method is applied to the datasets of surveillance to evaluate the effectiveness, where different kinds of anomalies were detected. Fig 2 shows the anomalies detected a bicycle and a skater moving with different speeds and passing through the crowd, respectively are correctly identified as anomalies, which are in cases their detection is a challenging task for a human observer. A remarkable achievement of the proposed system is that deviations from normal patterns can be detected. Our approach exhibits low cost and leads to better and accurate performance.

V. CONCLUSION

In this system, we propose a method for anomaly detection and classification of different scenarios is recorded from the surveillance cameras which are static. Swarm intelligence is used for the extraction of motion characteristics and the appearance is extracted by the histogram of oriented gradients together form the final descriptor, for anomaly detection. The ANN classifier classifies the normal and abnormal objects, which aiming to identify the any irregularities deviating from the normal pattern. The proposed system can be effectively used for challenging crowd videos with many occlusions and local noise. Also, it is capable of detecting different kind of anomalies, even when appearing for the first time in the dataset.

REFERENCES

[1] R.Raghavendra,A.Del Bue,M. Cristani, and V.Murino, "Optimizing interaction force for global anomaly detection in crowded scenes," in Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW), Nov. 2011.
[2] R. Mehran,A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in Proc. IEEE Conf.Comput. Vis. Pattern Recognit. (CVPR), Jun. 2009.
[3] Y.Cong, J. Yuan, and J.Liu, "Sparse reconstruction cost for abnormal event detection."in Proc. IEEE Conf. Comput. Vis. Pattern Recodnit. (CVPR), Jun. 2011.
[4] Y. Ito, K. Kitani, J. Bagnell, and M. Hebert, "Detecting interesting events using unsupervised density ratio estimation," in Proc. IEEE Conf. Eur. Conf. Comput. Vis. Workshop (ECCVW), vol. 7585. Oct. 2012.
[5] V. Kaltsa, A. Briassouli, I. Kompatsiaris, and M. G. Strintzis, "Swarm based motion features for anomaly detection in crowds," in Proc. IEEE Int. Conf. Image Process. (ICIP), Oct. 2014,
[6] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR), vol. 1. Jun. 2005.
[7] H. M. H. Teodorescu and D. J. Malan, "Swarm filtering procedure and application to MRI mammography," Polibits, vol. 42, no. 42, pp. 59–64, 2010.
[8] Vagia Kaltsa and Alexia Briassouli, "Swarm Intelligence for Detecting Interesting Events in Crowded Environments," in Proc. IEEE transaction on Image Processing, Vol. 24, No 7, July 2015.
[9] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," Phys. Rev. E, vol. 51, No. 5.
[10] Kim, K., Chalidabhongse, T.H., Harwood, D., Davis, L, "Real-time foreground-background Segmentation using codebook model," Real-Time Imaging **11**(3), 172–185 (2005).

