CrossMark

Visual data mining for crowd anomaly detection using artificial bacteria colony

Joelmir Ramos[1] · **Nadia Nedjah**[1] ·
Luiza de Macedo Mourelle[2] · **Brij B. Gupta**[3]

Abstract This paper presents a novel method for global anomaly detection in crowded scenes. The optical flow of frames is used to extract the foreground of areas with people motions in the crowd in the form of layers. The optical flow between two frames generates one layer. The proposed method applies the metaheuristic of artificial bacteria colony as a robust algorithm to optimize the extracted layers. Artificial bacteria cover all regions of interest that have high movement between frames. The artificial bacteria colony adapts quickly to the most varied scenarios. Moreover, the algorithm has low sensibility to noise and to sudden changes in video lighting as captured by optical flow. The bacteria population of the colonies, its food storage and the colony's centroid position regarding each optical flow layer, are used as input to train a Kohonen's neural network. Once trained the network is able to detect specific events based on behavior patterns similarity, as produced by the bacteria colony during such events. Experiments are conducted on available public dataset. The achieved results show that the proposed method captures the dynamics of the

✉ Joelmir Ramos
joelmiramos@gmail.com

Nadia Nedjah
nadia@eng.uerj.br

Luiza de Macedo Mourelle
ldmm@eng.uerj.br

Brij B. Gupta
bbgupta.nitkkr@gmail.com

[1] Department of Electronics Engineering and Telecommunications, Faculty of Engineering, State University of Rio de Janeiro, Rio de Janeiro, Brazil

[2] Department of Systems Engineering and Computation, Faculty of Engineering, State University of Rio de Janeiro, Rio de Janeiro, Brazil

[3] Department of Computer Engineering, National Institute of Technology Kurukshetra, Kurukshetra, India

crowd behavior successfully, revealing that the proposed scheme outperforms the available state-of-the-art algorithms for global anomaly detection.

1 Introduction

Recently many research works have focused on the detection and classification of abnormalities at real-time speed in crowded scenes. Nowadays, the large number of public events with a high density mass of people as well as the growth number of surveillance cameras available, make the automatic detection and classification of scenes with specific events possible and necessary. An automated detection of these specific events is important because human condition, in general, does not allow for a good performance at this task manually due to its monotony. Abnormality detection in crowded videos is a challenging task for many reasons: high density of people in handled scenes, difficulty in capturing the behavior of the people in the crowd, the quality of surveillance videos available and, mainly, the computational cost required in this kind of process.

The challenge of this work consists of developing a method capable of learning any behavioral pattern present in a crowd of people, with low computational cost to enable real-time surveillance applications. The optical flow is indicated by the literature as being a method capable of correctly extracting the movement registered between frames. Therefore, the main challenge occurs in the stage of analysis of the motion layers as generated by the optical flow method. The bacterial colony algorithm aims at obtaining the information regarding the movement spread, the movement density and the movement centroid of the crowd. For this purpose, the population attributes, stock quality and centroid of the colony, respectively, provide these information at low computational cost. The Kohonen's neural network groups the behavioral patterns of the bacteria colonies that emerge from distinct events occurring in the video, also at low computational effort.

This work allows for the detection of abnormalities in videos of crowded scenes. It does so optimizing the optical flow of scenes using an artificial bacteria colony. Abnormalities detection is achieved by a Kohonen's neural network, trained using the data regarding the bacteria behavior during the optical flow optimization process. The contribution of this work is four-fold: *(i)* a new method for the analysis and detection of abnormal events, which differ from those considered normal and are predominant in scenes from the same scenario; *(ii)* The method developed has a low computational cost in the analysis and detection stages; *(iii)* Moreover, the algorithm has low sensibility to noise and to sudden changes in video lighting as captured by optical flow; *(iv)* finally, this work reveals the potential of using the artificial bacteria colony algorithm together with Kohonen's neural network in real-time surveillance applications.

This article presents a method to detect abnormal events in specific environments by the analysis of surveillance videos using the artificial bacteria colony algorithm to optimize the motion layers obtained from pairs of frames. In this work, abnormality is defined as any odd event within cyclical events that occur in the scene. The Kohonens neural network classifies behavioral patterns from the evolution of bacterial colonies to detecting eccentric patterns in the given environment, such as suspicious activities to be reported to the human operator. Thus, this work fits the special issues scope as a real-world application using interactive

computational visual analytics, which is one of the sub-topics of the special issue call for papers. This has now been emphasized in the paper.

This rest of this paper is organized into 6 sections. First, in the Section 2 we describe the crowd anomaly. After that, in the Section 3, we describe some existing works related with human activities recognition and crowd abnormalities detection. Then, in Section 4, we explain the workflow of the proposed method. Subsequently, in Section 5, we provide details about the methodology used to approach the abnormalities detection problem. Subsequently, in Section 6, we present, analyze and discuss the experimental results. Last but not least, in Section 7, we draw some conclusions about the work and point out some exciting directions for future works.

2 Crowd anomaly detection

In general, the detection of anomalies in crowd videos occurs in the temporal domain, with the identification of features outside the range of normality or specific events [17]. In the literature, some of crowd abnormalities include individuals walking on a determined place or individuals running when other people around them are walking. Anomalous events are contextual and depend on other items of the scene [13]. Emergency situations, such as panic in a crowd or a threat to the people on the scene, are considered anomalies. Identifying these events is important for video surveillance applications. Spectral approaches based on clustering [1] and social force model [14] are widely used. Methods based on optical flow perform spatio-temporal volume analysis, wherein the movement is used to detect anomalous events [20].

Starting from the premise that the crowd tends to behave in a stable and homogeneous way under normal conditions, an abnormality may then occur when there is a rapid and sudden variation in the motion conditions within the considered scene. Research works indicate that the optical flow is among the best techniques for extracting motion from motion scenes [9]. However, this technique is sensitive to the video lighting conditions. In addition, it requires significant computational cost. The optical flow provides the brightness variation intensity of pixels between two frames. Thereby, it is possible to generate a spatio-temporal volume representing the motion variations within the video. In order to optimize the search for the areas of the video, wherein relevant activities that should be analyzed, the application of swarm intelligence based methods is promising [10, 20]. In this work, the meta-heuristic inspired by real colony behavior of bacteria shows attributes that allow a precise optimization while entailing a low computational cost. Another aspect, which makes possible the application of bacteria colony optimization, is its ability to mirror the crowd behavior in the scene by the bacteria behavior. In other words, the way the artificial bacteria move has a direct relation with the way people move within a crowded scene. Therefore, understanding the motion pattern of a colony of bacteria is, somehow, equivalent to understanding approximately the motion pattern of crowd.

In this work, the behavior patterns are unique for each of the studied scenarios and the classifier should be capable to recognize the abnormalities in diverse videos of the same scenario. Furthermore, the set of input data for machine learning, in this case, do not have the corresponding desired output, because these represent unpredictable crowd movement. Thus, this is a case of unsupervised leaning. So, only after training is completed, which is usually done using data clustering techniques, it would be possible to label manually the movement patterns that emerge. In this work, to approach this specific problem, we

exploit a Kohonen's neural network for classification. This kind network model is chosen due to its intrinsic capability of self-organization via a competitive method. This method is capable of detecting similarities, regularities and correlations between the input patterns set by clustering them. The training set is extracted from the information about the artificial bacteria behavior during the optimization process. After training the network, the clusters are manually labeled to distinguish specific events.

3 Related works

The analysis of human behavior as captured on videos evolved gradually towards the detection and recognition of abnormal behavior and, eventually, for recognition and identification of specific events. The taxonomy for human behavior analysis, as described in [3], emphasizes the relevance of information related to movements in the video. This section presents a consolidated vision about critical aspects related to developing techniques to analyze and detect events within a given video.

A crowd abnormal behavior can be divided into two classes: either focusing on the individual [2] or on the crowd [7]. In the first case, the crowd is considered as a collection of individuals. Therefore, it is necessary to segment it and track trajectories therein. However, this approach is seriously affected by occlusions in the crowd scene. In the former case, the crowd is treated as a whole in the analysis of medium and high-density scenes. Instead to tracking individuals, this approach extracts features of the crowd to represent its state.

Approaches based on supervised learning [22] use data, which are hard to be obtained, for appropriate training of support vector machine (SVMs). For automated surveillance, supervised approaches are less attractive because they usually require a new training every time an unpredictable alteration of scene occurs. Approaches based in multiple learning are techniques that are more widely used to recognize actions without supervision [12]. Unsupervised approaches offer an advantage over supervised ones in terms of feasibility, due to their ability to adapt to the most diverse situations that usually arise in applications, such as automated surveillance.

The use of swarm intelligence produces improvements in optical flow analysis, optimizing detection and localization of abnormalities in videos with crowds. By performing a rapid convergence by swarm of agents to the areas of interest, this could lead to a sharp decrease in terms of the time required for analysis of motion between frames. In [18] the detection of abnormalities in crowd scenes is performed using particle swarm optimization, which optimizes the interaction forces computed by the social force model. A new approach to detecting abnormal events in crowded scenes is presented in [11]. The swarming theory is applied for the identification of pedestrian movement characteristics, through the histogram of the so-called Oriented Swarm Acceleration (HOSA), which can efficiently capture the dynamics of crowd scenes.

4 System architecture

In this section, we present the basic architecture of the proposed system. Figure 1 shows the main development stages of the proposed method to approach the problem of the global abnormality detections. First stage uses short video as input. Then, the optical flow is extracted for consecutive frames, generating spatio-temporal volumes of vector fields. Each
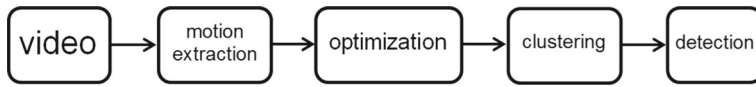
Fig. 1 Main steps of the proposed system

movement layer containing the magnitudes of optimal flow is treated as a cost function and optimized using the colony algorithm of artificial bacteria colony. The behavior records of the bacterial colonies are used as attributes to train a Kohonen's neural network. The neural network groups the data that presents similarities into classes and classifies the different events occurring in the scene.

In the proposed method all the frames of video are then converted from RGB color system to grayscale. Each layer of the optical flow, which consists of field of vectors with a magnitude variation according to the brightness patterns, is obtained considering each pair of consecutive frames of the video. The areas with movement between frames usually have the highest apparent velocities. Therefore, these areas represent the regions of interest. The ABC algorithm considers the field vector magnitudes of the apparent velocity as the cost function. It optimizes this proposed cost function by positioning artificial bacteria around all its local optima.

Colonies of virtual bacteria are scattered randomly in each layer of optical flow. In each layer, the bacteria are subjected to the Darwinian natural selection, in which only the more adapted survive and reproduce. The bacteria less adapted, that are those that during foraging could not obtain enough food as represented by the cost function, are inevitably eliminated. Considering each layer of optical flow, at the end of a defined number of epochs, the colony stabilizes and the bacteria are naturally distributed on existing local optima of fitness function. The amount of bacteria in the colony (population), the food stock of all the surviving bacteria (fitness accumulated) and the midpoint module of bacteria positions (centroid) are registered for each layer. These attributes represent the scatter of movement in the scene, the amount of movement between frames, and the position of the core of the crowd, respectively. In the scenes of video, where there are large regions with little movement, the bacteria population should be high. If there are small regions with large movement, the food stock should be bountiful. If the crowd of the video suddenly disperses, the centroid should reveal the magnitude and directions of crowd motion.

The optimization of the spatio-temporal volume of optical flow generates exit vectors containing the amount of bacteria in colonies, food stocks and centroid module. These attributes are used as the input variables of Kohonen's neural network. The network clusters, in classes, the colonies that have similar patterns of behavior. Once trained, the network performs the detection of a specific event every time that a colony of bacteria shows the same behavior pattern. In this work, the specific event is namely crowd abnormal behavior.

5 Proposed method

In this section, we give a detailed presentation of the proposed algorithm. The proposed scheme consists of three main components: *(i)* optical flow, *(ii)* artificial bacteria colony and *(iii)* Kohonen's neural networks.

5.1 Optical flow

Several methods use optical flow to perform event detection in frames sequence [19]. The optical flow estimates the motion between a pair of frames. The essence of the method is to find the estimate of the apparent motion between frames in relation to changes in the brightness patterns (pixel value of the grayscale image).

Optical flow is the distribution of the apparent velocities of brightness patterns in an image. Figure 2 shows the field of vectors with velocities associated to a sequence of images. This arises both from the movement of the scene objects and the movement of the camera. A sequence of images can be represented by its luminance function $I = I(x, y, t)$. The luminance conservation hypothesis means that the luminance of a physical point in the image sequence does not change over a short time interval, as defined in (1):

$$I(x, y, t) \approx I(x + \Delta x, y + \Delta y, t + \Delta t). \tag{1}$$

In order to determine the optical flow, it is assumed that at each pixel of the image, the constraint, as defined in (2) applies:

$$I_x u + I_y v + I_t = 0, \tag{2}$$

wherein I_x, I_y and I_t are the partial derivatives of the image, u and v are two components of the magnitude vector of the optical flow at a pixel position. This constraint alone is not enough to compute the vector field. Since the neighboring points of a moving object have similar velocities, it is possible to assume that the vector field of the optical flow is smooth. Thus, the vector field (u, v) is found by minimizing the function of (3):

$$\int \int [(I_x u + I_y v + I_t)^2 + \alpha^2 (v_x^2 + v_y^2 + u_x^2 + u_y^2)], \tag{3}$$

wherein the double integral extends to the entire image. The solution to this problem can be found, based on variational principles, by solving the following differential equations:

$$I_x^2 u + I_x I_y v = \alpha^2 \nabla^2 u - I_x I_t$$
$$I_x I_y u + I_y^2 v = \alpha^2 \nabla^2 v - I_y I_t, \tag{4}$$

wherein α represents a suitable weighting factor to compensate the error magnitude that is proportional to noise in the measurement.
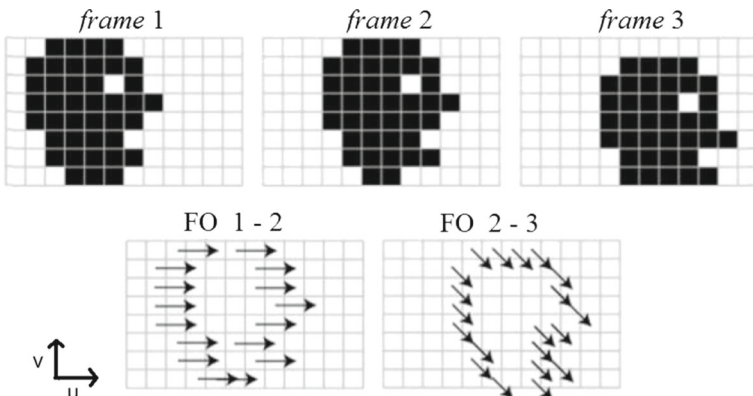

Fig. 2 Optical flow illustration

In this way, a solution can be found by an iterative procedure [8], using the following equations:

$$u^{n+1} = u^{-n} - \frac{I_x(I_x u^{-n} + I_y v^{-n} + I_t)}{\alpha + I_x^2 + I_y^2}$$

$$v^{n+1} = v^{-n} - \frac{I_y(I_x u^{-n} + I_y v^{-n} + I_t)}{\alpha + I_x^2 + I_y^2}.$$

(5)

Figure 3 shows an illustration of the procedure considering the optical flow between two adjacent frames. Figure 3a and b show two images, representing the two adjacent frames of a footstep movement as carried out by a man. Figure 3c shows the calculated optical flow between these two frames. Finally, Fig. 3d represents a detailed view of the optical flow, in which the direction field of the leg movement is shown.

In this work, we use the optical flow toolbox of MATLAB $^{®}$. The *FlowLK* function presents the improved optical flow by the local application of Tikhonov regularization, which aims at stabilizing the algorithm. The motion layer used in the optimization performed by the ABC algorithm consists of the vector field modulus of the apparent velocities of the optical flow, defined according to (6):

$$C_{OF}(x_{OF}, y_{OF}, k) = \sqrt{u(x_u, y_u, k)^2 + v(x_v, y_v, k)^2},$$

(6)

wherein $k = \{1, 2, ...,$ total of frames - 1 $\}$ and is associated with the spatio-temporal volume layer number of the optical flow. $OF(x_O F, y_O F, k)$ is the motion layer k obtained from the apparent velocity modulus $u(x_u, y_u, k)$ e $v(x_v, y_v, k)$, and x_u and y_u are the positions of the magnitudes of u, as well as x_v and y_v are the positions of the magnitudes of v.

In the present work, videos containing f frames generate spatio-temporal volumes with $k = f - 1$ layers. Each layer has a vector field of magnitudes of the optical flow. The

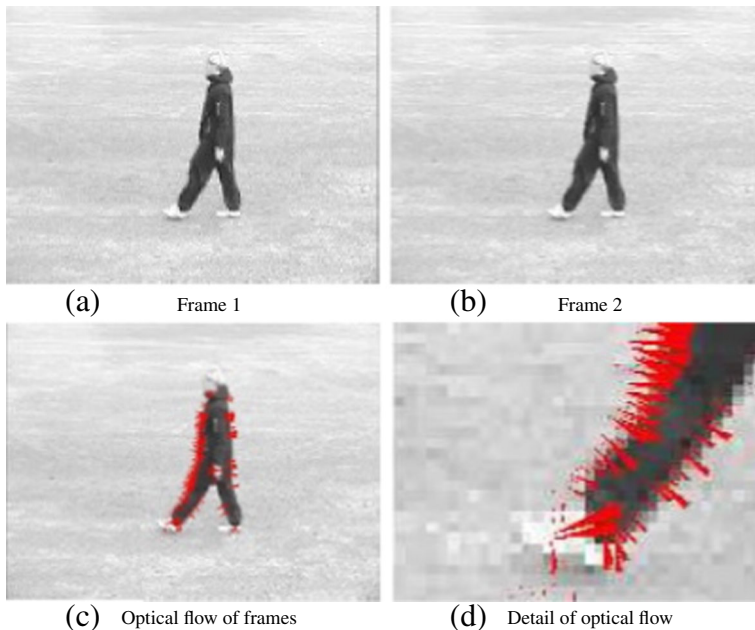|	
(a)　　Frame 1	(b)　　Frame 2
(c)　Optical flow of frames	(d)　Detail of optical flow

Fig. 3 Illustration about optical flow extraction of frames

regions of interest have the largest magnitudes. Each layer is then treated as a cost function. The layers of spatio-temporal volumes, or cost functions, are optimized by the algorithm of artificial bacteria colony.

5.2 Artificial bacteria colony algorithm

Bacterial foraging optimization (BFO) is a metaheuristic inspired by the behavior of *Escherichia coli* bacteria in the search for nutrients in their environment [16]. BFO is widely applied in numerical optimizations and uses the following basic steps: chemotaxis, reproduction and elimination-dispersion.

Chemotaxis is the chemically directed movement that some living beings develop. Chemotaxis, and the chemical substances involved in it, are used by some single-celled organisms, insects, mammals and even men. It is used for various purposes, such as during the search for nutrients, in order to avoid predators, generate communication between individuals in the formation of colonies or groups, for sexual attraction, or in the territorial demarcation [15]. In addition to this broad definition, the term chemotaxis, in the scientific literature, is almost always used to refer to cell movement in response to the concentration gradient of chemicals present in the environment.

Evolved over millions of years by nature, the chemotaxis of bacteria is a highly optimized process of searching and exploring unknown environments. Due to advances in the field of computation, chemotactic strategies of bacteria and their excellent search capability can be modeled, simulated and emulated to develop nature-inspired optimization methods, which are an alternative to existing methods. In this work an algorithm based on chemotactic strategies of bacteria is developed.

In the present work, a bacteria colony consists in all artificial bacteria of a specific optical flow layer. The artificial bacteria are points in a Cartesian Plane superimposed on a layer of optical flow. The Cartesian plane has the same dimensions of the optical flow layer. Each bacterium evaluates its position through a correspondence with the position of the optical flow layer, on which it is superimposed. The value obtained in this evaluation is considered its fitness. The fitness found by each bacterium is the magnitude of the optical flow.

In conjunction with chemotaxis, the Darwinian natural selection mechanism applies. Artificial bacteria use the fitness of the cost function as food to remain alive. At each epoch, in a single layer, the bacterium evaluates the fitness of its position and stores it, thus generating a food stock. For each layer, there are several values for the magnitudes of brightness variation. Therefore, bacteria that are exploring regions with low values, will have a much lower stock than those exploring regions with high values. In this way, the layer with its various regions becomes an environment that stimulates competition.

The survival condition (SC) is defined by the mean and standard deviation of the magnitudes of each layer. The reproduction condition (RC) is defined by the mean and standard deviation of the values of nutritional stock accumulated by the surviving bacteria. Thus, the thresholds are estimated dynamically according to the particularities of the layers. Therefore, the algorithm of artificial bacteria colony (ABC) adapts itself quickly to the changes of scenery. The mean and standard deviation of the layer are calculated according to (7):

$$\mu(k) = \frac{1}{n} \sum_{j=1}^{n} C(x_j, y_j, k)$$

$$\sigma(k) = \sqrt{\frac{1}{n} \sum_{j=1}^{n} \left[C(x_j, y_j, k) - \mu(k) \right]^2}, \tag{7}$$

wherein $\mu(k)$ and $\sigma(k)$ are respectively the mean and the standard deviation of the layer magnitudes k, n is the total number of considered magnitudes e $C(x_j, y_j, k)$ represents the layer of motion as generated by optical flow.

The fitness of bacterium b_i is measured by the association between its position in the Cartesian plane and the position of the magnitude of the movement in the layer according to (8):

$$\text{Position}(b_i) = P(x_{bi}, y_{bi})$$

$$\text{Fitness}(b_i) = C(x_{bi}, y_{bi}, k),$$

(8)

The stock, represented by the accumulation of the fitness values of the bacteria is defined as in (9):

$$stock(b_i) = \sum_{epoch=1}^{Ne} C(x_{bi}, y_{bi}, k),$$

(9)

wherein N_e is the number of evolutionary epochs of the algorithm in each layer. At each epoch the number of bacteria in the colony changes due to elimination and reproduction. The stock at the last epoch consists, therefore, of the values accumulated only by the surviving bacteria in evolution process. The mean and standard deviation of the nutrients accumulated by the surviving bacteria is calculated according to (10):

$$\mu_{stock}(k) = \frac{1}{N_f} \sum_{i=1}^{N_f} stock(b_i)$$

$$\sigma_{stock}(k) = \sqrt{\frac{1}{N_f} \sum_{i=1}^{N_f} [stock(b_i) - \mu_{stock}]^2},$$

(10)

wherein the $\mu_{stock}(k)$ is the average stock of the surviving bacteria, N_f is the number of bacteria in current epoch, $stock(b_i)$ is the surviving bacterial stock b_i and $\sigma_{stock}(k)$ is the standard deviation of the nutritional stock of the layer. Thus the SC and RC are estimated according to (11):

$$SC(k) = \mu_{magnitude}(k) + s * \sigma_{magnitude}(k)$$

$$RC(k) = \mu_{stock}(k) + r * \sigma_{stock}(k),$$

(11)

wherein s and r are the parameters of the ABC algorithm that define the sensitivity of the natural selection process. Note that variations in the parameter s influence the spread of bacteria in the regions of interest and variations in r influence the concentration of bacteria in these regions.

The centroid is the point whose coordinates are obtained via the means of those of the points that form the considered geometric figure. In the case of this application, the geometric form is the one resulting from the distribution of the bacteria. The centroid of the colony and the centroid module are computed from the position of the surviving bacteria in the last epoch, according to (12):

$$\text{centroid}(k) = \left(\frac{1}{N_b} \sum_{i=1}^{N_b} x_b(i), \frac{1}{N_b} \sum_{i=1}^{N_b} y_b(i) \right)$$

$$\text{centroid module}(k) = \sqrt{\left(\frac{1}{N_b} \sum_{i=1}^{N_b} x_b(i) \right)^2 + \left(\frac{1}{N_b} \sum_{i=1}^{N_b} y_b(i) \right)^2},$$

(12)

wherein N_b is the total number of surviving bacteria in the last epoch while $x_b(i)$ and $y_b(i)$ are the coordinates of the bacteria regarding the horizontal and vertical axis, respectively.

During the elimination step, bacteria that have a fitness value lower than that specified un the survival condition (SC) of the layer do not survive. During this step, mainly in the first epoch, the elimination occurs for a big part of the bacteria located in regions of the layer with little or no movement. Thereafter, the reproductive process gains emphasis and there is a population explosion in regions with intense movement regarding the optical flow layer. This population explosion, whose growth is exponential, allows to reasonably delimit the outline of the region with high magnitudes, since the bacteria that distance themselves from this region are quickly eliminated.

During the reproduction step, bacteria that survive and have a stock value above the reproductive condition (RC) of the epoch reproduce by cloning itself into two identical bacteria, with the same stock that are positioned randomly in the vicinity of the point where the reproduction occurred. Thus, at each epoch new bacteria are randomly positioned near of regions with magnitude values well above average. Bacteria that are very distant from these regions are eliminated in the early stages. Therefore, the outline of the colony of bacteria is shaped by the contour of the regions of interest.

Algorithm 1 presents the steps of the proposed CBA. First, we define the size of the population and the randomly initial position of the bacteria in the first epoch of the layers. The number of epochs N_e for each layer is then defined, considering that at each epoch, starting from the second one, the number of bacteria practically doubles. This is a parameter that has a strong influence on the computational effort required during the execution of the algorithm. It was noted that a number of times less than 3 creates instabilities in the final population of the colonies while a value greater than 6 does not generate significant differences in results. Therefore, the value used in this work is set as 5 during all the experiments. The parameter s of the survival condition is defined to adjust the sensitivity of the elimination of the worst bacteria. Finally, the parameter r of the reproduction condition, which may be thought of as an elitization parameter, is defined as to allow only the best bacteria to reproduce. The reproduction condition is responsible for the high concentration of bacteria at the highest points of the cost function, equivalent to the regions of greatest movement between frames. After parameter settings, there is a loop to optimize the layers, another loop for the evolutionary epochs regarding each layer and 2 other loops to execute the eliminations e reproduction for each bacterium at a given epoch. Finally, the algorithm provides the population records, nutritional stock and the centroid module of the k bacterial colonies for the k layers of movement.

This work proposes the use of a modified version of the BFO algorithm to locate areas of the frames that have optical flow magnitudes above a dynamic threshold. The dynamic threshold adapts to the content of each layer. This dynamic threshold is associated with the mean and standard deviation of magnitudes of brightness variation of each optical flow layer. The processes of chemotaxis, reproduction and death of bacteria allow the used algorithm to be able to discover information about the behavior of the crowd in the sequence of frames via the analysis of the population behavior of the bacteria.

Figure 4 allows an observation of the analogy between the foraging of a real *vs.* virtual bacteria, because both seek the points of higher nutrient density, as the optimal ones of a multimodal function. Figure 4d shows an image of the implementation carried out in this work, in which the bacteria are positioned in the regions that agree with the presence of

intense movement. In these areas where many people congregate, the number of bacteria increases.

Algorithm 1 Artificial Bacteria Colony – ABC

Require: volume spatio-temporal with k layers;

  1:    Define the inicial number of bacteria N in each layer;
  2:    Define the position of the bacteria in the first epoch;
  3:    Define the number of epochs N_e for each layer;
  4:    Define the survive parameter s;
  5:    Define the reprodution parameter r;
  6:    **for** each layer k **do**
  7:        calculate $\mu_{magnitude}(k)$ e $\sigma_{magnitude}(k)$;
  8:        calculate $CS(k)$;
  9:        **for** each epoch e **do**
10:            **for** each bacterium i **do**
11:                calculate fitness(i) of bacterium(i);
12:                **if** fitness of bacterium(i) $\geq$ threshold(k) **then**
13:                    stock(i) := stock(i) + fitness(i);
14:                **end if**
15:                **if** fitness of bacterium(i) $\leq SC(k)$ **then**
16:                    eliminate bacterium(i);
17:                **end if**
18:            **end for**
19:            calculate $\mu_{stock}(k)$ e $\sigma_{stock}(k)$;
20:            calculate $RC(k)$;
21:            **for** each bacterium i **do**
22:                **if** stock of bacterium(i) $\geq RC(k)$ **then**
23:                    duplicate bacterium(i) ;
24:                **end if**
25:            **end for**
26:        **end for**
27:        calculate centroid module of colony k;
28:    **end for**
29:    **return** population of all layers;
30:    **return** stock of all layers;
31:    **return** centroid module of all layers;

The evolution of the bacteria occurs along some epochs in a layer of movement, and the algorithm converges for the formation of the colony around the areas with the highest magnitudes of the layer. Figure 5 shows the evolution of the colony of bacteria over 5 epochs.

Figure 6 shows, in perspective, the distribution of a colony of bacteria as distributed across an optical flow layer, after few epochs. It is possible to see that the noise does not represent a good nutritional option for the bacteria, due to its relatively low value. Bacteria always opt for more dense regions, where parts of the colony can establish themselves with stability.
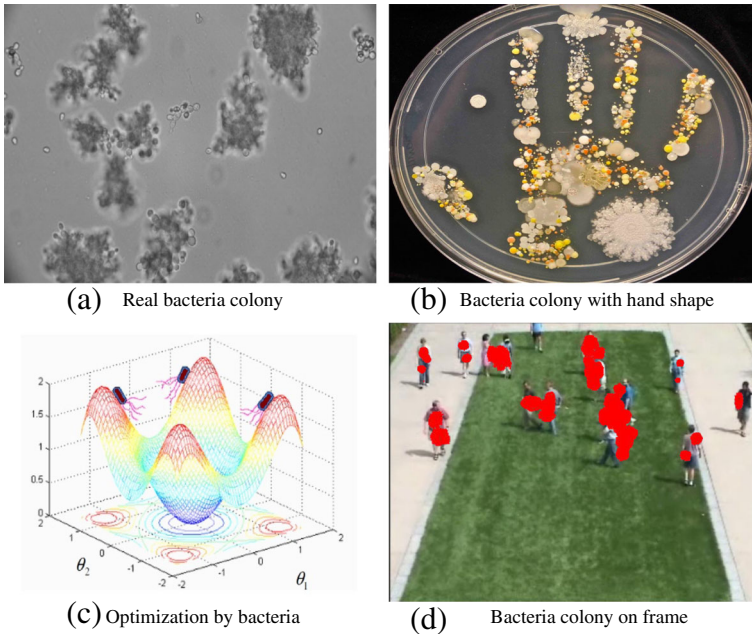
(a) Real bacteria colony   (b) Bacteria colony with hand shape

(c) Optimization by bacteria   (d) Bacteria colony on frame

Fig. 4 In **a** and **b** real bacteria colony are shown. In **b** specifically, bacteria demarcate the nutritional environment in the form of a hand. In **c** an cost function is shown where bacteria seek out the local optima by chemotaxis. In **d** it is shown, on a frame, the distribution of the colony of artificial bacteria to a layer of optical flow corresponding to that frame and its consecutive one

5.3 Kohonen's neural network

Kohonen's neural network or self-organizing Kohonen map (SOM) is part of a group of competitive neural networks that use competition strategies to adjust their weights during the learning process. SOMs can preserve the neighboring relations of input data and use unsupervised training to find similarities based only on these input patterns. The main purpose of Kohonen's neural networks is to group incoming data that are similar, thus forming classes or clusters. Figure 7 shows a Kohonen's map, where colors highlight a cluster formed after training.

Kohonen's neural networks are nonlinear projections of high dimensional spaces for a low-dimensional map M. The two-dimensional mapping, often adopted in the Kohonen's model, is a function $f : X \rightarrow M$, wherein $f : X \rightarrow \mathbb{R}^n$ and $f : M \rightarrow \mathbb{R}^2$, which assigns each element $x \in X$ a pair $(i, j) \in M$. The elements $m_{i,j}$ of the map M, as well as the input data are n-dimensional vectors that adjust the values of the synaptic weights of the neural network. Kohonen's neural networks owns features that make it appropriate for modeling various crowd behavior patterns. First, it performs topological ordering while computing the feature map, which makes the spatial location of neurons in the lattice correspond to a particular domain of input patterns and the neighborhood of a neuron share similar characteristics which can tolerate a certain degree of variance between different patterns through smooth transition [6].

Due to their characteristics, Kohonen's neural networks can be well applied in several areas, such as voice recognition, video behavior analysis and combinatorial optimization.
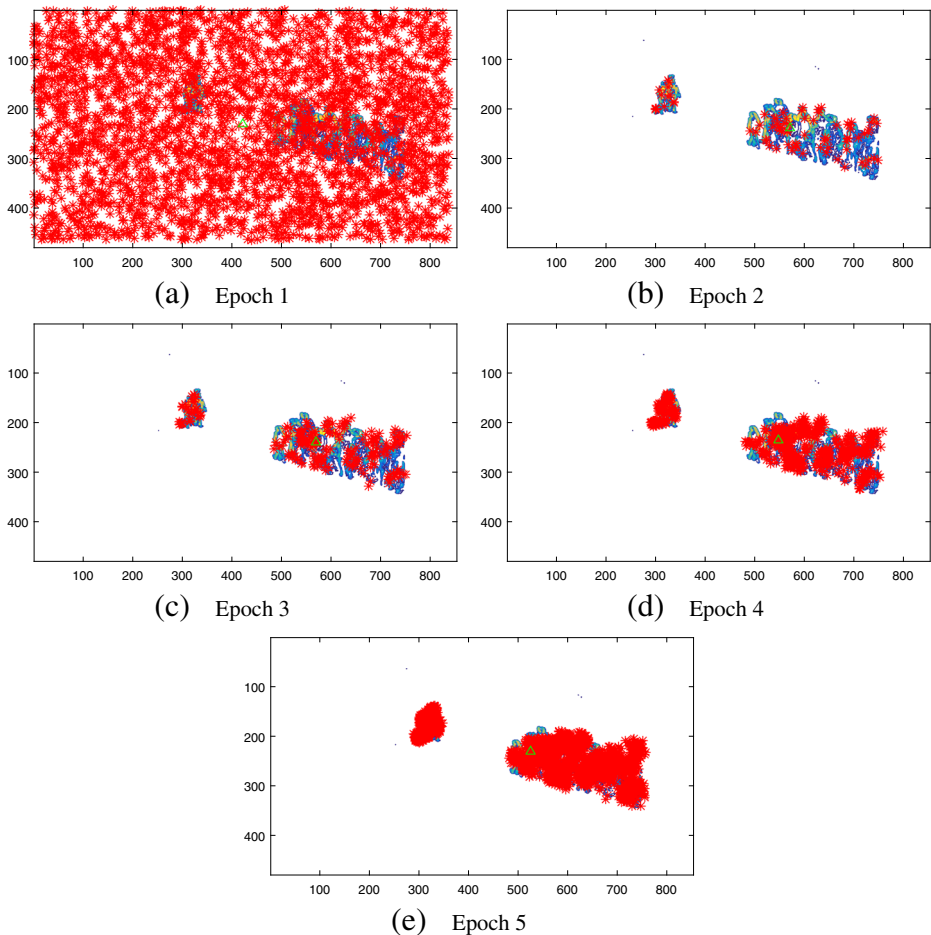
Fig. 5 Evolution of the one colony of bacteria over 5 epochs of the motion layer. The small green triangle in (**e**) represents the position of the centroid

The fact that similar mappings can be found in several areas of the human brain and other animals, indicates that the topology design is an important principle in signal processing systems, in the field of artificial intelligence.

In this work, after optimizing the spatio-temporal volume of the optical flow, the amount of bacteria in each colony, the food stock of all bacteria of the last epoch and the centroid module of the colonies are recorded for each layer. These values form a training set. These patterns need to be clustered according to their similarities. For this, a Kohonen's neural network is used.

The training set has three vectors: population of the colonies, food stock and centroid module. The Kohonen's neural network based algorithm, used in this work, is implemented using the SOM tool of MATLAB. The number of neurons for each video is set according to Table 1. The weight updating procedures occurs based on the batch training algorithm. The distance between the neurons is calculated using the Euclidean distance. The number of neurons is chosen to get the best rating of the different events, which occur in the video,
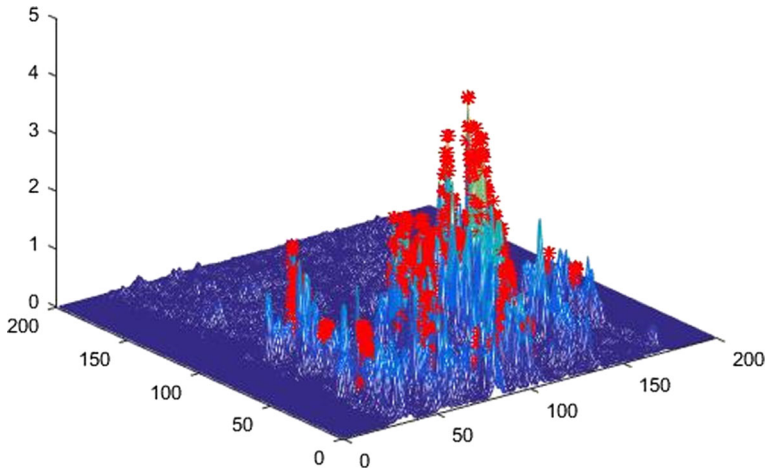
Fig. 6 The cost function is the vector field containing the magnitudes of brightness variation of an optical flow layer. Artificial bacteria are located in areas of highest magnitude

wherein each neuron represents a distinct class. During training, the network determines the neuron that best responds to the input vector. The vector of weights for the winning neuron is adjusted according to the training algorithm. During the process of network self-organization, the neuron whose vector of weights is closest to the vector of the input patterns is selected as winner. The weight of the winning neuron is adjusted. Neighboring neurons are also adjusted, but their weights are updated in the inverse proportion of the distance of the winning neuron.
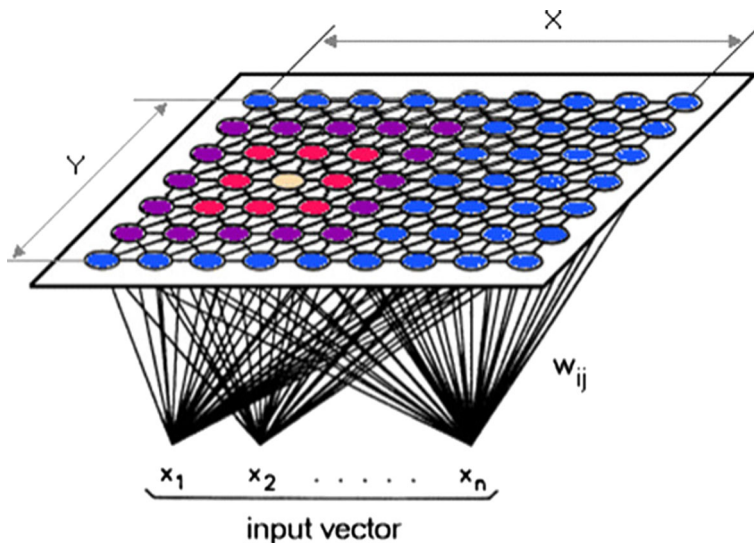

Fig. 7 Kohonen's neural network

Table 1 Number of neurons, average training time and standard deviation for each video of the UMN dataset

Video	#Neurons	Training average time (s)	Standard deviation
1	3	2.7562	0.4120
2	4	2.8436	0.4678
3	5	2.3217	0.4678
4	4	3.5428	0.3628
5	3	2.1842	0.4578
6	3	2.3567	0.4231
7	3	2.1547	0.1937
8	3	2.7853	0.2825
9	4	3.0621	0.2445
10	3	2.2754	0.3584
11	4	3.2621	0.2665

In this work, Kohonen's neural network cluster into classes the input data that show similarities. This allows that population, stock, and centroid values define a specific event in the scene.

6 Performance results

In order to validate and evaluate the performance of the proposed method, extensive experiments were performed using public dataset UMN [5]. The UMN dataset has 11 videos, with 3 different scenarios. There are crowds of low and average densities, moving indoors and outdoors. All video sequences exhibit panic situations among individuals. The video begins with individuals exhibiting normal behavior followed by sudden abnormal activity. Figure 8 shows frames from each video of the dataset.

All UMN dataset videos have resolution of $480x854$ pixels. All algorithms were run with Matlab R2015a. All simulations were performed on PCs with Intel Core i7 950 3GHz, 8Gb RAM and Microsoft Windows 7 Home Premium operating system.

Figure 9 shows the step sequence followed during the experiments for each video of the dataset. Figure 9b depicts the movement layer extracted by the optical flow (foreground) of the scene in Fig. 9a, as a colormap. In Fig. 9d, it is possible to notice that bacteria do not cover the areas with less motion than the others.

During the first epoch of optimization processes, the bacteria are randomly scattered across each layer in the spatio-time volume of the optical flow. For all experiments, 5 epochs are used for each layer. For experiments with more than 5 epochs, the algorithm requires excessive processing time, because the bacterial colony grows exponentially. The parameter s is set to 1.5 and the parameter r is set to 2. Regarding Kohonen's network, 1000 epochs of training are used for all videos. The number of neurons defined for each video in the dataset is shown in Table 1. The second column shows the average training time for each network while the third one shows the standard deviation.

During the first stage of the system, the preprocessing of the videos is performed to obtain the layers of movement by the optical flow. This stage is important because it is associated with the quality of the generated layers and usually requires a great deal of computational effort due to the complexity of the entailed mathematical computation required therein.
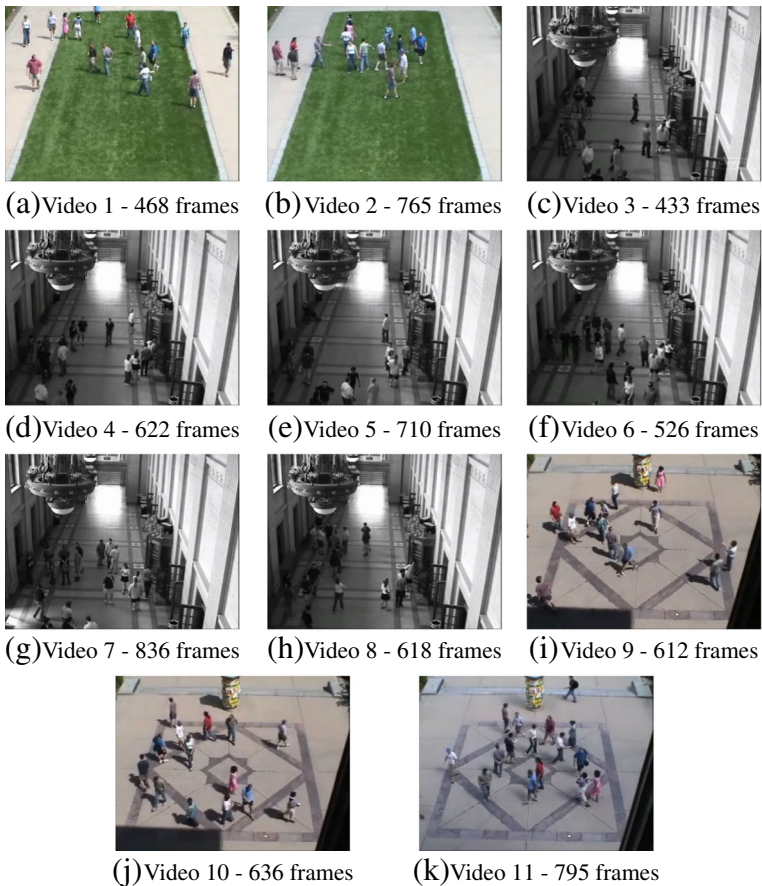
(a) Video 1 - 468 frames   (b) Video 2 - 765 frames   (c) Video 3 - 433 frames

(d) Video 4 - 622 frames   (e) Video 5 - 710 frames   (f) Video 6 - 526 frames

(g) Video 7 - 836 frames   (h) Video 8 - 618 frames   (i) Video 9 - 612 frames

(j) Video 10 - 636 frames   (k) Video 11 - 795 frames

Fig. 8 Sample frames from the videos of dataset UMN

Figure 10 shows the results obtained regarding the processing time average for all frame pair in each video of the dataset.

Experiments were performed with the initial populations of 250, 500, 1000, and 2000 bacteria. Figure 11 presents the optimization time average values for each layer as obtained for a pair of consecutive frames by the optical flow. The results shows that for the UMN dataset, 250 and 500 bacteria allow convergence of the optimization process at intervals that are on average less than 0.033 seconds, which the main requirement for real-time application where videos run are recorded at 30 frames per second.

Figure 12, which refers specifically to the use of 500 bacteria in the initial population, reveals the history of the bacterial colonies generated along video #1. The variation in the final amount of bacteria of each epoch is associated with the scattering and intensity of the movement between frames. The depicted graph illustrates the way bacteria colonies mirror the behavior of the crowd. In this video people walk into the environment at first then suddenly start running in all directions. The turmoil is characterized in the graph as being the moment when the population, stock and centroid values suddenly rise above average.
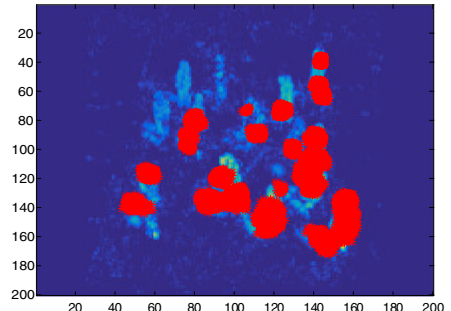
(a) Crowd walking
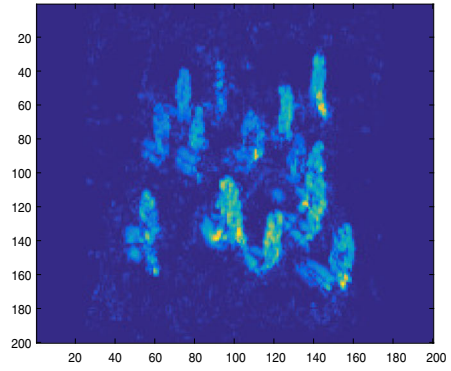
(b) Foreground of frame
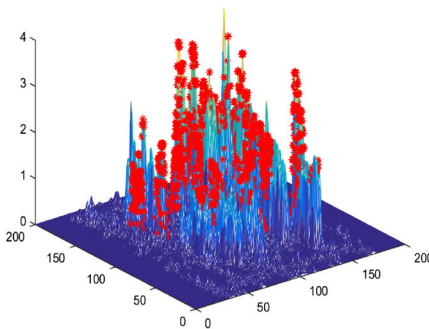
(c) Bacteria on best areas of fitness function

(d) Bacteria colony on foreground
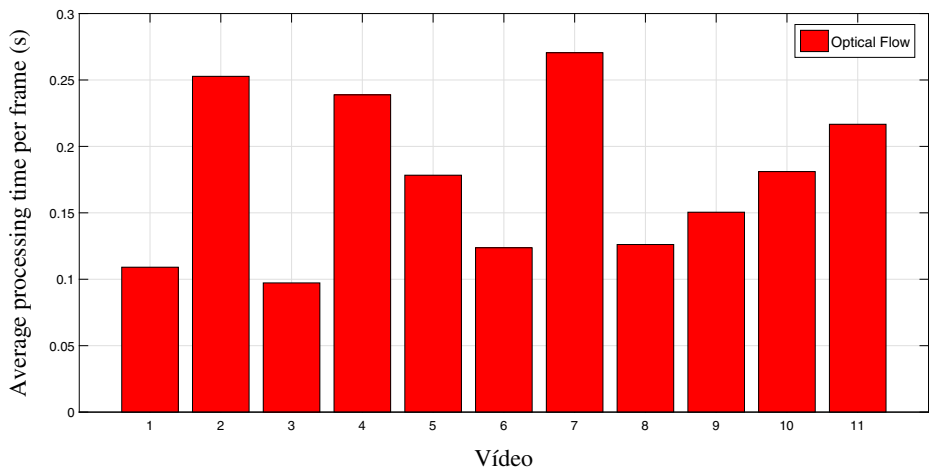
Fig. 9 Optimization steps of proposed system for UMN

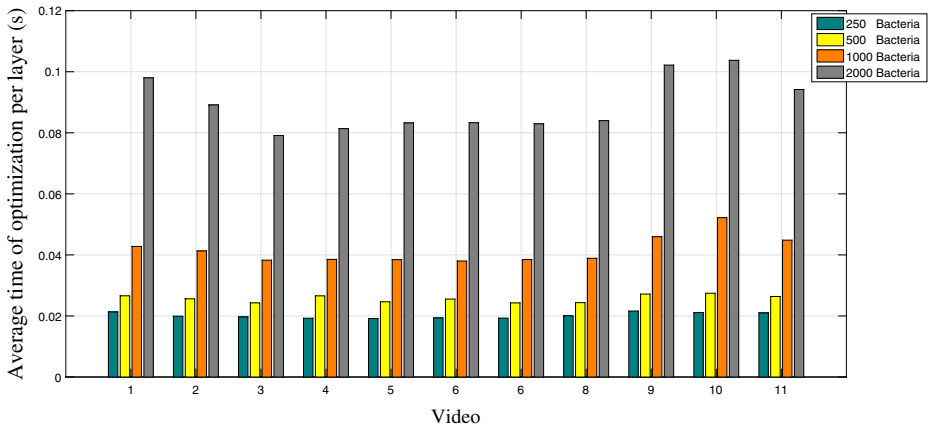**Fig. 10** Average processing time for generating a 1 layer of moviment with optical flow per frame pair

Fig. 11 Average optimization time per layer for 250,500,1000 and 2000 bacteria in the initial population of the first epoch of each layer
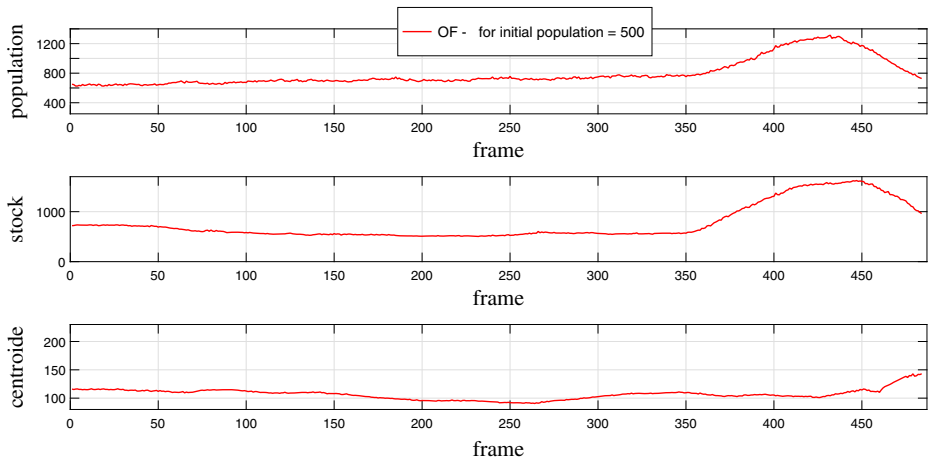

Fig. 12 Record the populations, stocks, and centroids of bacterial colonies obtained along optimization of the 1 video for layers of motion generated by optical flow

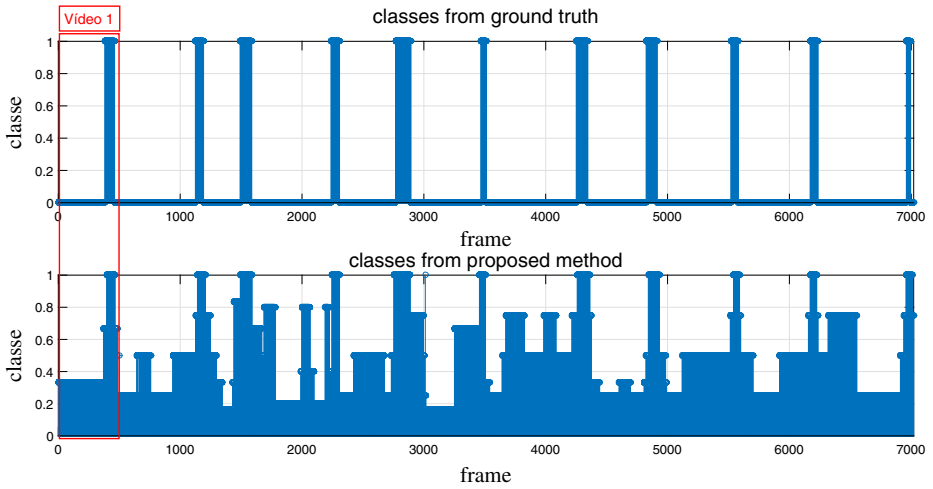**Fig. 13** Detection by Kohonen network with data from bacteria colony behavior

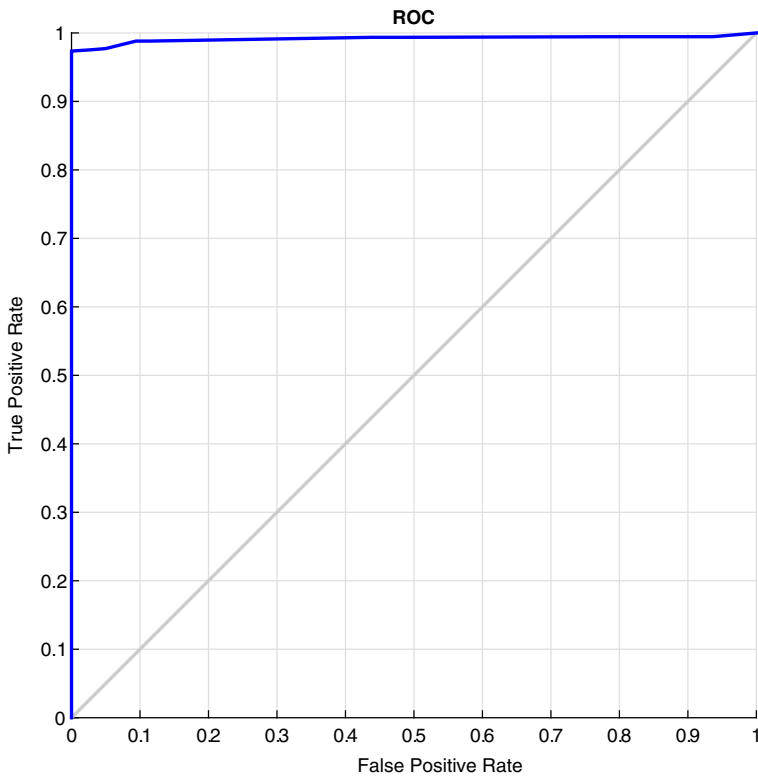Fig. 14 Classes detection for UMN dataset. The abnormalities is associeted with the class 1

Fig. 15 The ROC curve for detection os abnormal frames in UMN dataset

Table 2 Area under ROC curve of the proposed method *vs.* other methods for the UMN dataset

Method	Area under ROC	Improvement
Optical Flow [14]	0.8400	–
Social Force Model [14]	0.9600	14.28%
Chaotic Invariant [21]	0.9900	17.85%
Sparse Reconstruction [4]	0.9780	16.42%
Artificial Bacteria Colony	0.9929	18.02%

Figure 13 shows the comparison between the detection of the ground truth for video 4 and the detection as yielded by the proposed method for the UMN dataset.

Figure 14 shows the classification performed by the proposed method, in which class 1 represents the abnormal behavior. The other classes, whose number is smaller than 1 are considered normal events. The graph with the classes, based on the ground truth of all the UMN dataset videos illustrates the comparison with the detections as yielded by the proposed method. It is noteworthy to point out that all the abnormalities of the videos are accurately captured by the proposed method.

In order to evaluate the quality of the classification performed by the Kohonen's neural network, Fig. 15 shows the ROC curve of the classification performed by the proposed method for the detection of abnormalities. The curve is obtained by comparing the classes detected by the application of the proposed method and those of the ground truth of the videos. This curve was generated by the command *plotroc* available in Matlab.

Table 2 reports the area under the ROC curve for the proposed method using artificial bacteria colony as opposed to those obtained by other methods for detection of abnormality behaviors regarding the UMN dataset. It is noteworthy to point out that the proposed method offers the highest value of area under the ROC curve regarding state-of-the-art existing related works, with an 18% improvement. The improvement is computed using the work [14] as reference.

7 Conclusion

This work introduces a new method to perform the detection of global abnormalities in videos of crowded scenes using optical flow, artificial bacteria colony and Kohonen's neural network. This work demonstrates the ability of the method to capture the dynamics of the crowd by optimizing the optical flow of the videos. One of the main advantages of the method is that it does not need to carry out individual tracking to detect abnormalities. Another advantage of the proposed methodology is that it requires a reduced computational effort for the analysis and detection phases, allowing its application within real-time software.

The experiments revealed that the colony of artificial bacteria can comprehend the behavior of the individuals in the videos with crowds and was robust enough to ignore the noise captured by the optical flow. The detection of abnormalities is performed by the Kohonen's neural network and the results indicate that the network is able to adapt to the most diverse scenarios quickly, accurately and with low computational cost.

Improvements of this work could be obtained by using the space-temporal volume containing magnitudes of the interaction forces, using the Social Force Model [14] instead of pure optical flow. Improvements could also be implemented in the bacterial colony algorithm to optimize spatio-temporal volumes faster and with less computational cost. These are some directions for near future works that we intend to take over to improve the achieved results.

References

1. Andrade EL, Blunsden S, Fisher RB (2006) Modelling crowd scenes for event detection. In: 18th International Conference on Pattern Recognition (ICPR'06), vol 1. IEEE, pp 175–178
2. Brostow GJ, Cipolla R (2006) Unsupervised bayesian detection of independent motion in crowds. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol 1, pp 594–601
3. Chaaraoui AA, Climent-Pérez P, Flórez-Revuelta F (2012) A review on vision techniques applied to human behaviour analysis for ambient-assisted living. Expert Syst Appl 39(12):10873–10888
4. Cong Y, Yuan J, Liu J (2011) Sparse reconstruction cost for abnormal event detection. In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp 3449–3456
5. Dataset UMN (2006) Unusual crowd activity dataset of university of minnesota, available in http://mha.cs.umn.edu/movies/crowdactivity-all.avi. Minessota
6. Feng J, Zhang C, Hao P (2010) Online learning with self-organizing maps for anomaly detection in crowd scenes. In: 2010 20th International Conference on Pattern Recognition (ICPR). IEEE, pp 3599–3602
7. Garate C, Bilinsky P, Bremond F (2009) Crowd event recognition using hog tracker. In: 2009 Twelfth IEEE international workshop on performance evaluation of tracking and surveillance (PETS-Winter), pp 1–6
8. Horn BKP, Schunck BG (1981) Determining optical flow. Artificial intelligence 17(1–3):185–203
9. Kajo I, Malik AS, Kamel N (2015) Motion estimation of crowd flow using optical flow techniques: a review. In: 2015 9th International Conference on Signal Processing and Communication Systems (ICSPCS). IEEE, pp 1–9
10. Kaltsa Vagia, Briassouli Alexia, Kompatsiaris Ioannis, Hadjileontiadis LeontiosJ, Strintzis Michael-Gerasimos (2015) Swarm intelligence for detecting interesting events in crowded environments. IEEE Trans Image Process 24(7):2153–2166
11. Kaltsa V, Briassouli A, Kompatsiaris I, Strintzis MG (2014) Swarm-based motion features for anomaly detection in crowds. In: 2014 IEEE international conference on image processing (ICIP). IEEE, Paris, pp 2353–2357
12. Lui YM (2012) Tangent bundles on special manifolds for action recognition. IEEE Transactions on Circuits and Systems for Video Technology 22(6):930–942
13. Mahadevan V, Li W, Bhalodia V, Vasconcelos N (2010) Anomaly detection in crowded scenes. In: CVPR, vol 249, p 250
14. Mehran R, Oyama A, Shah M (2009) Abnormal crowd behavior detection using social force model. In: IEEE conference on computer vision and pattern recognition, 2009. CVPR 2009. IEEE, p 2009
15. Murray JD (2002) Mathematical biology i: an introduction, vol 17 of interdisciplinary applied mathematics
16. Passino KM (2002) Biomimicry of bacterial foraging for distributed optimization and control. IEEE Control Syst 22(3):52–67
17. Popoola OP, Wang K (2012) Video based abnormal human behavior recognition a review. IEEE Trans Syst Man Cybern Part C Appl Rev 42(6):865–878
18. Raghavendra R, Del Bue A, Cristani M, Murino V (2011) Optimizing interaction force for global anomaly detection in crowded scenes. In: 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), IEEE, pp 136–143
19. Rao AS, Gubbi J, Marusic S, Stanley P, Palaniswami M (2013) Crowd density estimation based on optical flow and hierarchical clustering. In: 2013 international conference on advances in computing, communications and informatics (ICACCI). IEEE, p 2013
20. Thida M, Eng H-L, Remagnino P (2013) Laplacian eigenmap with temporal constraints for local abnormality detection in crowded scenes. IEEE Trans Cybern 43(6):2147–2156

21. Shandong W, Moore BE, Shah M (2010) Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes. In: 2010 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, p 2010
22. Xinxiao W, Dong X, Duan L, Luo J, Jia Y (2013) Action recognition using multilevel features and latent structural svm. IEEE Transactions on Circuits and Systems for Video Technology 23(8):1422–1431

Joelmir Ramos holds a Master degree in Electronics Engineering from State University of Rio de Janeiro. Also, he holds an engineering graduation degree in electronics from the same university. His research interests include visual data mining, hardware architecture, FPGA in particular, and intelligent systems in general. Contact him at joelmiramos@gmal.com.
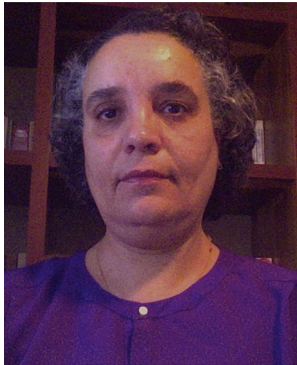
Nadia Nedjah is an associate professor in the Department of Electronics Engineering and Telecommunications at the Faculty of Engineering, State University of Rio de Janeiro, Brazil. Her research interests include functional programming, embedded systems and reconfigurable hardware design as well as cryptography. Nedjah received her Ph.D. in Computation from the University of Manchester - Institute of Science and Technology (UMIST), England, her M.Sc. in System Engineering and Computation from the University of Annaba, Algeria and her Engineering degree in Computer Science also from the University of Annaba, Algeria. Contact her at nadia@eng.uerj.br.

Luiza de Macedo Mourelle is an associate professor in the Department of System Engineering and Computation at the Faculty of Engineering, State University of Rio de Janeiro, Brazil. Her research interests include computer architecture, embedded systems design, hardware/software co-design and reconfigurable hardware. Mourelle received her PhD in Computation from the University of Manchester - Institute of Science and Technology (UMIST), England, her MSc in System Engineering and Computation from the Federal University of Rio de Janeiro (UFRJ), Brazil and her Engineering degree in Electronics also from UFRJ, Brazil. Contact her at ldmm@eng.uerj.br.

Brij B. Gupta received PhD degree from Indian Institute of Technology Roorkee, India in the area of Information and Cyber Security. In 2009, he was selected for Canadian Commonwealth Scholarship and awarded by Government of Canada Award ($10,000). He spent more than six months in University of Saskatchewan (UofS), Canada to complete a portion of his research work. He has published more than 80 research papers (including 01 book and 08 chapters) in International Journals and Conferences of high repute including IEEE, Elsevier, ACM, Springer, Wiley Inderscience, etc. He has visited several countries, i.e. Canada, Japan, Malaysia, Hong-Kong, etc. to present his research work. At present, Dr. Gupta is working as Assistant Professor in the Department of Computer Engineering, National Institute of Technology Kurukshetra, India. His research interest includes Information security, Cyber Security, Mobile Security, Cloud Computing, Web security, Intrusion detection, Computer networks and Phishing.