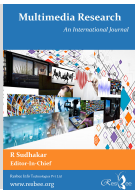
Vol.3 No.3 July 2020

Multimedia Research

Received 11 May, Revised 25 June, Accepted 3 July



Crowd Behaviour Recognition using Enhanced Butterfly Optimization Algorithm based Recurrent Neural Network

Yuying Chen
Western University of Ontario, Ontario, Canada
yuyingchen127@gmail.com

Abstract: The crowd emotion recognition is a motivating research area that helps the security personals by means of the public emotions to interpret the crowd activity in a region. Approximately several conventional techniques exploit the low-level visual features to comprehend the behaviors of a crowd which widen the gap between the high as well as the low-level features. The objective model is used to expand the automatic algorithm for emotion recognition; hence this work uses the Recurrent Neural Network (RNN). The Bhattacharya distance is used for effectual emotion recognition, which is necessary to choose video keyframes. The keyframes are subjected to the Space-Time Interest Points (STI) descriptor which extracts features that structure input vector to the classifier. RNN is trained by exploiting the enhanced Butterfly Optimization Algorithm (Enhanced-BOA). The developed classifier identifies the crowd emotions, like Escape, Angry, Happy, Fight, Running/Walking, Normal, as well as Violence. The experimentation of the developed technique revealed that developed technique obtained a maximum accuracy, sensitivity as well as specificity, correspondingly.

Keywords: Crowd Behaviour; Video Surveillance; Emotion Recognition; Optimization; RNN

Nomenclature

Abbreviations	Descriptions
STI	Space-Time Interest Points
SVR	Support Vector Regression
ROI	Regions of Interest
WT	Wavelet Transform
LSTM	Long Short-Term Memory

1. Introduction

For various applications like video surveillance CROWD behavior analysis encompasses enticed ever raising research attentiveness. Examining the visual crowd behaviors is demanding in modern trends and advances, mainly because of its variations, for instance: subtle changes, complex interactions, several semantics, etc [1].

The most prominent direction in this considered understands the crowd motion patterns in outdoor/indoor spaces. This task is represented preferably complex because of its motion flow variability of diverse individuals, chiefly in thick settings, whereas intricacy expands [2]. Precisely, motion directions of an individual are accurate possessing and frequently contingent on her/ his spatial context in a presented dense crowd setting [5]. This recommends that comprehension behavior for example: walking direction of a person have to be designed by evaluating her/ his context [17].

The crowd behaviors analysis covers diverse sub-issues like crowd alters or anomaly recognition [16], and crowd event characterization or recognition for that the aim is to naturally recognize adjustments or to conversely identify crowd events in video sequences [6]. Generally, the classification of conventional techniques for crowd behavior analysis and is categorized into two classifications such as object-based techniques as well as holistic techniques.

Numerous researchers investigate the crowds' structures via supervised learning models, like the graphical model, Bayesian model, as well as mixture agent's model. By integrating a data-driven learning technique, superior analysis is stated for crowd tracking on the basis of the motion patterns [1]. Nevertheless, crowd patterns stated in these researchers exploit merely minimum or middle-level features that consequences in the intricacy of digesting high-level crowd behaviors from a philosophically

https://doi.org/10.46253/j.mr.v3i3.a3

holistic insightful. Further, to encode more semantic information, high-level representations like group model, energy possibles as well as streakline potentials are developed, for crowd analysis to recognize more perceptive structures. In comparing with the medium-level features, and low-level motion features, the objective at designing object interactions when high-level semantic features by means of the affluent preceding information present a robust means towards superior stating as well as increasing a broad comprehension of crowd behaviors. High-level representations psychology, combine sociology as well as physics, to raise illustrative diversity and power [3].

The most important objective of the work is to present the application of the Space-Time Interest descriptor to the keyframes in order to attain the interesting points to facilitate precise emotional identification exploiting the crowded videos. By exploiting the RNN model the crowd emotion recognition is performed which is tuned optimally employing the enhanced-BOA that is the enhancement of the BOA to enhance the convergence of the proposed method.

2. Literature Review

In 2016, Yanhao Zhang, et al [1], proposed a novel crowd portrayal called Crowd Mood. This was established based on the identification which the social-emotional hypothesis of crowd behaviors perchance showed in order to examine the spacing interactions as well as motion patterns structural levels in crowds. Finally, structured trajectories were learned initially for the crowds using the particle advection exploiting the minimum rank estimation with group sparsity restraint that essentially distinguishes coherent motion patterns. Subsequently, affluent emotional motion features were explicitly extracted as well as combined using SVR to givebacks the social characteristics. Particularly, weighted features were constructed in an improved way using considering the features' importance.

In 2019, YAN MAO et al [2], developed a unified structure to design the model feature of the groups with emotion for a crowd suffering explosions as well as fires in public regions. The developed evacuation structure comprises 2 modules such as an emotion model for diverse roles in the scenario of a crisis. Emotional varying was a multifarious procedure whereas diverse roles were prejudiced by diverse factors. Both inter-group relationships as well as intra-group structure were considered to reproduce the group occurrence in crisis. The formation and group behaviors were associated with the emotions of diverse roles.

In 2018, Pei Lv et al [3], developed a new crowd behavior development technique by means of emotional contamination in political rallies. Initially, the majority envoy political rally prospects were analyzed in feature as well as design them into 2 categories of abstract cases. Moreover, the "empathy", as well as "extroversion" factors from the OCEAN model, was selected to explain the majority of significant individual personalities in such cases. Based on this, a developed emotional contamination design was developed by integrating the Susceptible-Infected-Recovered design and individual personality in diverse political perspectives. As a final point, the crowd in a political rally was driven to go consistent with the novel possible moving direction produced using emotional contamination as well as the unique way of the individual together.

In 2016, Hajer FRADI et al [4], presented quantify crowd properties using a prosperous set of illustration descriptors. The computation of these descriptors was understood by a new Spatio-temporal model of the crowd. It comprises of designing time-varying dynamics of the crowd by exploiting the local feature tracks. Several medium level representations were extracted to resolve the enduring crowd behaviors from the graph.

In 2018, Yuke Li [5] developed a deep end-to-end method, which collectively regarded as the spatiotemporal information, important to a prosperous considerate of crowd behavior. The derived illustrations were exploited as inputs to an LSTM-based construction to study fundamental spatiotemporal prompts in a single operation for the whole crowd in a certain scene.

In 2019, Xiaofei Wang et al [6], developed a crowd behavior recognition algorithm by integrating the streakline based on the fluid technicalities with a high-accurate variational optical flow model. In scenes, the angular histogram, as well as the ROI, were attained by clustering as well as the computing the dasymetric dot maps of the initial as well as finish points of trajectory, subsequently, integrating the dasymetric dot map and angular histogram information to examine if there were exact crowd behaviors in ROI, as well as therefore to recognize dissimilar kinds of crowd behavior in such scene.

3. Recognition of Crowd Emotion Exploiting the RNN Classifier with the Enhanced-BOA

The behavior of crowd recognition is to recognize the person's emotions from the crowd videos which increase the high objective in the research area to recognize the crowd's emotions. For recognition of emotion, the main requirement is to forecast the crowd's emotions as well as assure few safeguards regarding what is going to take place in the future. One of the major disadvantages related to recognizing the crowd's emotions is to exploit the video cameras, face emotional features activity, as well as voice features turn out to be inefficient. Since those features tolerate mixing up voices, absence of crowd history, uncertainty related to movement and posture in the crowd, etc, there is a requirement for the efficient recognition of emotion approach. Moreover, the automatic emotion of crowd recognition is developed by exploiting the RNN classifier. At first, in the video, the keyframes are selected exploiting the Bhattacharya distance among frames, for that in the first instance the WT of individual frames is decided. The low-frequency band of the WT is fed to the keyframe chosen exploiting the least Bhattacharya distance value. From chosen keyframes, the key points are attained exploits STI descriptor as well as an image with key points is fed to the input of RNN classifier that is tuned optimally exploiting the enhanced-BOA. Fig. 1 shows the developed crowd behavior recognition model of the employing the RNN classifier.

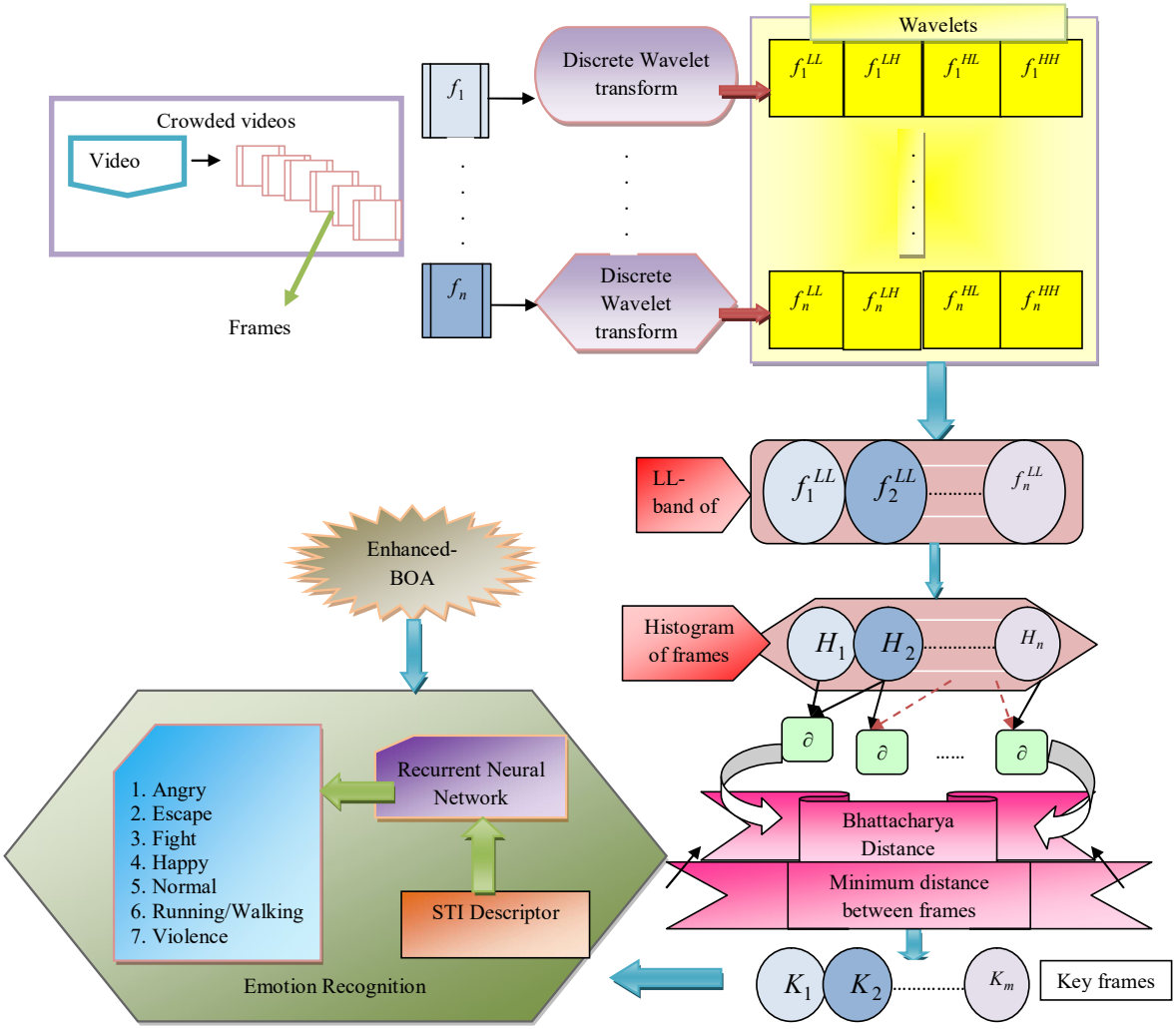


Fig. 1. *Architecture model of the crowd behaviour recognition*

Consider video be V with n number of frames, which is stated in eq. (1), whereas, f_i states i^{th} frame in the video V .

$$V = \{f_1, f_2, ..., f_i, ..., f_n\} \quad (1)$$

3.1 Wavelet Transform Application to the Individual Frames

In the video the individual frames are fed to WT in order to extract the frequency bands from the individual frames. The WT of frames produces important accurate data information so that it decomposes the signal as a function of time, whereas, $f_i^{LH}, f_i^{HL}, f_i^{LL}, f_i^{HH}$ indicates the high and low-frequency bands of i^{th} video frame V.

$$\left[f_i^{LL}, f_i^{LH}, f_i^{HL}, f_i^{HH}\right] = D\left[f_i\right] \tag{2}$$

The low-frequency band comprises important image information as well as hence it is selected for closer dispensation. For all frames, the LL-band is selected as well as a histogram of the band for frames is decided. The individual histogram size is, $[1 \times 256]$.

The wavelet histogram equivalent to i^{th} frame is indicated as, H_i^{LL}. The LL band's histogram of all frames is stated in eq. (3).

$$H = \{H_1, H_2, ..., H_i, ..., H_n\} \tag{3}$$

By exploiting the Bhattacharya coefficient these frames are subjected to selection step of the key frame.

3.2 Keyframe Selection using Bhattacharya Coefficient

The importance of the keyframe chosen is to reduce complexity as well as to discover the main important frame which produces precise information to carry out the classification. The selection of the keyframe principle eases the frame motion detection parameters, audio, effects, as well as other properties which differ by means of time. Therefore, the keyframe maps an object regarding time, whereas property attributes, like audio volume, spatial position, or opacity is indicated, goes after by means of interpolation of the values between keyframes. The keyframes state change over time, as well as these, alter define the state/location of objects. To determine keyframe the distance among 2 keyframes is calculated which defines the state of an object in the establishment as well as the novel state of the object latterly otherwise extinction of object movement. The coefficient of Bhattacharya calculates the similarity between 2 frames defining the conventional overlap between the populations. The coefficient of the Bhattacharya exploits the relation nearness among 2 frames as well as this measure is extremely consistent than any other similarity measures. As stated in eq. (4), the similarity measure is decided which formulates the similarity among the neighboring frame, H_{i+1} as well as the frame H_i.

$$\partial(H_i, H_{i+1}) = \sum_{i=1}^{n} \sqrt{(H_i)(H_{i+1})} \tag{4}$$

In eq. (4), $\partial(H_i, H_{i+1})$ states the coefficient of Bhattacharya for the 2 frames. The number of the keyframes is chosen which is formulated by exploiting the coefficient of the Bhattacharya is stated in eq. (5), m indicates the count of the chosen frames from the total count of frames in attendance in the video.

$$K = \{K_1, K_2, ..., K_i, ..., K_m\}; m < n \tag{5}$$

3.3 STI Descriptor Interesting Points Representation in Frame

The STI features [7] guarantee the motion patterns extraction to be owned by the prehistoric events which consequence because of the object movement that is shown among the frames. On the basis of the small-scale representation, the local features are decided from the frame $K_i(u, v, \delta)$. The small scale space of frame K_i decided exploiting the extraction of STI feature is stated in eq. (6).

$$M\left(., \nu^2, \gamma^2\right) = K_i(u, v, \delta) * N\left(., \nu^2, \gamma^2\right) \tag{6}$$

The eq. (6) is on the basis of the Gaussian convolution kernel and the phrase N is stated in eq. (7).

$$N = \exp\left(\frac{-(u+v)^2}{\left[2\nu_c^2 - \frac{\delta^2}{2\gamma_c^2}\right]}{(2\pi)^3 . \nu_c^4 . \gamma_c^2}\right) \tag{7}$$

The second-moment gradient is on the basis of the spatiotemporal image gradients and it is indicated in eq. (8) and (9).

$$\nabla M = (M_u, M_v, M_\delta)^T \tag{8}$$

$$\varphi\left(.,\nu^2,\gamma^2\right)=N\left(.,s\nu^2,s\gamma^2\right)*\left(\nabla M(\nabla M)^T\right) \quad (9)$$

In eq. (9), local maxima are indicated which discovers the location of the feature and the features size, which are exploited to evaluate the Spatio-temporal features of the frame by means of the automatic chosen of the scaling parameters, (ν,γ). Conversely, the features shape describes the pattern velocity which permits to select the stable features.

$$X=\Delta\eta-\ell\ \text{trace}^3(\eta)\quad \text{over}\quad (u,v,\delta) \quad (10)$$

In eq. (10), (u,v) indicates the spatial-temporal scale parameters. The features associated with the scale and the object velocity is decided to calculate the invariance of the object relating to the object size in the frame. From the local features of Spatio-temporal neighborhoods, spatial facade, and motion of activity/actions in the frame are extracted. To calculate local features, the spatial-temporal jets are calculated, as well as are indicated as eq. (11).

$$J=\left(M_u,M_v,M_\delta,..,M_{\delta\delta\delta\delta}\right) \quad (11)$$

Hence, the STI features indicate interesting points in the frame, as well as a feature vector, is consists of interesting points in image hence object emotion in the frame is familiar efficiently by exploiting the classifier. Indicate the output from the STI descriptor as, P.

4. Enhanced BOA based RNN for Emotion Recognition

In individual frames, interesting points are classified based on the RNN classifier and results in the classes, such as fight, angry, happy, escape, running/walking, normal, as well as violence. For the efficient images, the RNN classifier is exploited and the weights-biases are decided to exploit the developed method optimally.

4.1 RNN Classifier

RNNs are an excellent solution to the issue of designing dynamic alters in a time series. In natural language processing [8], handwriting recognition tasks [10], and speech recognition [9], they are extensively exploited. Fig. 2 shows the structure of RNN. The inputs of the RNN are changed on the basis of the time vector series Y_{t-1},Y_t,Y_{t+1}. Since the series carries on proceed, the hidden layer, H_t, the previous hidden layer, H_{t-1}, as well as it is concurrently affected by the input, Y_t. The eq. (12) and (13) can be exploited to properly explain the process of RNN.

$$H_t=f\left(U\cdot Y_t+P\cdot H_{t-1}\right) \quad (12)$$

$$O_t=g\left(V\cdot H_t\right) \quad (13)$$

In eq. (12), H_t indicates the sample memory at a time, t, i.e. the hidden layer value, as computed using eq. (12). P indicates the previous moment output that is exploited as the weight input at this instant, as well as U indicates the sample weight of input. The value of output, O_t, indicates computed by exploiting the eq. (13), by means of V being the output sample weight. Both g as well as f represent the activation functions, whereas f represents an activation function like ReLU, tanh, else sigmoid. Generally, g represents a softmax activation function. Since the RNN model gets deeper, gradient computed using backpropagation of the hidden layer might disappear else detonate. Even though gradient cropping has the ability to manage by means of gradient explosions, it cannot resolve gradient disappearing. Hence, in the text sequence of language representation, RNN cannot simply detain belief among text elements athwart large distances in series. The exploit of an LSTM can resolve these types of issues. The central part of an LSTM is the shape of the cell (that is cell state). Additionally, it comprises 3 types of gate structures such as the forget, input, as well as output gate. The significant formulations are stated as below:

$$f_t=\lambda\left(W_f\cdot\left[s_{t-1},y_t\right]+a_f\right) \quad (14)$$

$$i_t=\lambda\left(W_i\cdot\left[s_{t-1},y_t\right]+a_i\right) \quad (15)$$

$$o_t=\lambda\left(W_o\cdot\left[s_{t-1},y_t\right]+a_o\right) \quad (16)$$

$$f_t=\lambda\left(W_f\cdot\left[s_{t-1},y_t\right]+a_f\right) \quad (17)$$

$$C_t=f_t*C_{t-1}+i_t*\tanh\left(W_c\cdot\left[s_{t-1},y_t\right]+a_c\right) \quad (18).$$

$$s_t=o_t*\tanh\left(C_t\right) \quad (19)$$

Eq. (14)–(16) states three multiplicative gates such as the input gate, i_t; forget gate, f_t; as well as the output gate, o_t. In eq. (14)–(16), the input is $\left[s_{t-1},y_t\right]$, however, the parameters are diverse. λ indicates the sigmoid activation function. In Eq. (17) C_t indicates the cell state that is attained from C_{t-1} as well as

input at the preceding time step. If the forget gate, f_t, is 0, subsequently state at the preceding moment is turn out to be entirely vacant, with the intention that only the input at this time step will be contemplated. The i_t decides if to obtain input at this time. The ultimate o_t decides if to output cell state.

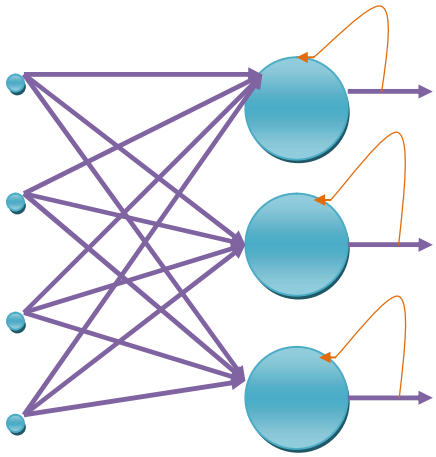


Fig. 2. *Structure model of RNN*

4.2 Enhanced Butterfly Optimization Algorithm

In 2019, BOA is developed and the population of this algorithm is the butterflies group, which is represented as the search agents. In BOA the cost of the objective model varies on the basis of the position of the butterflies [11]. BOA represents a swarm optimization approach in that every agent shares its involvement by other butterflies on the basis of the distribution of fragrance over extent. Using the random generation, this segment is done. The BOA approach is based on the trade-off among smell senses as well as fragrance.

4.2.1 Fragrance

In BOA, the fragrance possesses two kinds of power exponent such as stimulus intensity (I) and sensory modality [13]. The phrase power is the exponent to choose important intensity which performed in a linear response, compression response, and regular. The next phrase, sensory, indicates the procedure the energy form in the same manner as well as the modality states the used raw input in fragrance using the sensors. At last, the phrase stimulus intensity indicates the physical stimulus size which is corresponded using the fitness solution that is while the maximum value of the fragrance in a butterfly, the surrounding butterflies have the ability to sense as well as they attract towards it. The butterflies substance is done using 2 important scenarios cases such as the variation of stimulus intensity (I) as well as the fragrance formulation (f) [12].

The fragrance model is stated in eq. (20), whereas, α and c are in the range [0, 1].

$$f = cI^{\alpha} \tag{20}$$

4.2.2 Butterflies Movement

The approach comprises 3 important stages such as initialization stage, searching stage, as well as finalizing stage at the time of the termination conditions fulfilled. After the initialization of the predominant butterfly swarm, the cost function quantity is estimated for butterflies. Here, the parameters of the approach are also set. Subsequent to approach parameters setting, approach initiates for optimization. In the solution space, the first position of the butterfly is arbitrarily generated. Subsequent to the beginning of the iteration, artificial butterflies move toward novel positions in search space as well as their values of the cost is attained. Then, butterflies produce the fragrance at their positions using eq. (21).

$$z_i^{t+1} = z_i^t + \left(r^2 \times g^* - z_i^t\right) \times f_i \tag{21}$$

In eq. (21), z_i^t indicates the solution vector z_i for i^{th} butterfly, g^* indicates the optimal solution for the iteration t, the fragrance of the i^{th} butterfly is stated using f_i and r is an arbitrary constant among 0 and 1. Eq. (22) is used for the local search in the approach is attained.

$$z_i^{t+1} = z_i^t + \left(r^2 \times z_j^t - z_k^t\right) \times f_i \tag{22}$$

In eq. (22), z_j^t and z_k^t states j^{th} and k^{th} butterflies swarm member in search space. The parameters of the BOA, such as food searching and partner mating for the butterflies, and it is executed in both local and global scales.

Here, a key parameter vector for the BOA is used, that is. $M=[a,c,r]$ are contemplated on the basis of chaos theory. Chaos science examines the regarding the arbitrarily and unpredictable procedures. Chaos theory methods maximum sensitive dynamic systems that affect using any minute alterations. This feature produces points with simpler complexity as well as superior distribution to enhance the distribution points' in the search space. This features enhances the convergence speed of BOA [14] [15]. A common model for chaos theory is stated in eq. (23).

$$M_{i+1}^j = f\left(M_i^j\right),\ j=1,2,....l \tag{23}$$

In eq. (23), $f\left(M_i^j\right)$ is the chaotic model generator function and l indicates the map dimension. Here, Logistic Mapping was exploited as below.

$$a_{k+1} = \gamma a_k\left(1-a_k\right) \tag{24}$$

$$c_{k+1} = \gamma c_k\left(1-c_k\right) \tag{25}$$

$$r_{k+1} = \gamma r_k\left(1-r_k\right) \tag{26}$$

In eq. (24), (25) and (26), k indicates the iteration number, $a_0, c_0, r_0 \in [0, 1]$ indicates the initial arbitrary values and γ indicates a control parameter in the interval. It can be shown that, the formulations $\gamma \in [0, 1]$ —$[0.25, 0.5, 0.75]$ $\gamma = 4$ will be in chaos state. The flowchart diagram of the proposed method is shown in Fig 3.

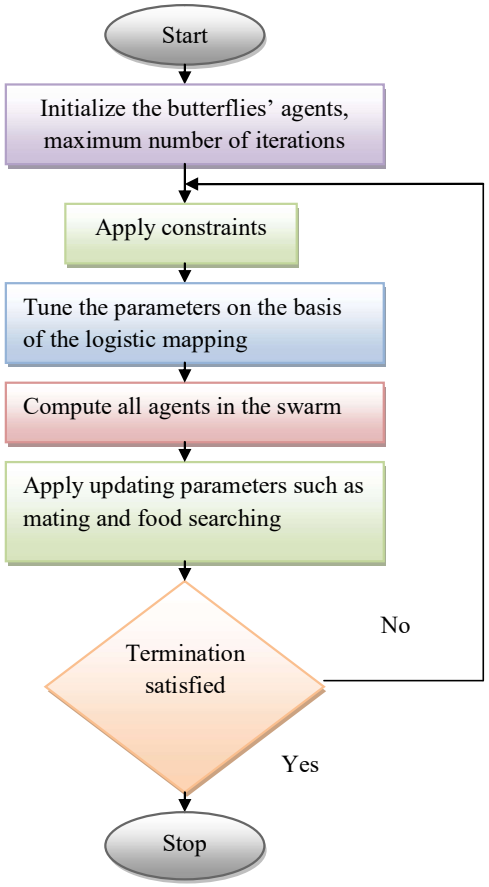


Fig. 3. *Flow chart of the proposed Enhanced BOA*

5. Results and Analysis

5.1 Experimental Procedure

The experimentation was done exploiting a real dataset gathered from the Youtube with a user query. A total of nearly 70 videos were gathered for the detection analysis with the analysis of abnormal and normal crowd videos. The classifier was trained exploiting original features attained from these videos as

well as therefore, upon the arrival of test video, object in crowd video was positioned as well as corresponding behaviors with respect to persons were identified in emotions form.

5.2 Performance Analysis

The performance analysis of the techniques is analyzed on the basis of the accuracy, sensitivity, and specificity, shown in Fig 4 and 5 based on varying K-Fold values and training percentages. From Fig 4 and 5, it is clear that the developed technique obtained a maximum value of accuracy while comparing with the conventional techniques.

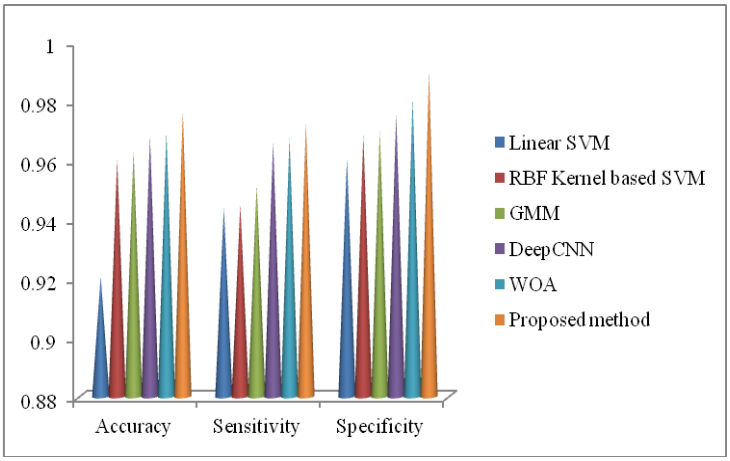


Fig. 4. *Analysis of developed and existing methods for K-Fold*

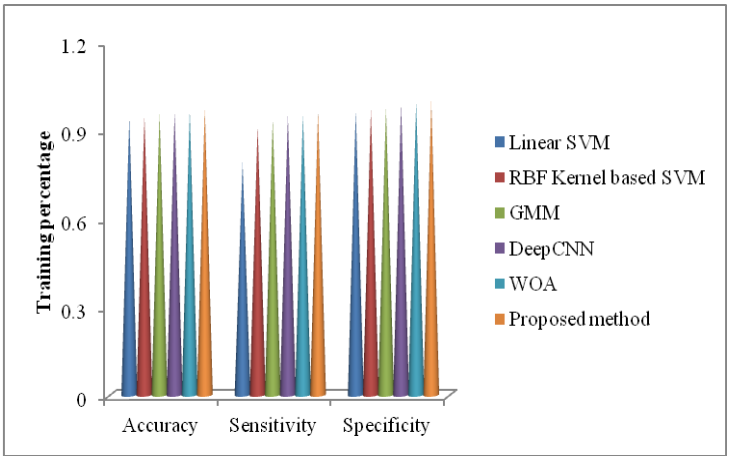


Fig. 5. *Analysis of developed and existing methods for Training percentage*

5. Conclusion

- In this paper, the automatic Crowd Emotion Recognition was developed by exploiting the RNN classifier.
- The conventional techniques used the voice, face, and etc, to identify the crowd behaviors that were inefficient as they are unable to deal with the dynamics of the object in the video.
- At first, from the input video, the keyframes were selected in the developed technique of automatic recognition so the frames with important features were selected.
- In the video, the frames were fed to the wavelet transform and the Bhattacharya distance was exploited, in order to choose the keyframes.
- After the keyframes selection, by exploiting the STI descriptor, the features from the frames were selected, as well as the features were subjected as input to the RNN classifier.
- By employing the enhanced-BOA method the classifier was tuned optimally.
- At last, the analysis of the approach exploiting the dataset reveals that the developed technique obtained a superior accuracy, specificity, and sensitivity.

References

[1] Y. Zhang, L. Qin, R. Ji, S. Zhao, Q. Huang and J. Luo, "Exploring Coherent Motion Patterns via Structured Trajectory Learning for Crowd Mood Modeling," in IEEE Transactions on Circuits and Systems for Video Technology, volume. 27, number. 3, page no. 635-648, March 2017.

[2] Y. Mao, X. Fan, Z. Fan and W. He, "Modeling Group Structures With Emotion in Crowd Evacuation," IEEE Access, volume. 7, page no. 140010-140021, 2019.

[3] P. Lv, Z. Zhang, C. Li, Y. Guo, B. Zhou and M. Xu, "Crowd Behavior Evolution With Emotional Contagion in Political Rallies," in IEEE Transactions on Computational Social Systems, vol. 6, no. 2, pp. 377-386, April 2019.

[4] H. Fradi, B. Luvison and Q. C. Pham, "Crowd Behavior Analysis Using Local Mid-Level Visual Descriptors," IEEE Transactions on Circuits and Systems for Video Technology, volume. 27, number. 3, page no. 589-602, March 2017.

[5] Y. Li, "A Deep Spatiotemporal Perspective for Understanding Crowd Behavior," in IEEE Transactions on Multimedia, volume. 20, number. 12, page no. 3289-3297, Dec. 2018.

[6] X. Wang, Z. He, R. Sun, L. You, J. Hu and J. Zhang, "A Crowd Behavior Identification Method Combining the Streakline With the High-Accurate Variational Optical Flow Model," in IEEE Access, vol. 7, pp. 114572-114581, 2019.

[7] C. Schuldt, L. Barbara, and S.- Stockholm, “Recognizing Human Actions : A Local SVM Approach ∗ Dept . of Numerical Analysis and Computer Science,” Pattern Recognition,” ICPR. Proc. 17th Int. Conf., volume. 3, page no. 32–36, 2004.

[8] W. Yin, K. Kann, M. Yu, H. Schütze, Comparative study of cnn and rnn for nat- ural language processing, 2017.

[9] S. Han , J. Kang , H. Mao , Y. Hu , X. Li , Y. Li , D. Xie , H. Luo , S. Yao , Y. Wang , Ese: efficient speech recognition engine with sparse lstm on fpga, in: Proceedings of the ACM/SIGDA International Symposium On Field-Programmable Gate Arrays, ACM, page no. 75–84, 2017 .

[10] C. Wigington , S. Stewart , B. Davis , B. Barrett , B. Price , S. Cohen , Data augmenta- tion for recognition of handwritten words and lines using a cnn-lstm network, in: 14th IAPR International Conference On Document Analysis and Recognition (ICDAR), IEEE, page no. 639–645. 2017 .

[11] R. B. Blair and A. E. Launer, "Butterfly diversity and human land use: Species assemblages along an urban grandient," Biological conservation, volume. 80, page no. 113-125, 1997.

[12] E. Pollard and T. J. Yates, Monitoring butterflies for ecology and conservation: the British butterfly monitoring scheme: Springer Science & Business Media, 1994.

[13] M. S. Bodri, "Puddling Behavior of Temperate Butterflies: Preference for Urine of Specific Mammals?,"The Journal of the Lepidopterists' Society, volume. 72, page no. 116-121, 2018.

[14] D. Yang, G. Li, and G. Cheng, "On the efficiency of chaos optimization algorithms for global optimization," Chaos, Solitons & Fractals, volume. 34, page no. 1366-1375, 2007.

[15] C. Rim, S. Piao, G. Li, and U. Pak, "A niching chaos optimization algorithm for multimodal optimization,"Soft Computing, volume. 22, page no. 621-633, 2018.

[16] R. Cristin,Dr.V.Cyril Raj and Ramalatha Marimuthu,"Face Image Forgery Detection by Weight Optimized Neural Network Model",Multimedia Research,Volume 2, Issue 2, April 2019.

[17] Raviraj Vishwambhar Darekar,Ashwinikumar Panjabrao Dhande,"Emotion Recognition from Speech Signals Using DCNN with Hybrid GA-GWO Algorithm",Multimedia Research,Volume 2, Issue 4, October 2019.