# Localization of Strangeness for Real Time Video in Crowd Activity Using Optical Flow and Entropy

https://doi.org/10.3991/ijoe.v19i07.38869

Ali Abid Hussan Altalbi(✉), Shaimaa Hameed Shaker, Akbas Ezaldeen Ali
Computer Science Department, University of Technology, Baghdad, Iraq
CS.20.02@grad.uotechnology.edu.iq

**Abstract—**Anomaly detection, which is also referred to as novelty detection or outlier detection, is process of identifying unusual occurrences, observations, or events which considerably differ from the bulk of data and do not fit a predetermined definition of typical behavior. Medicine, cybersecurity, statistics, machine vision, law enforcement, neurology, and financial fraud are just a handful of the industries where anomaly detection is used. In the presented study, an online tool is utilized to identify crowd distortions, which could be brought on by panic. An activity map is produced with the use of numerous frames to show the continuity regarding the flow over time following the global optical flow has been calculated in the quickest time and with the highest precision possible utilizing the Farneback approach to calculate the magnitudes. Utilizing a specific threshold, the oddity in the video will be picked up by the activity map's generation of an entropy. The results indicate that the maximum entropy level for indoor video is <0.16 and the maximum entropy level for outdoor video is >0.45. A threshold of 0.04 is used to determine whether a frame is abnormal or normal.

**Keywords—**abnormal event detection, RGB frames, gray frames, optical flow, farneback method, entropy, activity map

## 1 Introduction

The recognition of human behavior depending on optic flow was discussed in the presented work. A difficult and crucial area of computer vision is human action recognition. Recognizing human action is actually human to many computer vision applications, including surveillance [1], video retrieval, HCI, and scene computer. Noise-filled, low-frame-rate, and low-resolution images provide a few significant challenges to action recognition for surveillance applications [2]. One of the crucial and difficult study areas is abnormal detection, which is based on techniques used in video image processing like visual target tracking and scene understanding [3]. Abnormal crowd behavior detection is the process of extracting specific data from a video or surveillance image sequence that represents the crowd's abnormal behavior, like group behavior characteristics and population density, and performing classification. With regard to computer vision, detecting abnormal crowd behaviors is crucial [4]. A broad spectrum of the economic and application values in the social safety and security exists when relevant information is extracted from large-scale surveillance videos and improved accuracy of recognition regarding the emergencies and abnormal behaviors [5]. In the case when there is an anomaly in a time, abnormal behavior detection indicates the initial frame which could categorize the event and identify the abnormal behavior in time [6]. The appropriate features must be extracted from the video sequence and classified in order to successfully identify between the abnormal and normal events in the video [7]. Researchers frequently utilize temporal and spatial features, like optical flow, in conventional feature extraction approaches [8].

## 2 Related work

Numerous effective approaches were put forth, such as the ones that efficiently expand beyond the image domain into the video and action recognition domains. Yet, there is still room for development with regard to present techniques, particularly for real-world movies and videos, which feature a dynamic background, people with varying postures, and partial occlusions [9].

Various studies concentrate on part-based techniques, which just analyze the 'interesting' parts regarding the video instead of the entire video, in order to overcome the shortcomings [10]. Those "parts" may be corner and spatiotemporal interest point trajectories or flow vectors. Even though part-based techniques are promising [11], they still struggle with erroneous background clutter and motion-induced background detection and tracking of important sections, which precludes a clear and useful representation [12].

Lei Wang (2019) utilized outdated manual video representation approaches for action recognition through using a CNN-based hallucination step. The I3D model (among others) thrives on the combination of its output with Improved Dense Trajectory (IDT) and extracted with its low-level video descriptors that have been encoded through Fisher Vectors (FV) and Bag-of-Words (BoW), the model saves 20h–55h of the computations and yields cutting-edge results on 4 publicly available datasets [13].

An AutoEncoder-based network termed the cascade deep AutoEncoder (CDA) is used after the feature descriptor extraction procedure described by Tian Wang (2020). A unique descriptor that captures the multi-frame optical flow information is used to express the movement information. Following that, CDA network is trained using the feature descriptor from the normal samples. Lastly, the reconstruction error of CDA during testing procedure helps to identify the abnormal samples. By testing the suggested technique on many video surveillance datasets [14].

Yu Hao (2021), when considering group behavior, the network's input is the optical flow data between the RGB images and the video frames. A global optical flow descriptor was created by weighting and fusing the crowd's velocity, direction, energy, and acceleration. In addition, a single frame's time of the original image is used for extracting the crowd trajectory map. In order to model dependence link between long-time video frames and produce the final network classification results, "enetwork" employs two network branches in order to learn information about the spatial and temporal dimensions, respectively. Results from simulation tests demonstrate that the suggested recognition could successfully recognize multiple datasets, and that incorporating interframe motion information could considerably enhance the effectiveness of abnormal behaviour detection [1].

Yiheng Li (2022), the optical flow predicted via the models trained with the use of various projection techniques was combined in this study utilizing a new multi-projection fusion framework. It gains the ability to merge the complimentary information in the results of optical flow under various projections. Create the first panoramic dataset for panoramic optical flow in flow to train NNs and test optical techniques for doing so. The dataset's experimental findings show that technique performs better than already-developed deep networks and other alternatives for handling 360° content [15].

Considerable progress was made thanks to the ongoing, in-depth study of crowd abnormal behavior detection algorithms. The detection of the technology for detecting abnormal behavior in complex settings is still a difficult challenge. This study suggests a crowd abnormal behavior detection algorithm depending upon global optical flow and entropy level for increasing the robustness and accuracy of crowd abnormal behavior detection algorithm, for low performance of the crowd abnormal behavior detection resulting from complex environment. More details will be indicated in the next sections [16].

## 3 Method and materials

The speed and direction of crowds typically follow a comparable pattern. People will, on the other hand, run away rapidly from an abnormal event out of worry for their safety [17]. Yet, the crowd's abnormal behavior is characterized by its rapid movement speed, abrupt acceleration, and noticeable concentration of movement in one direction or balance in several directions [18], as well as its broad movement range, large pace, chaotic trajectory, and panic expression. Calculating attributes like acceleration, speed, direction, and motion amplitude is one of them [19]. Conventional abnormal behavior detection methods simply employ RGB images as the network's input, ignoring the comparatively easy and expressible via optical flow concealed motion information in video sequence [20].

### 3.1 Estimating optical flow

The relation between temporal changes and spatial properties in images, or motion, is a key feature of frame sequences; it demonstrates the dynamics of frames [21]. The presentation of motion data from an image sequence has been referred to as the optical flow estimation; optical flow is a 2D motion map that projects the scene's 3D motion onto the image plane. Keep in vision that [22], if camera position parameters are given, the optical flow module may also be utilized for extracting parallax in static stereoscopic vision [23]. Based on the presumptions that pixel intensities don't fluctuate over time and that nearby pixels move similarly, optical flow operates [24]. Additional presumptions can include the motion being locally smooth or the visual gradients being constant. The first two hypotheses result in equation (1):

$$U.I_x + V.I_y = I_t \tag{1}$$

In which $V$ and $U$ represent the optic flow components in vertical and horizontal directions. Also, $I_x$ and $I_y$ and $I_t$ are brightness function derivatives with respect to $x$, $y$ (coordinates of the image) and $t$ (time) [25].

The pattern of apparent motion regarding surfaces, objects, and edges in a visual image that is brought on by the motion of the observer in motion to the scene is known as optical flow or optic flow [26]. The spatial and temporal features of the optical flow approach are good [27]. As illustrated in Figure 1, it may be utilized to represent movement information like the direction and speed of the moving target since it could recognize independent moving objects in crowd under scenes with unknown prior knowledge and accurately compute their speed of movement [28].

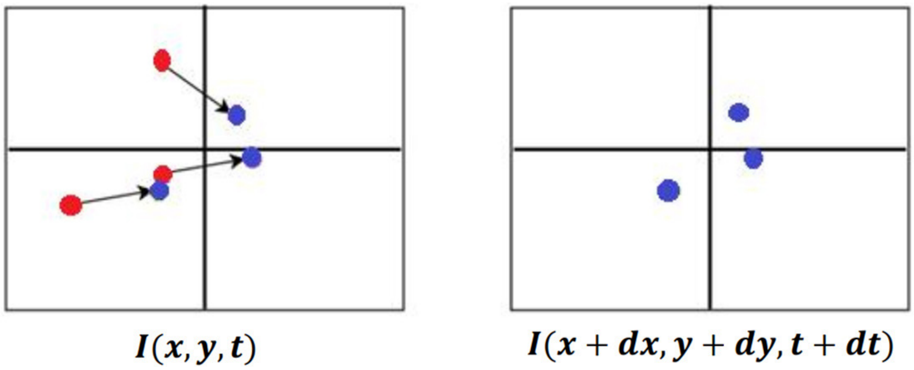


**Fig. 1.** Pixels displacement in two consecutive images. Blue pixels and red pixels are corresponding to image at time *dt* and *t* respectively [28]

Consider a pixel $I(x,y,t)$ in first frame, it moves to the next frame by taking $dt$ time.

Since the image pixels and intensity are the same from one frame to the next, equation (2) states that the pixel displacement is ($dx$, $dy$):

$$I(x, y, t) = I(x+\Delta x, y+\Delta y, t+\Delta t) \tag{2}$$

In which $I(x,y,t)$ represent the assumed pixel at location $(x,y,t)$, $\Delta x$, $\Delta y$, $\Delta t$ represent the movement between thetwo frames and c is a real valued constant number.

### 3.2 Global and local features

Local and global optical flow approaches are distinguished from one another. While global/dense approaches process every pixel in the image, local approaches only require us to process a small portion of total number of the pixels in the image. For additional flow data to be available in sparse/global flow extraction techniques [29]:

Descriptors regarding local image neighborhoods known as **local features** are evaluated at multiple interest points. We outline typical applications for local features in this section [30]. The possibility of different numbers of the feature points in every one of the images makes comparing images more difficult when handling local features, which is one of the main problems [31].

**Global features** are those where each image is captured by a single feature vector that contains data from the entire image. The components of the image, such as specific objects or regions, receive no consideration. Once the features of each image have been calculated, mighted use a distance metric to determine how comparable any two images [32].

A set of features known as a video global descriptor [23] indicates the video as a whole and is hence the ideal descriptor for describing typical video patches. Global features, which characterize a whole image, are frequently used by object recognition systems. The majority of texture and shape descriptors fall under this category [33]. These features are appealing since they result in relatively compact representations of images, in which each one of the images corresponds to a point in a high dimensional feature space, as shown in Figure 2 [34].

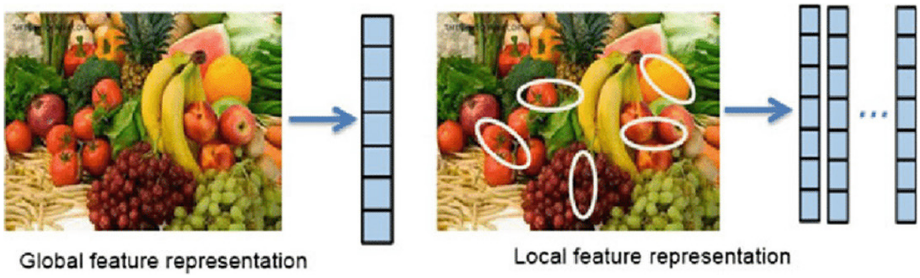


**Fig. 2.** Global and local image features representation [34]

Global features could be understood as an aspect of an image that affects all pixels. This characteristic could be the color histograms, edges, texture, or even a particular descriptor that was taken from one of the image's processing filters. While maintaining invariant to changes in illumination and viewpoint, local feature representation's fundamental objective is to distinguishably represent an image depending on a few key regions [35]. A set of local feature descriptors are therefore obtained from a set of the image regions that are known as regions of interest and used for representing the image depending on its local structures [36].

### 3.3 Farneback method

Figure 3 illustrates how the Farneback technique, a two-frame motion estimation algorithm, utilizes polynomial expansion to approximate the neighborhood regarding each one of the image pixels [28]:

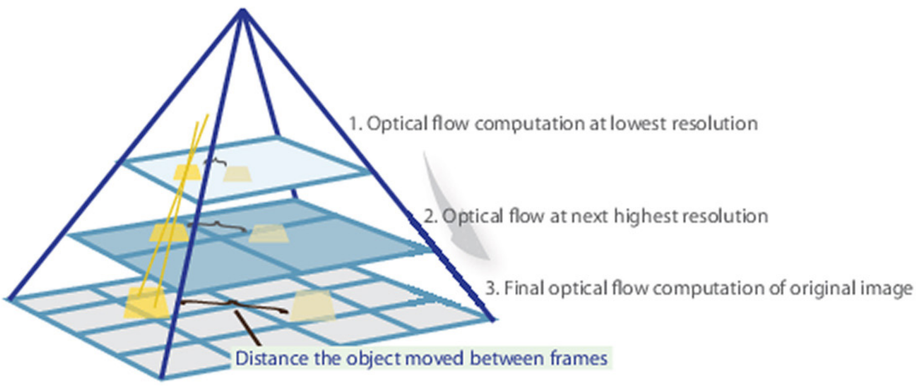


**Fig. 3.** The Farneback optical flow [28]

Each level of the image pyramid created by the Farneback algorithm has a lesser resolution than the level before it. The algorithm may follow points at multiple resolution levels, starting at the lowest level, in a case when choosing a pyramid level higher than 1. The algorithm can manage bigger point displacements between frames by expanding the number of pyramid levels. But there are also more calculations.

Returns an optical flow object that can be used for the estimation of speed and direction of moving objects in a video, as shown in Figure 4 [37].

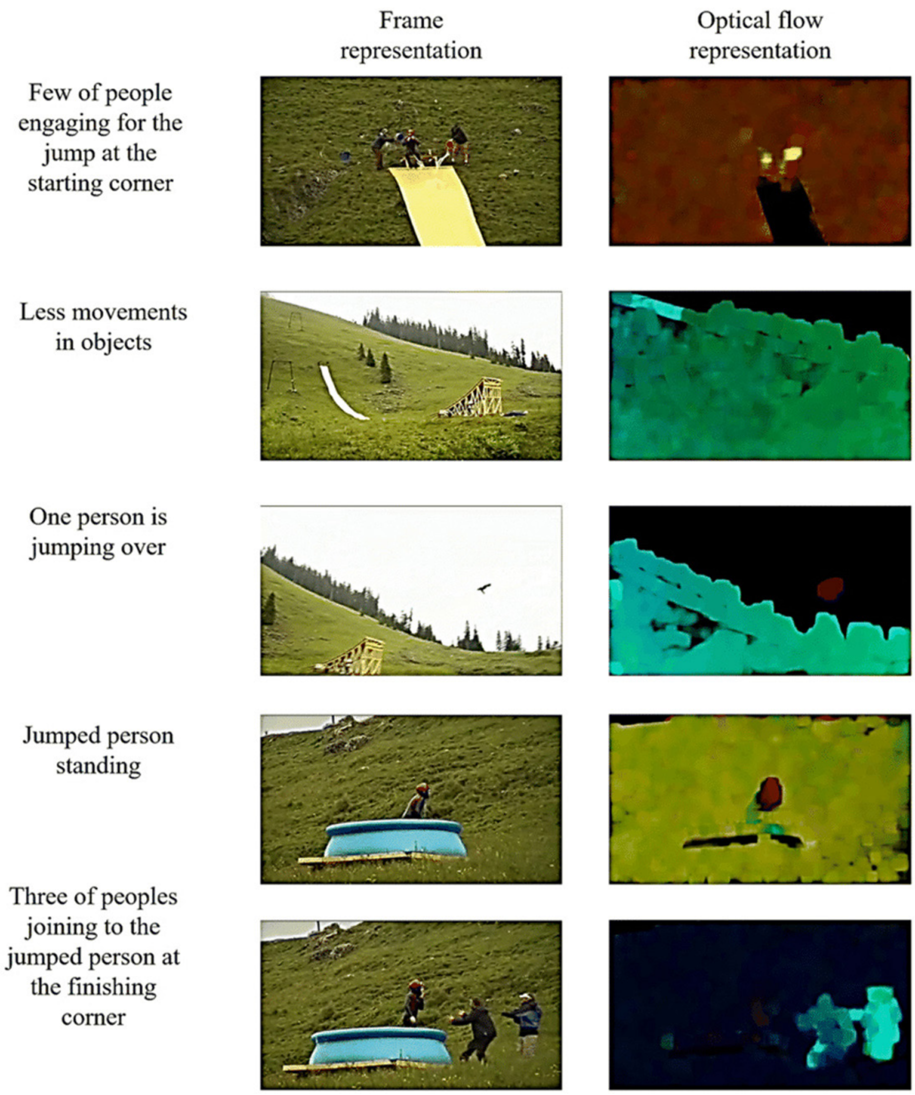


**Fig. 4.** Demonstration of behavior of dense optical flow (utilizing the Farneback approach) under various dynamic content types. Every one of the frames is obtained from a sequence that involves a movement of separate people or objects [37]

Up to convergence, tracking is performed starting at the lowest resolution level. The keypoints for the subsequent level are propagated from the point locations that were discovered at a level [38]. With each one of the levels, the algorithm improves the tracking in this way [39]. The algorithm might manage large pixel motions that could span distances bigger than the size of the neighborhood, thanks to the pyramid decomposition [40].

## 4 Proposed behavior detection Algorithm

focused on two mechanisms (optical flow and entropy level) for the purpose of extracting global features because A video global descriptor can be defined as a set of the features describing a video as a whole and thus, is best capable of describing the normal patches of the video.

UMN dataset was used for the information dataset. Also, the dataset includes a few recordings for each of the three types of scenes—Indoor, Lawn, and Plaza—as well as three different scenes with those names.

| ***Algorithm 1: Abnormal Behavior Detection*** |
|---|
| *Input:RGB video*<br>*Output: give ALARM on video when anomaly detected*<br>*Begin*<br>*Repeat*<br>*while (video frames not terminated) do*<br>*Convert video into Sequential frames*<br>*Convert RGB frames into gray*<br>*Apply Gaussian filter on each frame.*<br>*estimate activity map (Optical Flow)*<br>*Apply to post-preparing (Median Filter)*<br>*calculate the difference between two frame*<br>*estimate entropy from the difference*<br>*If the Entropy level > threshold*<br>*give ALARM on video (abnormal)*<br>*else*<br>*give normal video*<br>*end while*<br>*until frames eand*<br>*END* |

At first re-processing the video (by converting the video from colored to gray because the important thing for me is the movement of the object and not the color density of it. After that, that video is converted into a number of sequential frames through resolution and how many frames per second, after that, applying a Gaussian filter to reduce and delete noise).

The video stream is broken up into frames in the pre-processing module. For each frame using this approach, dynamic and static features were extracted. Therefore, the highlight generating algorithm considers decoding video into individual frames as a preprocessing step.

After that, for each old and new frame, the percentage change was calculated or by calculating the size of the optical flow, after which the entropy of the frames was measured, and through experiment a certain threshold was determined for each video showing abnormal conditions in the video by printing a sign or A warning informs about the presence of abnormal movement in that video and prints a graph of the entropy with respect to time to show how it has changed depending on the abnormal movement of objects in the video over time, as shown in Figure 5:

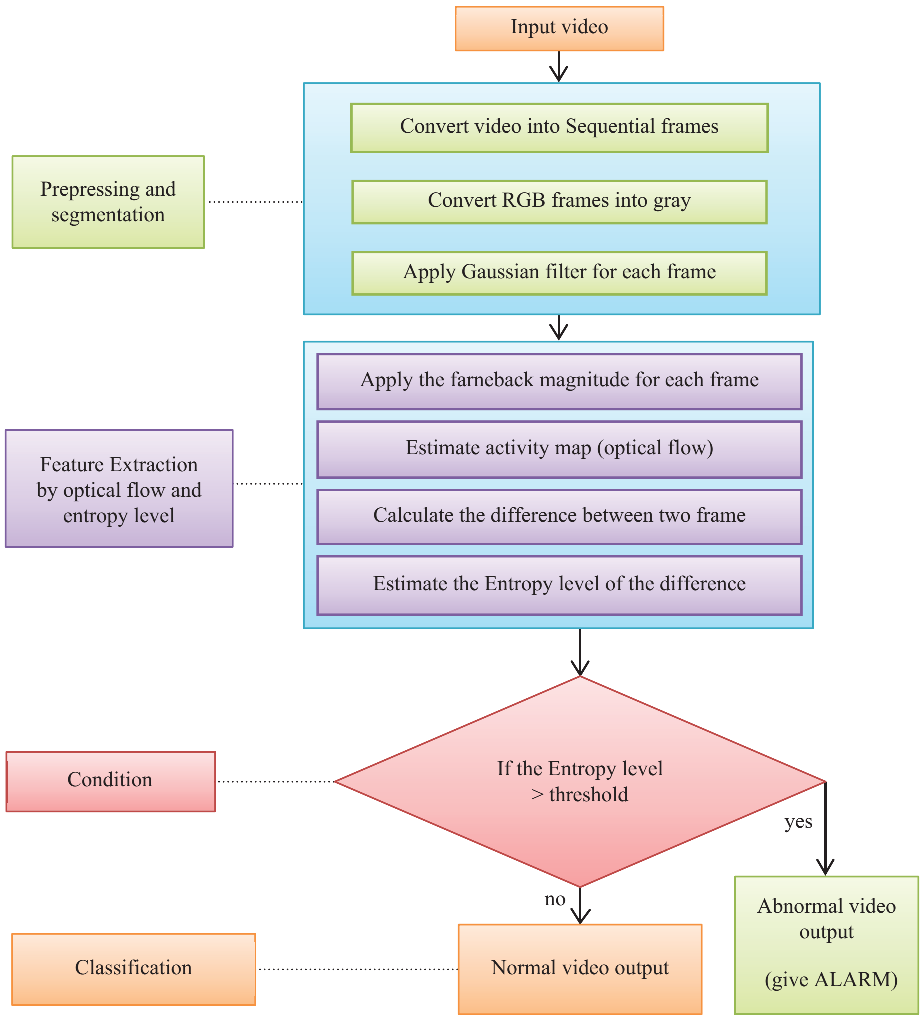


**Fig. 5.** Abnormal behavior detection algorithm based on global optical flow and entropy level

# 5 Experimental results

Utilized a dataset from the University of Minnesota (UMN) to find unusual crowd detection. UMN dataset includes three scenes—the first is outdoor, the second is indoors, and the third is outdoors—each video at a frame rate of 30. (320 × 240 × 3) is the RGB frame size; see Figure 6. The dataset contains the raw data needed to classify abnormalities.

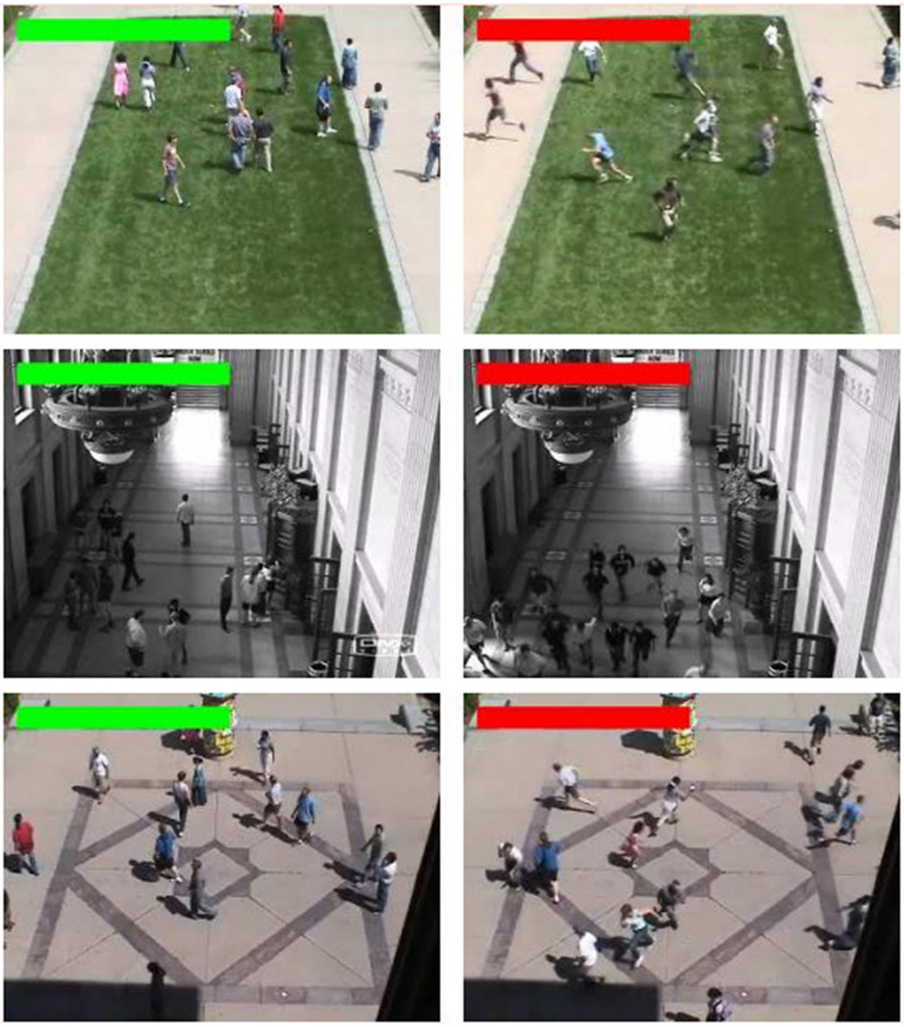

**Fig. 6.** UMN data-set samples for each of 3 scenes: Normal (green) and Abnormal (red)

The limitations of this dataset are some. Only 3 anomaly scenes are present in the data-set, and there are significant spatial and temporal differences between the abnormal and normal frames. There is no pixel-level ground truth in this data-set. Considering these restrictions. Only the global detector is employed since the data-set is straightforward and anomaly localization is not crucial. White shading denotes "there is movement in that piece of the mean picture," and the movement guide will contain shading spaces ranging from dark to white. At that pixel, nothing changed using the dark approaches. Additionally, there are dim pixels, which demonstrate less development. Entropy could also help in determining how much movement there is in the differentiation outline.

When applying the system to the first video of the data set, which is crowded and outdoor, it was noticed after taking pictures of the results before and after the occurrence of the strangeness in the video and the sudden movement of the crowds, the big difference in optical flow when the activity map changes for the people in that video, and calculating the difference between the previous and subsequent activity map in a sequence Frames, as shown in Figure 7.

In the first video, after producing the optical flow, it utilizing Farneback technique, an activity map has been created with the use of multiple frames for showing the continuity regarding the flow over time. Then, the activity map has been utilized to generate the first entropy and show the max entropy level for outdoor greater than 0.45, Depending on a threshold 0.04, anomaly detection in the video is identified and ALARM is given, as shown in Figure 8.

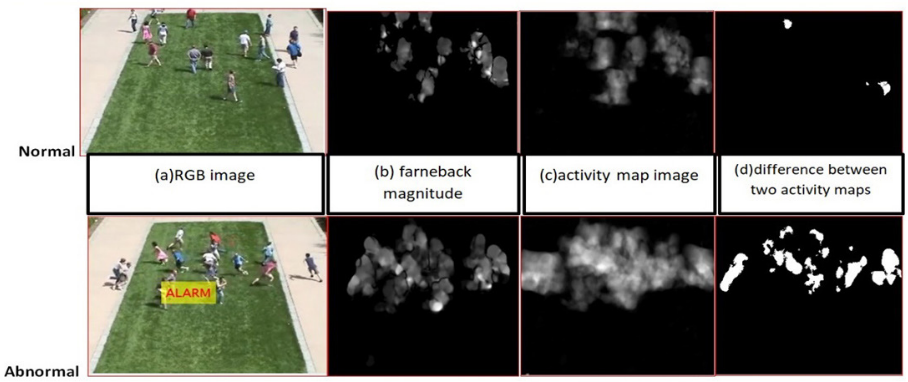


**Fig. 7.** Outdoor localization of strangeness from crowded video scenes based on optical flow

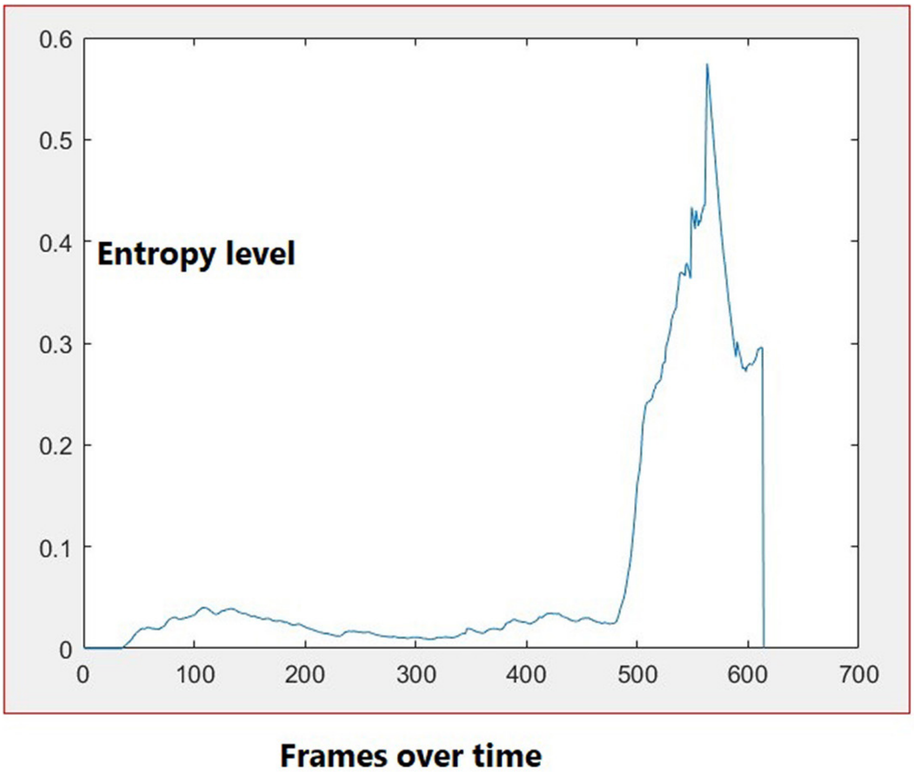


**Fig. 8.** The result of the entropy level of the outdoor video

When applying the system to the second video of the data set, which is crowded and indoor, it was noticed after taking pictures of the results before and after the occurrence of the strangeness in the video and the sudden movement of the crowds, the big difference in entropy and optical flow when the activity map changes for the people in that video, and calculating the difference between the previous and subsequent activity map in a sequence Frames, as shown in the Figure 9.

In the second video of the dataset, after producing the optical flow, it using Farneback technique, an activity map is created with the use of multiple frames to show the continuity of the flow over time. After that, activity map is utilized in order to generate the second entropy and show the max entropy level for outdoor greater than 0.165, Depending on a threshold 0.04, anomaly detection in the video is identified and ALARM is given, as shown in Figure 10.


**Fig. 9.** Indoor localization of strangeness from crowded video scenes based on optical flow

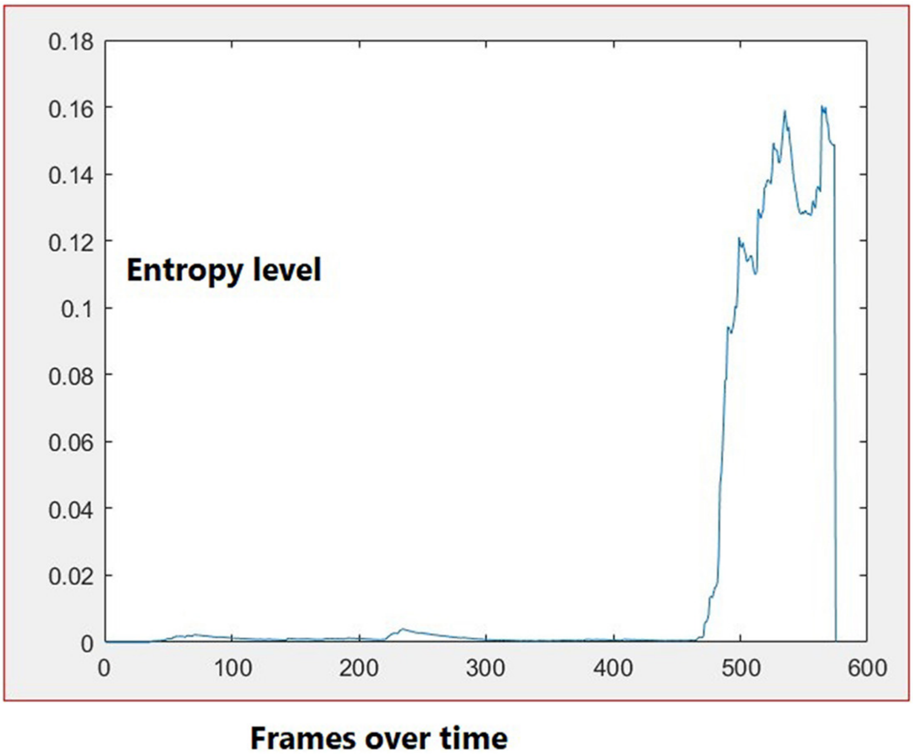


**Fig. 10.** The result of the entropy level of the indoor video

When applying the system to the third video of the data set, which is crowded and outdoor, it was noticed after taking pictures of the results before and after the occurrence of the strangeness in the video and the sudden movement of the crowds, the big difference in entropy and optical flow when the activity map changes for the people in that video, and calculating the difference between the previous and subsequent activity map in a sequence Frames, as shown in the Figure 11.

In the third video of the data set, after producing the optical flow, it with the use of Farneback technique, an activity map is created with the use of multiple frames for showing the continuity regarding flow over time. Then, the activity map is utilized to generate the third entropy. and show the max entropy level for outdoor greater than 0.48, Depending on a threshold 0.04, anomaly detection in the video is identified and ALARM is given, shown in the Figure 12.

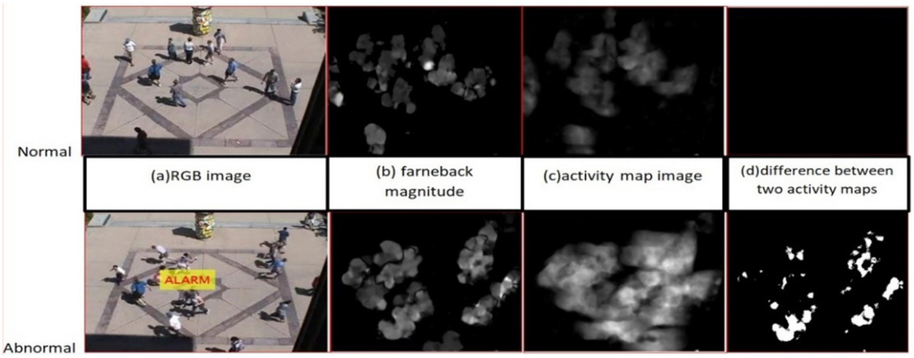


**Fig. 11.** Outdoor localization of strangeness from crowded video scenes based on optical flow

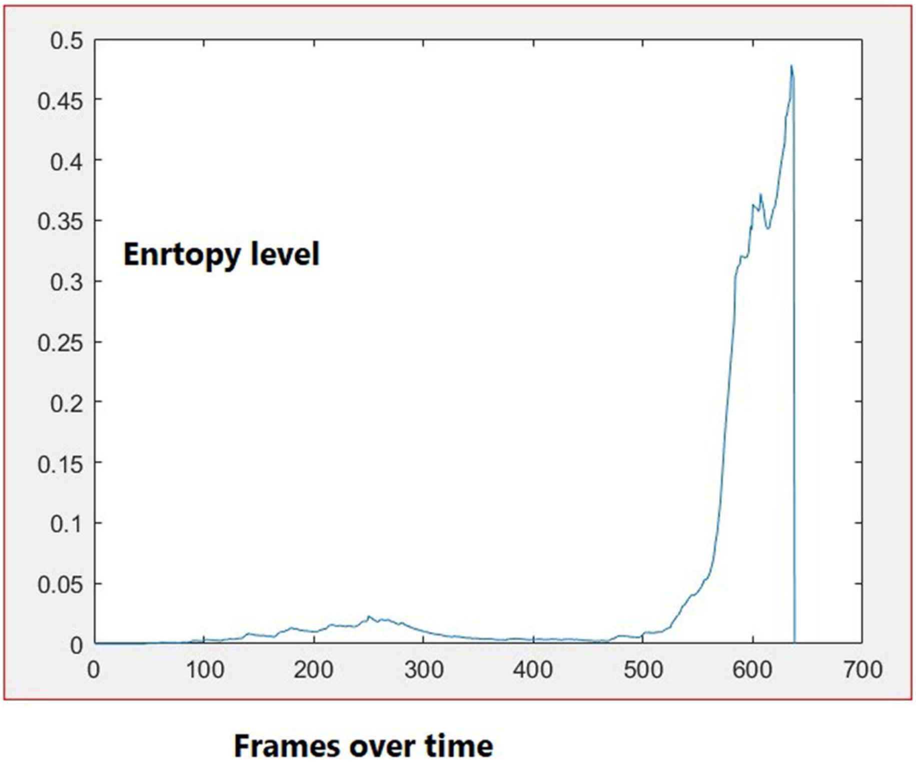


**Fig. 12.** The result of the entropy level of the outdoor video

The result shows the maximum entropy level for outdoor video is more than 0.45, however, the maximum entropy level for indoor video is more than 0.16. with a threshold 0.04 for classifying each frame as abnormal or normal in a given sequence, as shown in Table 1.

**Table 1.** Shown the entropy level and threshold for indoor and outdoor videos

| Type of Video | Number of Frames | Entropy Threshold | Max Entropy Level |
|---|---|---|---|
| First-outdoor video | 614 | 0.04 | 0.58 |
| Second-indoor video | 575 | 0.04 | 0.165 |
| Third-outdoor video | 638 | 0.04 | 0.48 |

## 6 Conclusions

One of the focus areas for vision research is crowd abnormal behavior detection, which is also a key component of intelligent surveillance. The smart security of airports, schools, shopping centers, and communities makes extensive use of this technology. The goal of this study article is to put into practice an on-line solution in order to identify abnormal crowd behavior, which could be brought on by panic. After reprocessing the video, optical flow was applied, the entropy level was assessed, and using a predetermined threshold, it was decided whether the video was abnormal or normal. The results show that the system could accurately identify and localize anomalies as they occur in the video, depending on the variation and change in events and movement through time.

## 7 References

[1] Y. Hao, Y. Liu, J. Fan, and Z. Xu, "Group abnormal behaviour detection algorithm based on global optical flow," *Complexity*, vol. 2021, 2021, https://doi.org/10.1155/2021/5543204

[2] A. Padilha-Feltrin, D. A. Q. Rodezno, and J. R. S. Mantovani, "Volt-VAR multiobjective optimization to peak-load relief and energy efficiency in distribution networks," *IEEE Trans. Power Deliv.*, vol. 30, no. 2, pp. 618–626, 2014, https://doi.org/10.1109/TPWRD.2014.2336598

[3] S. J. Berlin and M. John, "Spiking neural network based on joint entropy of optical flow features for human action recognition," *Vis. Comput.*, vol. 38, no. 1, pp. 223–237, 2022, https://doi.org/10.1007/s00371-020-02012-2

[4] S. P. Smith, "Comparing virtual environments for cybersickness using a cumulative optical flow entropy metric," *IEEE Access*, vol. 9, pp. 68898–68904, 2021, https://doi.org/10.1109/ACCESS.2021.3077899

[5] X. Zhang, X. Shu, and Z. He, "Crowd panic state detection using entropy of the distribution of enthalpy," *Phys. A Stat. Mech. its Appl.*, vol. 525, pp. 935–945, 2019, https://doi.org/10.1016/j.physa.2019.04.033

[6] H. Hao, X. Li, and M. Li, "A Detection Method of Abnormal Event in Crowds Based on Image Entropy," in *Proceedings of the 2019 4th International Conference on Intelligent Information Processing*, 2019, pp. 362–367, https://doi.org/10.1145/3378065.3378134

[7] C. Xin, C. Wang, Z. Xu, J. Wang, and S. Yan, "Marker-free fatigue crack detection and localization by integrating the optical flow and information entropy," *Struct. Heal. Monit.*, p. 14759217221103252, 2022, https://doi.org/10.1177/14759217221103251
[8] G. Bao, D. Li, and Y. Mei, "Key frames extraction based on optical-flow and mutual information entropy," *J. Phys. Conf. Ser.*, vol. 1646, no. 1, 2020, https://doi.org/10.1088/1742-6596/1646/1/012112
[9] F. Afza *et al.*, "A framework of human action recognition using length control features fusion and weighted entropy-variances based feature selection," *Image Vis. Comput.*, vol. 106, 2021, https://doi.org/10.1016/j.imavis.2020.104090
[10] S. Pinitkan and N. Wisitpongphan, "Abnormal activity detection and notification platform for real-time ad hoc network," *Int. J. online Biomed. Eng.*, vol. 16, no. 15, pp. 45–63, 2020, https://doi.org/10.3991/ijoe.v16i15.16065
[11] S. A. B. Salman, S. Al-Janabi, and A. M. Sagheer, "Valid blockchain-based e-voting using elliptic curve and homomorphic encryption," *Int. J. Interact. Mob. Technol.*, vol. 16, no. 20, pp. 79–97, 2022, https://doi.org/10.3991/ijim.v16i20.33173
[12] A. Nadeem, A. Jalal, and K. Kim, "Automatic human posture estimation for sport activity recognition with robust body parts detection and entropy markov model," *Multimed. Tools Appl.*, vol. 80, no. 14, pp. 21465–21498, 2021, https://doi.org/10.1007/s11042-021-10687-5
[13] L. Wang, P. Koniusz, and D. Huynh, "Hallucinating IDT descriptors and I3D optical flow features for action recognition with CNNs," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2019-Octob, pp. 8697–8707, 2019, https://doi.org/10.1109/ICCV.2019.00879
[14] T. Wang, M. Qiao, A. Zhu, G. Shan, and H. Snoussi, "Abnormal event detection via the analysis of multi-frame optical flow information," *Front. Comput. Sci.*, vol. 14, no. 2, pp. 304–313, 2020, https://doi.org/10.1007/s11704-018-7407-3
[15] Y. Li, C. Barnes, K. Huang, and F.-L. Zhang, "Deep 360° optical flow estimation based on multi-projection fusion," pp. 336–352, 2022, https://doi.org/10.1007/978-3-031-19833-5_20
[16] E. Agustsson, D. Minnen, N. Johnston, J. Ballé, S. J. Hwang, and G. Toderici, "Scale-space flow for end-to-end optimized video compression," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 8500–8509, 2020, https://doi.org/10.1109/CVPR42600.2020.00853
[17] I. W. Ghindawi, "A proposed registration method using tracking interest features for augmented reality," vol. 34, 2016, 2018.
[18] H. L. Masoner and A. Hajnal, "Does optic flow provide information about actions?" *Attention, Perception, Psychophys.*, vol. 85, no. 4, pp. 1287–1303, 2023, https://doi.org/10.3758/s13414-023-02674-9
[19] H. Xu, L. Li, M. Fang, and F. Zhang, "Movement human actions recognition based on machine learning," *Int. J. Online Eng.*, vol. 14, no. 4, pp. 193–210, 2018, https://doi.org/10.3991/ijoe.v14i04.8513
[20] P. A. S. Mendes, M. Mendes, A. P. Coimbra, and M. M. Crisóstomo, "Movement detection and moving object distinction based on optical flow," *Lect. Notes Eng. Comput. Sci.*, vol. 2240, pp. 48–53, 2019.
[21] A. A. A. Karim and R. A. Sameer, "Static and dynamic video summarization," *Iraqi J. Sci.*, vol. 60, no. 7, pp. 1627–1638, 2019, https://doi.org/10.24996/ijs.2019.60.7.23
[22] A. A. Karim, "Construction of a robust background model for moving object detection in video sequence," *Iraqi J. Sci.*, vol. 59, no. 2, pp. 969–979, 2018, https://doi.org/10.24996/ijs.2018.59.2B.19
[23] W. H. Ali, "A new method using naive bayes and RGBD facial identification based on extracted features from image pixels," *Eng. Technol. J.*, vol. 39, no. 4A, pp. 632–641, 2021, https://doi.org/10.30684/etj.v39i4A.1936

[24] S. N. Mohammed and A. K. Abdul Hassan, "The effect of the number of key-frames on the facial emotion recognition accuracy," *Eng. Technol. J.*, vol. 39, no. 1, pp. 89–100, 2021, https://doi.org/10.30684/etj.v39i1B.1806
[25] J. Wu, Y. Min, X. Yang, and S. Ma, "Micro-expression recognition algorithm based on information entropy feature," *J. Shanghai Jiaotong Univ.*, vol. 25, no. 5, pp. 589–599, 2020, https://doi.org/10.1007/s12204-020-2219-7
[26] A. E. Ali and N. F. Hassan, "Proposing a scheme for human interactive proof test using plasma effect," *Baghdad Sci. J.*, vol. 16, no. 2, 2019, https://doi.org/10.21123/bsj.16.2.0409
[27] M. N. Abdullah and Y. H. Ali, "Vehicles detection system at different weather conditions," *Iraqi J. Sci.*, vol. 62, no. 6, pp. 2040–2052, 2021, https://doi.org/10.24996/ijs.2021.62.6.30
[28] S. Husseini, "A survey of optical flow techniques for object," p. 29, 2017.
[29] D. A. Lisin, M. A. Mattar, M. B. Blaschko, M. C. Benfield, and E. G. Learned-Miller, "Combining local and global image features for object class recognition 1 introduction 2 classification with global features", https://doi.org/10.1109/CVPR.2005.433
[30] E. Hato and M. E. Abdulmunem, "Fast Algorithm for Video Shot Boundary Detection Using SURF features," *SCCS 2019 – 2019 2nd Sci. Conf. Comput. Sci.*, pp. 81–86, 2019, https://doi.org/10.1109/SCCS.2019.8852603
[31] M. Aly, P. Welinder, M. Munich, and P. Perona, "Automatic discovery of image families: Global vs. local features computational vision lab, Caltech Pasadena, CA, USA Evolution Robotics," *Image (Rochester, N.Y.)*, pp. 777–780, 2009, https://doi.org/10.1109/ICIP.2009.5414235
[32] T. H. Obaida, A. S. Jamil, and N. F. Hassan, "Real-time face detection in digital video-based on Viola-Jones supported by convolutional neural networks," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 3, pp. 3083–3091, 2022, https://doi.org/10.11591/ijece.v12i3.pp3083-3091
[33] R. A. Lateef and A. R. Abbas, "Human activity recognition using smartwatch and smartphone: A review on methods, applications, and challenges," *Iraqi J. Sci.*, vol. 63, no. 1, pp. 363–379, 2022, https://doi.org/10.24996/ijs.2022.63.1.34
[34] Ali Ismail Awad and Mahmoud Hassaballah, "Image feature detectors and descriptors," *Stud. Comput. Intell.*, vol. 630, 2016, https://doi.org/10.1007/978-3-319-28854-3
[35] W. S. Ahmed and A. A. A. Karim, "Motion Classification Using CNN Based on Image Difference," *CITISIA 2020 – IEEE Conf. Innov. Technol. Intell. Syst. Ind. Appl. Proc.*, 2020, https://doi.org/10.1109/CITISIA50690.2020.9371835
[36] S. H. Shaker and F. Q. Al-Khalidi, "Human gender and age detection based on attributes of face," *Int. J. Interact. Mob. Technol.*, vol. 16, no. 10, pp. 176–190, 2022, https://doi.org/10.3991/ijim.v16i10.30051
[37] P. Gunawardena *et al.*, "Real-time automated video highlight generation with dual-stream hierarchical growing self-organizing maps," *J. Real-Time Image Process.*, vol. 18, no. 5, pp. 1457–1475, 2021, https://doi.org/10.1007/s11554-020-00957-0
[38] S. S. Ghintab and M. Y. Hassan, "CNN-based visual localization for autonomous vehicles under different weather conditions," *Eng. Technol. J.*, vol. 41, p. 2, 2023, https://doi.org/10.30684/etj.2022.135917.1289
[39] A. K. Farhan, N. M. G. Al-Saidi, A. T. Maolood, F. Nazarimehr, and I. Hussain, "Entropy analysis and image encryption application based on a new chaotic system crossing a cylinder," *Entropy*, vol. 21, no. 10, pp. 1–14, 2019, https://doi.org/10.3390/e21100958
[40] S. Mahmood and L. Saud, "An efficient approach for detecting and classifying moving vehicles in a video based monitoring system," *Eng. Technol. J.*, vol. 38, no. 6A, pp. 832–845, 2020, https://doi.org/10.30684/etj.v38i6A.438

## 8 Authors

**Ali Abid Hussan Altalbi,** received the BSc Computer Sciences in 2006 from the University of Babylon, and MSc degrees in Computer Sciences in 2019 from the University of Technology, Baghdad, Iraq, current Ph.D. student in Department Computer Sciences, University of Technology, Baghdad, Iraq. His research focuses on Computer Vision, Video and Image processing and pattern recognition (CS.20.02@grad.uotechnology.edu.iq).

**Dr. Shaimaa Hameed Shaker** earned his Ph.D. in Computer science from the Department of Computer Science at the University of Technology-Baghdad-Iraq since 2006. Shaimaa earned her bachelor's and master's degree also in Computer Science at the University of Technology (UOT)-Baghdad-Iraq since 1996. Her research interested focus on Image processing and pattern recognition and security visual cryptography systems (120011@uotechnology.edu.iq).

**Dr. Akbas Ezaldeen Ali** earned her Ph.D., Msg, and BSc in Computer Science from the Department of Computer Science at the University of Technology, Baghdad, Iraq. Her research interested focus on Video and image Processing, Artificial Intelligent, and Computer Vision (110034@uotechnology.edu.iq).

Article submitted 2023-02-14. Resubmitted 2023-03-11. Final acceptance 2023-03-11. Final version published as submitted by the authors.