Recognition of Suspicious Human Activity in Video Surveillance: A Review

Neha Gupta

Computer Science & Engineering Department, IFTM University, India | Computer Science & Engineering Department, Moradabad Institute of Technology, India
discoverneha@gmail.com
(corresponding author)

Bharat Bhushan Agarwal

Computer Science & Engineering Department, School of Computer Science and Applications, IFTM University, India
bharat_agarwal@iftmuniversity.ac.in

ABSTRACT

Over the past few years, there has been a noticeable growth in the use of video surveillance systems, frequently functioning as integrated systems that remotely monitor key locations. In order to prevent terrorism, theft, accidents, illegal parking, vandalism, fighting, chain snatching, and crime, human activities can be observed through visual surveillance in sensitive and public places like buses, trains, airports, banks, shopping centers, schools, and colleges. In this paper, a review of the state-of-the-art is provided, showing the overall development of identifying suspicious behavior from surveillance recordings over the past few years. We give a quick overview of the issues and difficulties associated with recognizing suspicious human activity. The purpose of this publication is to give this field's scholars a literature evaluation of several suspicious activity recognition systems along with their general structure.

Keywords-suspicious activity; video surveillance; human activity; deep learning

I. INTRODUCTION

Suspicious Human Behavior (SHA) recognition through video surveillance is a well-known topic of research in the fields of image processing and computer vision and involves classifying human activity as normal and abnormal. Unusual or suspicious behaviors, which are rarely displayed by persons in public settings, can be classed as abnormal. Today, more and more people are using video surveillance to keep an eye on human activity and stop any questionable behavior. During the recent years, human activity recognition has been realized by a wide variety of applications in military, intelligence, mass-transit agencies, and research and academia as a measure to counter crime and terrorism, public health monitoring, detection of public violent protests and attacks, etc. [1-3]. Due to the technological improvement, the means of surveillance are easily available, yet the means of continuously and effectively processing, analyzing, and detecting are not. Intelligent systems that can detect and classify suspicious human activities have been established as a crucial means to carry the necessary counter-action mechanisms, for controlling the situation, and/or for post situation/scenario analysis [4].

In general, SHA recognition involves identification of human activities that can be classified as normal and abnormal.

The first classification refers to the common human activities at public places. Normal activities are the typical human behaviors that occur in public spaces, such as hand-waving, applauding, jogging, boxing, and walking. The latter category includes behaviors like leaving bags behind, dodging crowds, robbing, fighting and attacking, vandalism, crossing borders, etc. [5]. The focus is to employ efficient techniques/systems to identify the abnormalities in human activities which can be used to predict/decide the appropriate course of action. The abnormalities can be of several kinds and be mixed with normal activities. For example, in the first category, while running, a person might drop a bag with explosives, while a walking person might point a weapon towards another person, and so on. On the other hand, in the second category the abnormalities can be classified incorrectly, e.g. a running crowd in a marathon does not need to be identified as a negative behavior.

There is a wide range of activities to be identified in SHA recognition [5], as shown in Figure 1. Explosives left in unattended objects by terrorists in crowded places or secluded places have a complex range of tasks to be looked into before making appropriate decisions. Illegal parking on roads causing traffic jam or accidents needs quick decisions in real life applications. Activities involving violence such as street

fighting, vandalisms, etc. also require quick action to ensure minimum damage of life and property. Manual monitoring and intervention fail to deal with such situations. Thus, fully automatic, effective intelligent surveillance systems that are efficient enough to handle real-time scenarios are important to develop and install. This survey provides a comprehensive detail regarding the different aspects of developing such systems, discusses existing works, and finally puts forward the prospects.

Fig. 1.   Activity recognition.

II.   SOLUTION APPROACHES AND DISCUSSION

The following approaches are used by the existing solutions for SHA recognition:

- Object tracking and detection by background subtraction: This is the most common initial phase rule followed by the existing solutions for SHA recognition. In this phase, items are tracked and identified using changes in the order of the frames, and then the foreground objects are retrieved. Tracking-based or non-tracking-based methods are used to identify objects in video frames [6].

- Feature extraction: Certain features of objects such as motion and shape are extracted using various algorithms to identify objects [7].

- Object classification: In this step, a method is used to distinguish between the various things in the video, such as people, guns, cars, etc. Various methods, including SVM, Bayesian, Haar-classifier, KNN, face recognition, and skin color detection are used.

- Suspicious activity detection: The final stage after classifying the items in the video stream is a comparison with several threshold values categorically to check for aberrant behavior.

The specifications as to what can a suspicious event be are discussed below.

- Loitering individuals: We have loitering when a person is present at a particular location for a longer period than required. The related mathematical formulations can be quantified depending on the applications type [8].

- Unattended/missing objects: The suspicious objects have to be detected in the video frames. However, identification of abandoned/unattended object is very difficult in video frames in crowded environment. To tackle the problem different techniques have been proposed, such as frame differencing, optical flow, and background subtraction [9, 10].

- Intruder detection: Intruder detection approaches utilize multiple cameras for SHA analysis in cases such as bank robberies, chain snatching, etc. Another proposed technique is based on Ontology, and is also used in airports and banks to recognize SHA. The optical flow technique is employed to detect the snatch theft from crowd movement pedestrian video footage.

- Abnormal activity/behavior: It includes fall detection, accident detection, crowd detection, fire detection, etc. Fall detection is predominantly used in the health monitoring system. Herein, static cameras are used for patient care in hospitals and unattended elderly people at home in modern times [11].

III.   LITERATURE REVIEW

In this section we discuss the relevant works regarding SHA recognition.

Author in [12] considered the current camera surveillance systems that simultaneously capture dynamic images from areas using multiple web cameras. These images are matched against the enormous amount of dynamic images, which makes the process very tedious for the observer. Meanwhile, if some of the images are missed by the observer, the system will fail. Authors in [13] proposed a DCNN model called DIATRadHARNet designed for SHA classification. The following concepts serve as the framework for the scheme: depthwise separable convolutions, channel weighting (CHW) based on importance, multiple filter sizes in the depthwise section, and application of various size kernels to the same input tensor. Authors in [14] proposed an object detection algorithm for video surveillance and real-time security camera systems. To do this, a modified version of the Kanade-Locus-Thomasi extraction algorithm was presented for object tracking. The proposed scheme detects and analyzes a small number of features instead of a large number of objects. The problem of noise creation is dealt with a Kalman filter. Authors in [15] proposed a real-time SHA recognition scheme based on a Convolutional Neural Network (CNN) and 2D pose estimation technique which is beneficial in a wide range of surveillance areas. The skeletal images of humans are extracted from the input frames of the video through 2D pose estimations. Then, these are then fed to a pre-trained CNN to categorize them into different human activities like fall or not fall, trespassing or no trespassing, etc. Finally, based on the classification, appropriate action can be taken, such as sending messages on mobile phones, triggering alarms, etc. Authors in [16] proposed an SHA recognition technique known as YOLOv3, which handles the complex problem of human detection. For this, the video is first converted into a few frames then, each frame is analyzed to recognize and detect any suspicious activity.

In [17], the authors handle the complexity of associating the skeleton inputs for recognition of both abnormal human activity and daily activity by reducing the number of features of the datasets through an RNN-based approach. Moreover, an LSTM model was used to classify suspicious activities in medical applications. In [18], a hierarchical approach was used to detect different SHAs such as fainting, loitering, unauthorized entry, etc. At first, all SHAs are defined using a semantic approach. Then, background subtraction is employed to detect the object. Then, a correlation technique is used to track the objects. Authors in [19] based their method on two features: shape moments and Histogram of Normalized Distances (HND) from the center of gravity and contour points of the object shape. For categorizing human activities, Naive Bayes classifier and Multi-class Support Vector Machine were utilized. Authors in [20] focused on detecting gun-based crimes and abandoned luggage based on a deep neural network model capable of detecting handguns in images. In [21], the authors provided an SHA detection and tracking recognition system based on artificial neural networks. The silhouette pattern of the human blob created through segmenting the camera-captured video is used in this technique. In [22], the authors combined the Bi-LSTM network with Skeleton Activity Forecasting (SAF) to create a deep learning-based SHA recognition system. The pose estimation in this case uses the human skeletal joints, which are viewed as points. On a streaming video from an IP networked camera, the skeleton tracking and places of interest are approximated.

In [23], a deep learning based approach is used for SHA recognition wherein, at first, the features are calculated from video frames. Then, a feature classifier is used to predict whether the activity is suspicious. In [24], at first the videos were broken down into frames. Then these frames were analyzed using background subtraction to detect humans. Then, after using a CNN to extract features from the frames, a Discriminative Deep Belief Network (DDBN) was used. The DDBN is also given some labeled movies of various SHAs, and its associated features are also retrieved. The two sets of features are then compared to determine if there are normal or suspicious behaviors. In [25], a 63 layer-deep CNN model called L4-BranchedActionNet is proposed. The framework is first trained using an object detection dataset called CIFAR-100 with the help of the SoftMax function. After that, the dataset for SHA recognition is then input to the trained model to obtain a set of features. These are then improved utilizing an ant colony system and coded features with an entropy-based structure. To produce the final results, a variety of SVM and KNN-based classifiers are fed with the optimized features. Authors in [26] proposed an algorithm for tracking a moving object by the video samples captured during movement. This method consists of multiple subparts including video sample creation, an experimental setup for capturing the data, and applying the object tracking algorithm using mean shift algorithms. This system can be used for calculating vehicle speed, number of vehicles passed, etc. and it has been tested over multiple frame rates. The results were not compared with the existing work.

Suspicious activity detection will be a major breakthrough in the video surveillance for behavior identification, action recognition, activity classification, etc. Authors in [27] proved the usability of automated surveillance systems. Surveillance plays an important role in maintaining law and order and in detection of possible threat. Usually, this process happens manually and requires a significant amount of manpower, which can be reduced if we automate the process. Authors in [28] explored the Situation Awareness (SAW) which is used for many crucial applications. The main challenge in SAW is faced during the instant changes while identifying objects. It also becomes challenging, whenever it focuses on huge video frames. The developed system is question-answer based, and chooses content as per interest. The interest-based traits can be more intricate in some situations than the facial features. Authors in [25] developed a CNN model having 63 deep layers. The given name of the model is L4-Branched-ActionNet. The AlexNet has been modified using four blanched sub-structures. It is first turned into a network that has already been trained using the CIFAR-100 object detection dataset and the SoftMax algorithm and the crucial details are spotted. To optimize feature subsets, these features are used. Entropy is employed to code these traits initially, and an ant colony optimization method is then applied. Different variants of SVM and KNN were used for classification. Cubic SVM provides the highest efficiency of 0.99. Several machine learning approaches as described in [29-37] can be utilized in a similar way. Authors in [38] proposed a classroom activity detection approach using video surveillance. This method's drawback is that it is difficult to differentiate between students and teacher in the class and to develop a generalized model for all scenarios. Also, data from real surveillance can be noisy and of low-quality. The proposed work was tested on real environment using Siamese neural network for classification from classroom recordings. Authors in [39] also worked to improve the classroom learning outcome. A brief comparative summary of the relevant works is shown in Table I. Similar work has been done in [40-42].

IV. MOTIVATION AND CHALLENGES

The motivation of adopting smart and intelligent video surveillance systems is mainly to identify human activities which are suspicious in nature. In various highly sensitive places these systems will aid in the prevention of thefts, terrorist acts, hooliganism, fighting, and attacks, and fire. The following areas are shielded against shady behavior via intelligent video surveillance [7]:

- Universities and other academic institutions utilize video surveillance to keep an eye on student activity to protect property from theft and vandalism. Additionally, they assist in preventing student fighting and inappropriate behavior. When exams are being given, video surveillance may be also utilized.

- Video surveillance helps to prevent theft, vandalism, fighting, disease identification, growing crowds, and explosive attacks to protect the population and public infrastructures including borders, laboratories, jails, military bases, temples, etc.

- The use of video surveillance to identify SHA is expanding in the retail sector, both for internal security in places like

warehouses and stores as well as for external security in places like parking lots. Even tiny businesses are using cameras to keep an eye on people and record video proof in case of theft or an incident.

- In every nation, the safety of travelers, runways, and aircrafts is paramount in airports, which are highly sensitive security zones. Such security-sensitive places are protected with high levels of protection thanks to a real-time system that detects SHA in video monitoring.

- In the banking industry, video monitoring is crucial for ensuring security. The use of weapons during armed robberies and assaults is prevented by the presence of cameras. Automated banking devices are frequently targeted by criminals. One way a security camera can help in the identification of fraud is by installing a device to read the magnetic information on bank cards.

- Recognizing suspicious behavior from video surveillance can assist in finding frauds, heists, and other crimes. Intelligent video surveillance is an intriguing technique to assist security officers because monitoring a casino necessitates seeing human movement in a congested setting.

- In hospitals, video surveillance is also used to keep an eye on patients. At home, it can be used to keep an eye on elderlies or children.

TABLE I.   SHA LITERATURE REVIEW SUMMARY

Ref	Object detection method	Noise removal method	Remarks
[12]	Background subtraction	Smoothing filter	Finds the detection point of SHA and evalueates the corresponding degree of risk
[13]	Lighweight deep CNN	Spatial filter	Efficient classification with high accuracy
[14]	KLT and Kalman filter		Efficient real time tracking
[15]	2D pose estimation and CNN		Effective response system
[16]	YOLOv3		Quick processing, accurate detection
[17]	Multilayer LSTM network		Less training data, better cross-view and cross-subject evaluation
[18]	Background subtraction	A thresholding technique	Less complexity, high accuracy
[19]	Naïve Bayes classifier		Effective and accurate detection and prediction
[20]	Deep neural network		Gun-crime andabandoned laggage detection
[21]	ANN		High accuracy and robustness
[22]	Bi-LSTM network	Adaptive thresholding	Skeleton tracking
[23]	CNN and RNN		Apriori detection, simple, yet powerful
[24]	CNN and DDBN		High accuracy
[25]	Entropy-coded ant colony system		High accuracy
[26]	Mean shift algorithm		Efficient object tracking. Box and image sequence segmentation
[27]	Faster region-based CNN inception V2 framework		SHA detection in public places
[28]	I-ViSE and deep neural networks	Smoothing filter	Emphasis is given on ensuring that the photos analyzed have significant content and good quality
[29]	Siamese neural framework	Textual windows across segment	Superior prediction perfomance for online and offline classrooms
[30]	Deep fully connected convonutional and recurrent neural network		Precice estimates of the total amount of time spent in classes or activities

To develop an intelligent video surveillance system that can automatically detect SHA, there are several issues and challenges to overcome [7]:

- Moving object recognition is challenging to accurately process due to dynamic fluctuation in natural environments, such as slow illumination changes brought on by day-to-night shifts and quick illumination differences brought on by weather changes.

- The look of an object is altered by its shadow, making it difficult to track and identify the specific object in a video. Characteristics like shape, movement, and background, are more delicate for a shadow.

- Occasionally, noises made by swaying tree branches make it difficult to identify an object in a video.

- Finding the object in a busy location is a difficult process. It is quite challenging to discover abandoned objects, theft, or violence in such a circumstance.

- Full or partial object occlusions occur occasionally, the objects in video are entirely or partially obscured, making identification difficult.

- It might be quite difficult to identify foreground items in low-resolution films. Identification of object boundaries becomes extremely challenging.

- The creation of a real-time, intelligent monitoring system is the more difficult task. When extracting and tracking the foreground objects from films with complicated backgrounds, processing time increases.

- Since background subtraction in abandoned object detection only recognizes moving things as the foreground, static object detection is a difficult task.

V. MEASURES TO DETECT SHA

In general, the measures used to detect SHA follow a hierarchical structure as shown in Figure 2. At first the video is split into a sequence of frames. Then, background subtraction is used to detect changes in the sequences of the frame and then extraction of foreground objects. Following this, the different types of objects are attempted to be extracted by the system. Then, the different obstructions in the form of noise, shadow etc. are removed from the frame containing the object. Finally, the recognizable object is obtained from the surveillance

footage. Figure 3 represents the specific steps regarding the general steps in entity (object) detection in SHA recognition. The first step is the identification of the different kinds of steps. To stop terrorist bomb attacks, suspicious activity detection also includes the detection of abandoned objects. Background approaches in video surveillance treat stationary items as the background and moving things as the foreground. Therefore, a newly arrived object is absorbed into the backdrop when it becomes static.
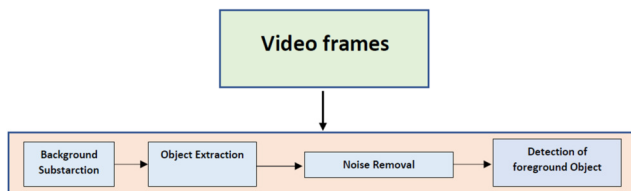

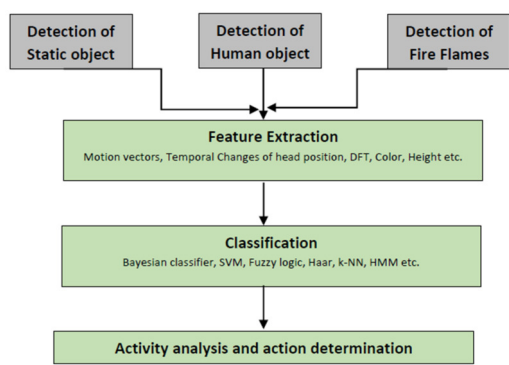Fig. 2.   General steps in entity detection.

Fig. 3.   Specifications regarding the general steps in entity (object) detection in SHA recognition.

Finding foreground objects that are free from noise, illumination, or shadow is incredibly challenging in computer vision. Noise makes it difficult to identify an object, illumination produces false detection, and shadow alters the object's appearance, making object tracking particularly challenging. The right features must be chosen in order for video surveillance to automatically detect anomalous activity. Finding the most valuable information in the captured video is the main goal of feature extraction. After determining whether there are foreground-moving or still objects in a frame, the object categorization stage is used to determine whether the behavior is normal or abnormal. A static human is distinguished from a static abandoned object, a fight is distinguished from a boxing match, a face is distinguished from an object that has skin color, fire is distinguished from a flashlight, the sun, or any other artificial light source, and so on.

VI.   CONCLUSION AND FUTURE WORK

The current paper reviews various intelligent and automatic SHA recognition surveillance techniques. The reviewed papers cover a wide range of applications associated with real-time implementation. The applications range from health-care systems to perimeter monitoring in war fields. The techniques and dataset used are the most common ones pertaining to the SHA domain. Note that recognition of human behavior is a complex process, and monitoring and analyzing human activities is difficult. Despite the complex nature of the problem, it is also required in day-to-day life in order to improve safety in modern society. On the other hand, recognition of suspicious activity of non-living things is relatively less complex but requires a great deal of expertise to deduce correct results. We discussed the related works and mentioned their shortcomings. The weaknesses can be significantly improved and pave the way for a wide range of open research.

REFERENECES

[1]   J. Candamo, M. Shreve, D. B. Goldgof, D. B. Sapper, and R. Kasturi, "Understanding Transit Scenes: A Survey on Human Behavior-Recognition Algorithms," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 1, pp. 206–224, Mar. 2010, https://doi.org/10.1109/TITS.2009.2030963.

[2]   S. A. Shah and F. Fioranelli, "RF Sensing Technologies for Assisted Daily Living in Healthcare: A Comprehensive Review," *IEEE Aerospace and Electronic Systems Magazine*, vol. 34, no. 11, pp. 26–44, Aug. 2019, https://doi.org/10.1109/MAES.2019.2933971.

[3]   O. D. Lara and M. A. Labrador, "A Survey on Human Activity Recognition using Wearable Sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013, https://doi.org/10.1109/SURV.2012.110112.00192.

[4]   W. Huang, L. Zhang, W. Gao, F. Min, and J. He, "Shallow Convolutional Neural Networks for Human Activity Recognition Using Wearable Sensors," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021, https://doi.org/10.1109/TIM.2021.3091990.

[5]   R. K. Tripathi, A. S. Jalal, and S. C. Agrawal, "Suspicious human activity recognition: a review," *Artificial Intelligence Review*, vol. 50, no. 2, pp. 283–339, Aug. 2018, https://doi.org/10.1007/s10462-017-9545-7.

[6]   J. M. McHugh, J. Konrad, V. Saligrama, and P.-M. Jodoin, "Foreground-Adaptive Background Subtraction," *IEEE Signal Processing Letters*, vol. 16, no. 5, pp. 390–393, Feb. 2009, https://doi.org/10.1109/LSP.2009.2016447.

[7]   A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Computing Surveys*, vol. 38, no. 4, Sep. 2006, Art. no. 13, https://doi.org/10.1145/1177352.1177355.

[8]   S. Patil and K. Talele, "Suspicious movement detection and tracking based on color histogram," in *International Conference on Communication, Information & Computing Technology*, Mumbai, India, Jan. 2015, pp. 1–6, https://doi.org/10.1109/ICCICT.2015.7045698.

[9]   T. Y. Lai, J. Y. Kuo, C.-H. Liu, Y. W. Wu, Y.-Y. Fanjiang, and S.-P. Ma, "Intelligent Detection of Missing and Unattended Objects in Complex Scene of Surveillance Videos," in *International Symposium on Computer, Consumer and Control*, Taichung, Taiwan, Jun. 2012, pp. 662–665, https://doi.org/10.1109/IS3C.2012.172.

[10]   T. T. Zin, P. Tin, H. Hama, and T. Toriu, "Unattended object intelligent analyzer for consumer video surveillance," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 2, pp. 549–557, Feb. 2011, https://doi.org/10.1109/TCE.2011.5955191.

[11]   K. K. Verma, B. M. Singh, and A. Dixit, "A review of supervised and unsupervised machine learning techniques for suspicious behavior recognition in intelligent surveillance system," *International Journal of Information Technology*, vol. 14, no. 1, pp. 397–410, Feb. 2022, https://doi.org/10.1007/s41870-019-00364-0.

[12]   M. Takai, "Detection of suspicious activity and estimate of risk from human behavior shot by surveillance camera," in *Second World Congress on Nature and Biologically Inspired Computing*, Kitakyushu, Japan, Dec. 2010, pp. 298–304, https://doi.org/10.1109/NABIC.2010.5716350.

[13] M. Chakraborty, H. C. Kumawat, S. V. Dhavale, and A. B. Raj A., "DIAT-RadHARNet: A Lightweight DCNN for Radar Based Classification of Human Suspicious Activities," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–10, 2022, https://doi.org/10.1109/TIM.2022.3154832.

[14] S. Nandyal and S. Angadi, "Recognition of Suspicious Human Activities using KLT and Kalman Filter for ATM Surveillance System," in *International Conference on Innovative Practices in Technology and Management*, Noida, India, Feb. 2021, pp. 174–179, https://doi.org/10.1109/ICIPTM52218.2021.9388322.

[15] A. S. Dileep, S. S. Nabilah, S. Sreeju, K. Farhana, and S. Surumy, "Suspicious Human Activity Recognition using 2D Pose Estimation and Convolutional Neural Network," in *International Conference on Wireless Communications Signal Processing and Networking*, Chennai, India, Mar. 2022, pp. 19–23, https://doi.org/10.1109/WiSPNET54241.2022.9767152.

[16] N. Bordoloi, A. K. Talukdar, and K. K. Sarma, "Suspicious Activity Detection from Videos using YOLOv3," in *17th India Council International Conference*, New Delhi, India, Dec. 2020, pp. 1–5, https://doi.org/10.1109/INDICON49873.2020.9342230.

[17] R. Nale, M. Sawarbandhe, N. Chegogoju, and V. Satpute, "Suspicious Human Activity Detection Using Pose Estimation and LSTM," in *International Symposium of Asian Control Association on Intelligent Robotics and Industrial Automation*, Goa, India, Sep. 2021, pp. 197–202, https://doi.org/10.1109/IRIA53009.2021.9588719.

[18] U. M. Kamthe and C. G. Patil, "Suspicious Activity Recognition in Video Surveillance System," in *Fourth International Conference on Computing Communication Control and Automation*, Pune, India, Aug. 2018, pp. 1–6, https://doi.org/10.1109/ICCUBEA.2018.8697408.

[19] H. Samir, H. E. Abd El Munim, and G. Aly, "Suspicious Human Activity Recognition using Statistical Features," in *13th International Conference on Computer Engineering and Systems*, Cairo, Egypt, Dec. 2018, pp. 589–594, https://doi.org/10.1109/ICCES.2018.8639457.

[20] S. Loganathan, G. Kariyawasam, and P. Sumathipala, "Suspicious Activity Detection in Surveillance Footage," in *International Conference on Electrical and Computing Technologies and Applications*, Ras Al Khaimah, United Arab Emirates, Nov. 2019, pp. 1–4, https://doi.org/10.1109/ICECTA48151.2019.8959600.

[21] M. K. Fiaz and B. Ijaz, "Vision based human activity tracking using artificial neural networks," in *International Conference on Intelligent and Advanced Systems*, Kuala Lumpur, Malaysia, Jun. 2010, pp. 1–5, https://doi.org/10.1109/ICIAS.2010.5716186.

[22] D. Kumar and S. R. Sailaja, "Abnormal Activity Recognition using Deep Learning in Streaming Video for Indoor Application," in *ITU Kaleidoscope: Connecting Physical and Virtual Worlds*, Geneva, Switzerland, Dec. 2021, pp. 1–7, https://doi.org/10.23919/ITUK53220.2021.9662095.

[23] C. V. Amrutha, C. Jyotsna, and J. Amudha, "Deep Learning Approach for Suspicious Activity Detection from Surveillance Video," in *2nd International Conference on Innovative Mechanisms for Industry Applications*, Bangalore, India, Mar. 2020, pp. 335–339, https://doi.org/10.1109/ICIMIA48430.2020.9074920.

[24] B. A. Alavudeen, P. Parthasarathy, and S. Vivekanandan, "Detection of Suspicious Human Activity based on CNN-DBNN Algorithm for Video Surveillance Applications," in *Innovations in Power and Advanced Computing Technologies*, Vellore, India, Mar. 2019, vol. 1, pp. 1–7, https://doi.org/10.1109/i-PACT44901.2019.8960085.

[25] T. Saba, A. Rehman, R. Latif, S. M. Fati, M. Raza, and M. Sharif, "Suspicious Activity Recognition Using Proposed Deep L4-Branched-Actionnet With Entropy Coded Ant Colony System Optimization," *IEEE Access*, vol. 9, pp. 89181–89197, 2021, https://doi.org/10.1109/ACCESS.2021.3091081.

[26] G. Mathur, D. Somwanshi, and M. M. Bundele, "Intelligent Video Surveillance based on Object Tracking," in *3rd International Conference and Workshops on Recent Advances and Innovations in Engineering*, Jaipur, India, Nov. 2018, pp. 1–6, https://doi.org/10.1109/ICRAIE.2018.8710421.

[27] R. Srinath, J. Vrindavanam, V. P. Vasudev, S. Supreeth, H. Raj, and A. Kesarwani, "A Machine Learning Approach for Localization of

Suspicious Objects using Multiple Cameras," in *IEEE International Conference for Innovation in Technology*, Bangluru, India, Nov. 2020, pp. 1–6, https://doi.org/10.1109/INOCON50539.2020.9298364.

[28] S. Yahya Nikouei, Y. Chen, A. Aved, and E. Blasch, "I-ViSE: Interactive Video Surveillance as an Edge Service using Unsupervised Feature Queries," Mar. 2020. https://doi.org/10.48550/arXiv.2003.04169.

[29] K. R. Kodepogu *et al.*, "A Novel Deep Convolutional Neural Network for Diagnosis of Skin Disease," *Traitement du Signal*, vol. 39, no. 5, pp. 1873–1877, Nov. 2022, https://doi.org/10.18280/ts.390548.

[30] N. Kumar and D. Aggarwal, "LEARNING-based Focused WEB Crawler," *IETE Journal of Research*, pp. 1–9, Feb. 2021, https://doi.org/10.1080/03772063.2021.1885312.

[31] M. Kaur, V. Kumar, V. Yadav, D. Singh, N. Kumar, and N. N. Das, "Metaheuristic-based Deep COVID-19 Screening Model from Chest X-Ray Images," *Journal of Healthcare Engineering*, vol. 2021, Mar. 2021, Art. no. e8829829, https://doi.org/10.1155/2021/8829829.

[32] N. Kumar, N. Narayan Das, D. Gupta, K. Gupta, and J. Bindra, "Efficient Automated Disease Diagnosis Using Machine Learning Models," *Journal of Healthcare Engineering*, vol. 2021, May 2021, Art. no. e9983652, https://doi.org/10.1155/2021/9983652.

[33] N. Kumar, M. Gupta, D. Gupta, and S. Tiwari, "Novel deep transfer learning model for COVID-19 patient detection using X-ray chest images," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 1, pp. 469–478, Jan. 2023, https://doi.org/10.1007/s12652-021-03306-6.

[34] M. Gupta, N. Kumar, B. K. Singh, and N. Gupta, "NSGA-III-Based Deep-Learning Model for Biomedical Search Engines," *Mathematical Problems in Engineering*, vol. 2021, May 2021, Art. no. e9935862, https://doi.org/10.1155/2021/9935862.

[35] N. Kumar, M. Gupta, D. Sharma, and I. Ofori, "Technical Job Recommendation System Using APIs and Web Crawling," *Computational Intelligence and Neuroscience*, vol. 2022, Jun. 2022, Art. no. e7797548, https://doi.org/10.1155/2022/7797548.

[36] M. Gupta, N. Kumar, N. Gupta, and A. Zaguia, "Fusion of multi-modality biomedical images using deep neural networks," *Soft Computing*, vol. 26, no. 16, pp. 8025–8036, Aug. 2022, https://doi.org/10.1007/s00500-022-07047-2.

[37] A. Hashmi *et al.*, "Contrast Enhancement in Mammograms Using Convolution Neural Networks for Edge Computing Systems," *Scientific Programming*, vol. 2022, Apr. 2022, Art. no. e1882464, https://doi.org/10.1155/2022/1882464.

[38] H. Li, Z. Wang, J. Tang, W. Ding, and Z. Liu, "Siamese Neural Networks for Class Activity Detection," in *21st International Conference on Artificial Intelligence in Education*, Ifrane, Morocco, Jul. 2020, pp. 162–167, https://doi.org/10.1007/978-3-030-52240-7_30.

[39] E. Slyman, C. Daw, M. Skrabut, A. Usenko, and B. Hutchinson, "Fine-Grained Classroom Activity Detection from Audio with Neural Networks." arXiv, Nov. 09, 2021, https://doi.org/10.48550/arXiv.2107.14369.

[40] Y. Said, M. Barr, and H. E. Ahmed, "Design of a Face Recognition System based on Convolutional Neural Network (CNN)," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5608–5612, Jun. 2020, https://doi.org/10.48084/etasr.3490.

[41] B. A. Mossaad, S. Elkosantini, and M. Abid, "An Automated Surveillance System Based on Multi-Processor and GPU Architecture," *Engineering, Technology & Applied Science Research*, vol. 7, pp. 2319–2323, Dec. 2017, https://doi.org/10.48084/etasr.1645.

[42] N. Kumar, A. Hashmi, M. Gupta, and A. Kundu, "Automatic Diagnosis of Covid-19 Related Pneumonia from CXR and CT-Scan Images," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 7993–7997, Feb. 2022, https://doi.org/10.48084/etasr.4613.