

Review

Taxonomy of Anomaly Detection Techniques in Crowd Scenes

Amnah Aldayri and Waleed Albattah * 

Department of Information Technology, College of Computer, Qassim University, Buraydah 52571, Saudi Arabia
* Correspondence: w.albattah@qu.edu.sa

Abstract: With the widespread use of closed-circuit television (CCTV) surveillance systems in public areas, crowd anomaly detection has become an increasingly critical aspect of the intelligent video surveillance system. It requires workforce and continuous attention to decide on the captured event, which is hard to perform by individuals. The available literature on human action detection includes various approaches to detect abnormal crowd behavior, which is articulated as an outlier detection problem. This paper presents a detailed review of the recent development of anomaly detection methods from the perspectives of computer vision on different available datasets. A new taxonomic organization of existing works in crowd analysis and anomaly detection has been introduced. A summarization of existing reviews and datasets related to anomaly detection has been listed. It covers an overview of different crowd concepts, including mass gathering events analysis and challenges, types of anomalies, and surveillance systems. Additionally, research trends and future work prospects have been analyzed.

Keywords: crowd; anomaly detection; abnormal behavior; surveillance system; CCTV



Citation: Aldayri, A.; Albattah, W. Taxonomy of Anomaly Detection Techniques in Crowd Scenes. *Sensors* **2022**, *22*, 6080. <https://doi.org/10.3390/s22166080>

Academic Editor: Mario Munoz-Organero

Received: 4 July 2022

Accepted: 9 August 2022

Published: 14 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The World Health Organization (WHO) clarifies significant gathering events as any occurrence, whether planned or unplanned, that attracts a substantial number of participants to strain the neighborhood, city, or nation hosting the event's planning and response resources [1]. The heterogeneous composition of the crowd in terms of color, age, language, and culture presents several administrative issues for local organizers focused on ensuring the event's efficient management. Administrative authorities are more concerned with understanding the crowd mechanics that explain what could harm large crowds [2]. An anomaly detection system is a monitoring program that automatically identifies and considers the signs of abnormal or irregular actions directly [3]. With the widespread usage of video surveillance techniques, manual evaluation of vast quantities of video data gathered from crowd surveillance CCTV cameras has become complicated, time-consuming, and ineffective in the case of large crowds [4]. It requires workforce and continuous attention to decide if the captured actions are normal or abnormal. Therefore, an automatic anomaly detection functionality is necessary for surveillance systems to accurately identify and detect anomalies in crowd scenes [5]. Detecting abnormal behaviors rapidly and automatically in crowded environments is significant for improving safety, preventing risks, and guaranteeing quick response. Anomaly detection in surveillance systems is critical for assuring safety, security, and in some cases, the prevention of possible disasters [6]. Anomaly detection intends to discover the anomalies in a quick time automatically. Recently, intelligent monitoring systems have become crucial for effective crowd management. Due to their importance, computer vision, video analysis, and automated crowd anomaly detection have become popular research topics.

Contribution

A comprehensive overview of the crowd concept, abnormal human behavior, and surveillance systems have been discussed. A new taxonomic organization of the recent

developments in abnormal human behavior detection techniques for large-scale (Danse) crowds has been proposed to identify subfields that are still unexplored or that are seldom approached from the perspective of deep learning. A wide range of recent deep learning approaches to detecting anomalies has been covered. It includes research papers and reviews published in the time interval from 2011 to 2022. Moreover, this review focuses on studying the human crowd, specifically human abnormal behavior.

2. Crowd and Mass Gathering Event

A large-scale event is crowded and attracts people from multiple locations with diverse cultural backgrounds, which generates significant management, control, and communication challenges due to their diversity [7]. These large gatherings are potentially dangerous for the public. Numerous physical characteristics describe crowd behavior, including the direction of motion, velocity, energy, and interaction force [8]. The field of crowd analysis includes three general concepts or levels: crowd management, crowd monitoring, and crowd control, see Figure 1. Crowd management is defined as using techniques to plan and manage mass gathering events before, during, and after the event. It ensures the safety of people, good event planning and managing, predicts and prevents unexpected issues, and prepares initial plans for emergencies. Crowd monitoring provides the opportunity to estimate crowd dynamics, detect and predict possible risks, track, support virtual simulation of crowd behavior, and develop automated systems [9]. Globally, security and event management agencies are beginning to realize the importance of crowd monitoring, considering the growing concern about public safety. Crowd control is a public security practice and actions taken during the situation to prevent abnormal behavior when such as fights, riots, or crowd crushes occur. An automated crowd scene analysis involves counting, tracking, and identifying the behaviors of a large crowd of individuals occupying the same physical space [10]. An estimate of the number of people in a certain area is known as crowd counting.

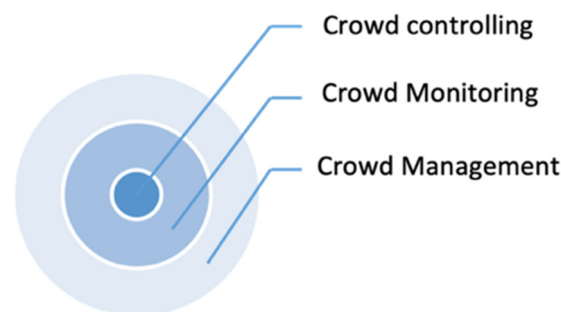


Figure 1. Crowd analysis concepts.

A critical aspect of crowd safety is crowd action recognition, which recognizes the different actions of an individual or group of individuals. The ability to track objects in crowded video sequences is critical to interpreting visual scenes [11]. In Islam, Hajj is considered one of the five pillars and a duty that all physically able, healthy, and financially capable people must fulfill once in a lifetime. Pilgrims worldwide arrive in one place over five days to perform religious rites. In the Muslim lunar calendar, this begins on day 8 of Zulhijjah and ends on day 13, Zulhijjah [12]. This annual mass gathering event is considered one of the biggest in the world. Since the number of pilgrims attending these events has increased over the years, many challenges could occur, such as overcrowding at the sites resulting in congestion, stampedes, damages, loss of pilgrims, violations, and fatalities [1]. The Kingdom of Saudi Arabia seeks to provide pilgrims with the best possible Hajj experience by providing infrastructure, safety, security mechanisms, and numerous other amenities to manage these crowds. However, it is still seeking more tools for this task [13].

3. Crowd Analysis

Most research works divide crowd analysis into two major branches: crowd statistics and crowd behavior analysis. This study introduces another branch, the tracking approach. Figure 2 illustrates the newly proposed taxonomy for crowd analysis.

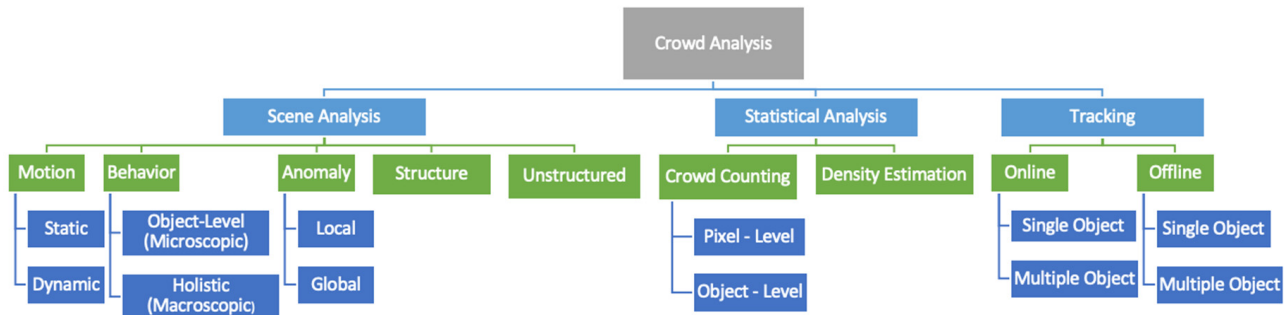


Figure 2. Proposed taxonomy for crowd analysis.

- **Scene Analysis:** Automatic video analysis is called video analytics, and it can detect and analyze temporal and spatial events. The usefulness in finding anomalies in real-time, monitoring crowds, detecting pedestrians, and tracking vehicles make video scene analysis an active research topic. The CCTVs distributed in crowded public areas facilitate the process of analyzing the motion, behavior understanding, anomaly detection, and determining the type of the crowd, whether it is structured or unstructured.
- **Statistical Analysis:** Crowd density estimation and crowd counting are examples of statistical analysis, which involves analyzing patterns and trends in quantitative data. The number of people per meter can be used to calculate crowd density. While crowd counting is a method of counting how many people are present in a space. These estimations are effective in controlling the flow of the crowd in a specific area and avoiding overcrowding, accidents, and stampedes.
- **Tracking:** Object tracking is the process of determining the location of moving objects over time [14]. An object can be tracked online or offline, and one object or several objects can be tracked simultaneously. The changes in features over time can be used to track anomalies detected by object detection.

4. Crowd Scene Analysis Challenges

Occasionally, some moving entities in videos do not appear clearly to the observer in some circumstances. There are diverse kinds of challenges, as illustrated in Figure 3:

- **Occlusion:** this happens when two or more objects come too close jointly and seem to merge, which leads to the system losing track of the trackable object or tracking the wrong object because of overlapping [15].
- **Scale Variation:** it occurs when there is a wide range of sizes of the tracked objects, which causes the tracking system to lose precise tracking.
- **Illumination Variation:** refers to the variation in the quantity of origin light mirrored on an image and can be caused by changes in lighting, shadows, or noise.
- **Speed:** while objects in a scene often move at different speeds, the predictor should recognize objects in motion videos accurately by being fast during prediction.
- **Background Clutter:** it refers to the existence of large numbers of objects in the image, which makes it difficult for a detector to recognize individual objects due to their non-uniform arrangement. There is a possibility that objects that need identifying will blend into the background, making them difficult to detect.
- **Variety:** occurs when an object has more than one shape and size.
- **Camera Position and Angle:** it is possible to have inconsistencies in perspective due to different angles and camera positions, as well as the tilting and up-and-down motion of the camera.

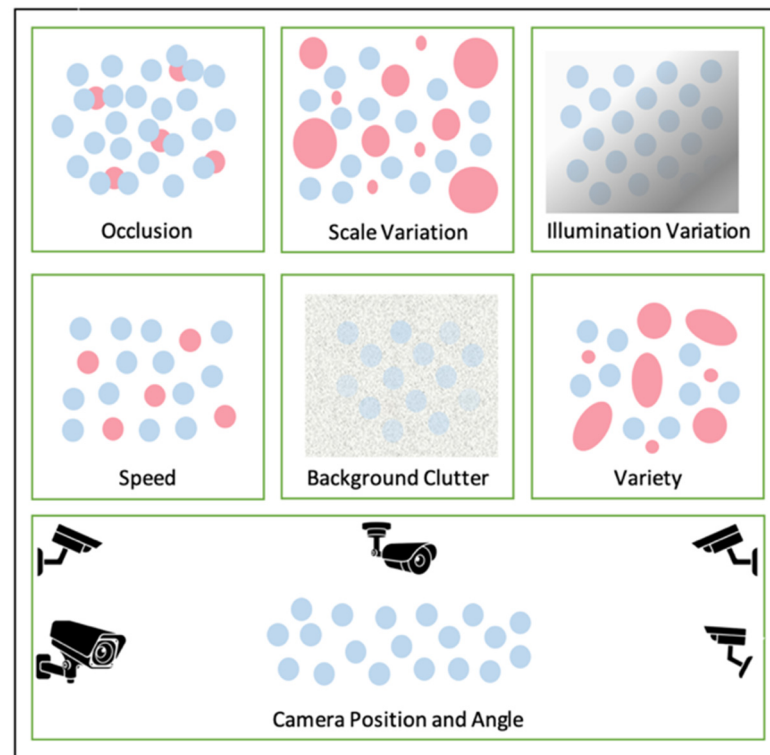


Figure 3. Crowd scene analysis challenges.

5. The Concept of Anomaly

The word anomaly comes from the Greek word “anomolia”, which indicates irregular or uneven patterns [16]. In the data mining and statistics communities, anomalies are also known as abnormalities, deviants, and outliers [17]. It can be defined as an unusual pattern that does not conform to expected behavior or place. For textual data, the anomaly can be detected by plotting the data; the data points that are greater than or less than other data are referred to as anomalies or outliers, which is inconsistent with other data. On the other hand, for videos or image data, the anomaly can be identified by analyzing and understanding the behavior or patterns of objects in that area; the object that behaves differently from expected patterns is an anomaly object. Anomaly detection refers to the detection and localization of patterns or any behavior that does not correspond to expectations. A person may exhibit abnormal behavior in public alone or as part of a group. Thousands of pilgrims gather simultaneously in the Hajj area, which is an illustration of a heavily populated place. Different abnormal activities could happen, such as congestion, walking against the pedestrian path, standing in places not designated for standing and obstructing the movement of pilgrims, sitting in places other than those designated for sitting, and running and scrambling at the gates and the train station. Moreover, violence is a representation of abnormal behavior, which is a physical force that affects the surrounding area and people; it can be detected through a smart surveillance system that helps to control the safety of the environment and limit violations and other accidents [18].

6. Anomaly Detection

Monitoring public security often involves the detection of abnormal behavior in surveillance videos of crowds. Anomaly detection in crowded scenes refers to the detection of irregularities, abnormalities, or discovering patterns that are out of alignment with normal behavior in images or video sequence data. In [19], anomaly detection is described as identifying patterns that are extremely distinct from the rest. According to [20], anomaly detection refers to the identification of crowd movements, where the abnormal behaviors in crowded locations usually emerge as crowd commotion. The detection of anomalies

aims to identify and categorize anomalies in given datasets [21]. Anomaly detection can be classified into three categories: supervised, unsupervised, and semi-supervised. The dataset that has both data and labels can be used for supervised anomaly detection. The labels determine the type of event, whether it is “normal” or “abnormal.” Unlabeled datasets are employed for the unsupervised anomaly detection method. The unsupervised method considers that most of the events in the dataset are normal and otherwise assumed an anomaly. In situations where the dataset has not been completely labeled nor unlabeled, meaning that some data are labeled, and some are unlabeled, semi-supervised anomaly detection techniques are used. As a rule, anomaly detection procedures assess the patterns in the available normal data, illustrate them, and then model them in order to uncover new patterns in the new data [22]. Surveillance systems [23–27], intrusion detection [28–30], fraud detection [31,32], and health monitoring [33–35] are just a few of the domains where anomaly detection has applications.

7. Types of Anomalies

Basically, the term anomaly refers to anything that is unusual, irregular, or uncharacteristic and differs from the normal event [36]. An anomaly can be broken down into three types: point anomalies, extended anomalies, and collective anomalies.

- **Point Anomalies:** occur when a single individual entity has observed irregular behavior from the rest of the data [37].
- **Contextual Anomaly:** An instance that could be considered anomalous in some specific circumstances is called a contextual anomaly, which is also called a conditional anomaly [36]. When a data value has irregular behavior compared to the rest of the data in a specific context, but not in all circumstances [38], therefore, if something is anomalous in some specific context, then it can be classified as a contextual anomaly.
- **Collective Anomalies:** often represent a collection of related entities as a correlated group that has observed anomalies against the remaining data. They are called collective anomalies [39].

8. Surveillance System

The surveillance system is a real-time administration program developed to identify and detect irregular activities directly automatically; it can be called an anomaly detection system [3]. Using advanced technology to manage crowds could be the proper approach to avoid any potential issues [40]. CCTV cameras are the most common safeguard instrument used to observe individuals and their activities. It is a typical policy to ensure safety. Approximately 770 million CCTV cameras have been installed around the world [41]. The constant observation of these cameras by humans is very difficult in a large crowd. The limitations of CCTV raised the requirements of continuous manual monitoring of the screens by the workers, which is very difficult to respond immediately to any actions and time-consuming. To overcome this limitation, an automatic system that could detect and identify abnormal behavior automatically and notify the authority to act at the same time is required. To detect the actions and categorize them effectively, deep learning techniques such as CNN, RNN, LSTM, and more are used, which produce outstanding results. This intelligent system is qualified to detect objects that differ significantly from the normal state, such as fighting, vagrancy, stampede, and incidents [42]. It is equipped in various areas, such as academies, roads, playgrounds, and hospitals, to encourage the management process [18].

9. Previous Reviews on Anomaly Detection

Several surveys have been published on crowd analysis and abnormalities detection. Some of the studies focus on general ideas and concepts, and some other research focuses on a specific area. As summarized in Table 1, this section presents some of the most significant reviews published between 2011 and 2022, which represent significant results and contributions. The main contribution of this review is the depth of concentration on

the deep learning methods, role, and performance in human abnormalities detection in crowded areas. For dense crowds, a new taxonomic organization presents recent developments in human abnormal behavior detection. Furthermore, participate in the discovery of subfields that are still unexplored or that are rarely covered. Future directions and trends are demonstrated as open challenges for future research. This review focuses on studying the human crowd, specifically abnormal human behavior.

Table 1. Summarized presentation of review papers in anomaly detection.

Ref.	Year	Focus
[43]	2011	Computer vision techniques for analysis of urban traffic
[44]	2012	Anomaly detection in automated surveillance systems
[45]	2012	Detecting abnormal human behavior in the context of a video
[46]	2012	Discuss frameworks for recognizing human activity
[47]	2012	Human behavior analysis with semantic enhancement
[48]	2013	Intelligence video surveillance system (IVSS) using a multi-camera network
[49]	2014	Machine learning techniques for novelty detection
[50]	2015	Describe the difficulties that come with modeling for video anomaly detection
[51]	2016	Currently available anomaly detection video datasets issues
[52]	2017	Computer vision techniques used for crowd disaster avoidance
[53]	2017	Computer vision techniques for analyzing dense crowd scenes
[54]	2017	Explore various available methods used to identify abnormal crowd behavior
[55]	2017	Crowd statistics and behavior understanding
[56]	2018	Implementation of deep learning techniques for video anomalous detection
[57]	2018	Available methods for human abnormal behavior detection
[58]	2018	Unsupervised- and semi-supervised learning-based for video anomaly detection
[59]	2018	Feature extraction and description techniques for abnormal behavior recognition
[17]	2019	Deep-learning-based anomaly detection techniques for various domains
[60]	2019	Object trajectories, clustering, anomaly detection, summarization, and synopsis generation
[61]	2020	Video anomaly detection in road traffic
[62]	2020	Deep learning-based methods for analyzing crowded scenes
[63]	2021	Deep learning technique used for anomaly detection
[64]	2021	State-of-the-art deep learning-based approaches for detecting video abnormalities
[2]	2021	Explore various studies related to crowd analysis
[42]	2021	Deep learning-based algorithms for recognizing video anomalies, opportunities, and challenges
[65]	2021	For security systems, automated and real-time surveillance technologies of irregular action recognition are used to identify dynamic crowd behavior
[66]	2021	Analyzed and compared crowd anomaly detection methodologies
[67]	2022	Crowd count, human detection and behavior, anomaly detection, and importance of crowd analysis
[68]	2022	Crowd modeling and analysis
[69]	2022	Comparative analysis of existing crowd behavior analysis methods
[70]	2022	Deep learning framework for anomaly detection
[71]	2022	GAN-based anomaly detection
[72]	2022	Summarization of video analytics deep learning techniques in the Hajj scenes
[73]	2022	Evolution of anomaly detection methodologies in intelligent video surveillance

10. Taxonomy of Anomaly Detection in Crowd Scenes

With the increasing demand for security and safety of people in large-scale crowd areas, CCTV is used to monitor the crowds. Analyzing the video streams provided by CCTV is an important task to detect and localize the anomaly behavior in the crowd. However, the literature includes many studies that need to be classified to understand deep learning for efficient crowd management better. This section presents a classification of the previous studies according to classical ML vs. DL, anomaly type, the scope of application, real-time vs. offline, and human crowd vs. non-human crowd, as described in Figure 4.

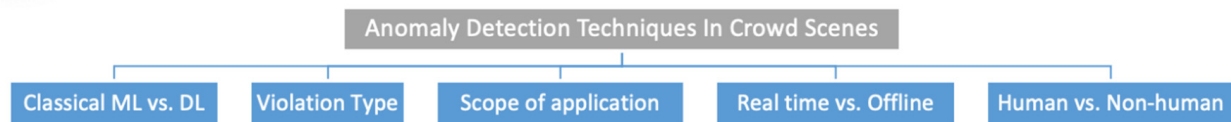


Figure 4. Anomaly detection techniques in crowd scenes.

10.1. Classical ML vs. DL

Classical ML is a collection of algorithms and techniques used to build a model that can learn from existing observations and exploit the learned data to predict new observations by finding patterns in data; it works better with small data [74–82]. DL is a subset of ML, which is based mostly on artificial neural networks [83–93]. When classical ML techniques are compared with DL techniques, the DL techniques achieve more significant performance and accuracy in many domains such as natural language, object detection, speech, image classification, and semantic segmentation [94]. ML acts satisfactorily on small to medium datasets, while DL acts sufficiently on large datasets. According to hardware requirements, ML can work on a local CPU, while DL requires powerful computing power such as a GPU. For feature engineering, ML needs to be explicitly identified or annotated features by humans, while DL can learn and discover features automatically by neural networks. For the training time, ML models usually take a short training time, while DL requires computationally intensive time and power for training. ML utilizes many automated algorithms that allow the model to generate predictions from employed data. While DL uses a quite different and complicated architecture called a neural network, that hands data via processing layers to interpret data and generate predictions. Current works on crowd anomaly detection can generally be classified into unsupervised, supervised, semi-supervised, and reinforcement learning. The dataset with both data and labels can be used for supervised anomaly detection. In anomaly detection, the labels indicate the type of event, whether it is “normal” or “abnormal”, or determine a specific type of anomaly such as fighting, burglary, and more. Unlabeled datasets have been employed for the unsupervised anomaly detection method. The unsupervised method considers that most of the events in the dataset are normal and otherwise assumed an anomaly. Semi-supervised anomaly detection techniques are used in situations where the dataset has not been completely labeled nor unlabeled, meaning that some data are labeled and some are unlabeled. The Supervision type column determines the type of the method: unsupervised [74–77,85,86], supervised [78–81,83,84,87–92], semi-supervised [82], and reinforcement learning [93]. The model column determines the algorithm used in work, CNN, RNN, GAN, KNN, SVM, GMM, . . . , etc. The anomaly column determines the type of anomaly that each study tries to detect. The dataset column names the dataset used in each study.

10.2. Violation Type

It is well known that the physical world produces abnormal behaviors that appear beyond explanation. Detecting these abnormal behaviors is not easy because it comes in several types. The studies [74,76–78] focus on detecting non-pedestrians and escape panics as abnormal behavior in a crowded place. Moreover, the study [79] includes more abnormal behavior such as irregular pedestrian movement and action differences from regular recognized events. It uses a single shot multi-box detector (SSD) to detect abnormal behavior in three different datasets PASCAL, VOC, and High-Speed Railway. The proposed

improved SSD network achieved increased results on the three datasets by 2.52% and 4.74%, respectively. While [75] proposes a novel Gaussian kernel-based integration model (GKIM) for anomalous entities detection and localization in pedestrian flows. Then, a block-based detection framework was developed by training a recurrent conditional random field (R-CRF) using the GKIM features. This [75] study divides the detection process into two types, groups and individuals.

The group anomaly behavior includes suddenly scattered crowds, and individuals include non-pedestrians, escape panics, and action differences from regular recognized events. The proposed framework outperforms the compared methods in terms of equal error rate (EER) and detection rate (DR) in both frame-level and pixel-level with three different datasets UCSD, UMN, and UCD. A deep learning model that can detect normal or abnormal actions on an academic campus using CCTV footage has been introduced in [80], which uses three different datasets, UCSD, UMN, and LV, to detect the anomaly. The model consists of two parts and two neural networks, CNN and RNN: First, CNN is used for high-level feature extraction from video frames. Second, based on the obtained features, the RNN classifier predicts the class as normal or abnormal using LSTM architecture. A pre-trained model VGG-16 was used in image feature extraction with videos obtained from CCTV cameras. The results show that the introduced model allows for the prevention of crimes before it occurs. The real-time CCTV images were tracked and analyzed automatically and achieved an accuracy of 87.15%. Fighting and violence are the most common abnormal behaviors that occurred in public places, which are addressed for detection in [81,82,85–89]. The vehicles and bicycles that drive oppositely, at fast speed, or at not allowed places, for example, on the pedestrian side, which cues as dangerous for people, is considered abnormal behavior in [82,85,91,92]. In universities, specifically inside the campus, fighting, and fainting are irregular behaviors. An abnormal behavior recognition system based on 3D-CNN and LSTM has been developed in [87] to detect abnormal behavior in universities. The 3D-CNN and LSTM models are employed to maintain motion correlation between consecutive feature images using 3D-ResNets architecture. Crossing the track at the train station or railway outside the pedestrian zone is considered a wrong behavior that should be avoided [83,84].

10.3. Scope of Application

Managing a large-scale crowd in crowded places is a business solution that offers an intelligent analysis of crowd mobility. It can be applied to applications that contain crowds, such as at shopping centers, queue detection, cultural events, public places theft detection, playgrounds, streets and highways, sports stadiums, train stations, and airport terminals, see Figure 5. For the Hajj aspect, an abnormal behavior detection approach based on optical flow and generative adversarial network (GAN) for crowd scenes anomaly detection has been proposed [27]. The optical flows are used to identify dynamic features. Then, an optical flow framework based on GAN has been employed with a transfer learning strategy to identify abnormal human behavior in large-scale crowd scenes during the Hajj. To differentiate between normal and abnormal behaviors, the U-Net and Flownet have been used. The suggested approach is evaluated using three datasets: UMN scenes 1, 2, 3, UCSD, and Abnormal Behaviors HAJJ datasets. The results indicate that the accuracy achieved with UMN scenes is 99.4%, 97.1%, and 97.6%, respectively.

Moreover, it achieves 89.26% with UCSD and 79.63% with the proposed Abnormal Behaviors HAJJ dataset. The model can work perfectly, but the accuracy requires to be enhanced by training the model with more samples and annotating more details. Moreover, a new crowd density prediction model for Hajj and Umrah crowd video analytics system has been proposed [95] to enhance the protection and safety of pilgrims in Makkah. CNN analyzed the crowd by counting the number of people in a specific area. The suggested model exceeds the state-of-the-art methods with a considerable decrease of MAE, which results in 240.0 and improved by 177.5 degrees, and MSE, which results in 260.5 and improved by 280.1 degrees, with the HAJJ-Crowd dataset. Indeed, COVID-19 is also spread

by crowds, which are classified as sensitive sources. For crowd management during the pandemic, a recommendation system has been developed that suggests the closest shopping centers or stores with the least estimated crowds near the user's location [96], which helps to avoid crowding and scrambling in stores. The top-K approach and behavioral game theory have been used to predict the user's choice and estimate the crowd level for the requested place. The model outcomes indicate an increase in the trust factor from 0.5 to 0.76 and reduce the crowd level by an average of 40%. A lightweight CNN framework [84] has been proposed for anomaly detection in smart cities that is functional for a real-world surveillance environment. The introduced framework contains three key phases: First, the lightweight CNN model is used to extract spatial features from sequence surveillance video frames. Second, create a feature vector from a series of 30 frames of the video. Third, the feature vector is fed to the residual LSTM to identify abnormal activities in a real-world environment. The system's outcome shows that using CNN features with the residual blocks in LSTM for sequence learning is effective for anomaly detection and recognition. A deep learning model that can detect normal or abnormal actions on an academic campus using CCTV footage has been introduced [80]. In case of an abnormal event, the model sends an alert message to the authority. The proposed model achieved an accuracy of 87.15% in abnormality detection in the academic campus area. Moreover, an industrial aspect required an intelligent real-time video surveillance system for anomaly detection to protect safety, which was developed in [97] and achieved good results.



Figure 5. Scope of application.

10.4. Real-Time vs. Offline

Locating moving objects in videos over time is known as object tracking [98]. It has a variety of applications in computer vision, such as analyzing human behavior in crowds [99,100], pedestrian tracking systems [101], body motion tracking in crowds [102], detecting anomalies in crowds [103,104], and monitoring traffic flow [105]. The capability to comprehend and model an object's motion is crucial to the success of a tracker. Tracking can be performed with one object or several objects simultaneously. Even if an environment contains several objects, only a single object is tracked in single tracking regardless of how many objects are presented. While multiple object tracking involves observing all the objects in the environment over time [106]. However, tracking and abnormality detection are imperative, whether performed online in real-time or offline. An online approach gathers real-time data about people and their behavior to achieve an understanding of their behavior so that abnormalities can be detected immediately. Since the frames are processed at the same time of occurrence, the subsequent frames cannot be used to predict and improve results; only previous frames can. In another situation, offline trackers will be used when tracking an object in a stream that has been recorded. Using the previous and subsequent frames, the program conducts batch processing of the frames to analyze the video stream and provide accurate results. An intelligence system that controls the crowd by detecting abnormal behavior using deep learning techniques through a real-time video surveillance system has been introduced [107]. This system allows avoiding injury or any other action which causes harmful effects to the community because of the crowd using both CNN and KNN. A real-time lightweight computational architecture for violence detection in a crowded public place using convolutional long short-term memory (Conv-LSTM) has been developed [108]. A dataset of crowd anomalies was used to validate the algorithm, which achieved 95.16% accuracy.

Moreover, an efficient system that can detect and locate abnormal behavior in surveillance videos in crowded events has been introduced [109]. The proposed system is based on a new Motion Information Image (MII) model expressed using optical flow and CNN. The outcomes reveal that the introduced system is very efficient and can identify and locate abnormal behaviors in real-time. The algorithm's accuracy outperforms the existing algorithms at both pixel and frame levels. In addition, a new system for real-time anomalous event detection in videos called MOVAD has been proposed [25]. It achieved comprehensive performance that exceeds the current state-of-the-art methods. Intelligent anomaly detection and classification systems were introduced in [85] to detect abnormal behavior in surveillance videos using Faster RCNN with Deep Reinforcement Learning (DRL) techniques for offline tracking. The proposed model has outperformed the other methods with the maximum accuracy of 98.50% and 94.80% on the test004 and test007 datasets. Another accurate and effective deep learning framework for detecting abnormal behavior in videos with Vgg-16 and LSTM has been developed [110]. Experimental results show that the proposed method achieves the best detection results at the frame and pixel levels. Indeed, a recurrent neural network (RNNs) and two-dimensional convolutional neural networks (2D CNN) have been developed for violence detection [111], that achieved an accuracy of 99%, 93.75%, and 96.74%, respectively, on the Hockey dataset, Violent Flow, and Real-Life Violence Situations Dataset.

Moreover, an abnormal behavior recognition system based on 3D-CNN and LSTM has been developed [87]. The 3D-CNN and LSTM models are employed to maintain motion correlation between consecutive feature images using 3D-ResNets architecture. The experiments show that the proposed method has an excellent performance in abnormal behavior recognition on some challenging datasets. An automatic abnormal behavior detection system of videos based on VGGNet and BSVM has been developed [112], it was applied through transfer learning strategies to detect abnormal events. The results illustrated that the VGGNet-19 network obtained better accuracy than other hand-crafted descriptors, with an average accuracy of 97.44%. A new fully convolutional neural networks (FCNs) architecture system for global abnormal behavior detection and localization in videos has been developed [113]. The proposed architecture is fast and accurate for anomaly detection in video data, which achieved a 370-fps processing speed on a standard GPU.

10.5. Human Crowd vs. Non-Human Crowd

The concept of "crowds" is not limited to specific objects. The term "crowds" can include different types of objects such as human crowds, vehicle crowds, crowds of animals, crowds of birds, and many more examples. A fundamental requirement for analyzing crowd scenarios is identifying the kind of crowd. Table 2 presents a summary of deep learning anomaly detection projects with different targets, including humans and non-humans. An anomaly detection system combining the optical flow method and convolutional neural network (CNN) has been introduced to identify and inform the irregularities of human and vehicle crowds in difficult video scenes [83]. The proposed system achieved an average accuracy of 86.3% and an average time of 12 s with the human crowd, while vehicle anomaly detection achieved 89.7% accuracy with an average time of 11 s. A novel architecture called DeepCrowd, which can detect and classify five different types of a crowd (vehicle crowd, human crowd, bird crowd, animal crowd, mixed crowd), has been developed [114]. The DeepCrowd system achieved a good accuracy of 83.11% in detecting and classifying the type of crowd. A unified autonomous system has been developed to detect risky human behavior in video surveillance systems or monitor systems RGB image based on a deep convolutional network [82]. The result shows the potential and possibility of the proposed system, which provides adequate achievement in distinguishing abnormal behavior in a real-world situation. For sparse crowds, an adaptive training-less method for anomaly detection in surveillance videos has been introduced [115]. It achieves comparable performance results with several state-of-the-art methods on publicly available UCSD, UMN, CUHK Avenue, and ShanghaiTech datasets.

Table 2. Categorization of the state-of-the-art anomaly detection methods in crowd scenes.

Ref.	Type	Approach	Anomaly	Scope	Processing	Target	Dataset
Classical Machine Learning							
[74]	Unsupervised	K-means	Non-pedestrians, escape panics	Public Places	Offline	Human	UCSD, UMN
[93]	Unsupervised	Dictionary learning	Suddenly scattered, non-pedestrians, escape panics	Public Places	Offline	Human	UCSD, UMN PETS2009, Avenue
[92]	Unsupervised	Soft Clustering	Non-pedestrian, escape panics	Public Places	Offline	Human	UMN, UCSD
[91]	Unsupervised	k-means	Non-pedestrian	Public Places	Offline	Human	UCSD
[89]	Supervised	Optical flow	Non-pedestrians, escape panics	Public Places	Offline	Human	UCSD, UMN
[75]	Supervised	GKIM, R-CRF	Non-pedestrians, panics, irregular movement	Public Places	Offline	Human	UCSD, UMN, UCD
[76]	Supervised	K-means, Linear SVM	Crowd running, crash, kidnap, burglary, fighting	Public Places	Offline	Human	UCSD, UMN, LV
[77]	Supervised	SVM	Panics, fighting, running, standing	Public Places	Offline	Human	UMN, BEHAVE
[78]	Semi-Supervised	GMM, SVM	Violent, panics	Public Places	Real-Time	Human	UMN, Violent flows
Deep Learning							
[79]	Supervised	SSD, VGG-16	Bullet train, pedestrian	Railway	Offline	Human Train	PASCAL VOC, Railway
[90]	Supervised	SSD, VGG-16	Small object	Railway	Real-time	-	ILSVRC CLS-LOC, Railway
[88]	Unsupervised	GAN	Biking, fighting, vehicle, running	Public Places	Offline	Human Vehicle	CUHK Avenue UCSD, Campus ShanghaiTech
[87]	Unsupervised	3D-CNN LSTM	Panics, fighting, protest	Public Places	Offline	Human	UMN, CAVIA, Web
[94]	Supervised	Modified 3D ConvNet	Violent	Public Places	Offline	Human	Crowd violence

Table 2. Cont.

Ref.	Type	Approach	Anomaly	Scope	Processing	Target	Dataset
[80]	Supervised	CNN RNN	Use mobile in class, fighting, fainting	University	Offline	Human	KTH, CAVIAR
[81]	Supervised	CNN	Walking, jogging, fighting, kicking, punching	Public Places	Offline	Human	CMU, UTI PEL, HOF WED
[82]	Supervised	VGG-16 LSTM	Kicking, pointing punching, pushing	Public Places	Offline	Human	UT-Interaction-Data
[83]	Supervised	Optical Flow CNN	Panic, running fast speed, crash	Public Places	Offline	Human Vehicle	UCSD, UMN
[84]	Supervised	CNN Residual LSTM	Fighting, explosion, accidents, shooting, robbery, shoplifting, burglary	Smart Cities	Real-Time	Human	UCF-Crime, UMN, Avenue
[85]	Reinforcement Learning	Faster RCNN	Car, bicycle	Surveillance System	Offline	Vehicle	UCSD
[25]	Supervised	CNN, RNN KNN, Optical Flow	Bicycles, skateboards, wheelchairs	Public Places	Real-Time	Human vehicles	CUHK Avenue UCSD, campus, ShanghaiTeh, UR fall
[27]	Supervised	Optical Flow GAN	Standing, sitting, sleeping, running, moving in opposite, non-pedestrian	Hajj	Real-Time	Human Cars Wheelchairs	UMN, UCSD, HAJJ datasets
[95]	Supervised	CNN	Density	Hajj, Umrah	Real-Time	Human	HAJJ-Crowd
[96]	-	point-of-interests (POI)	Crowding, scrambling	Shopping Centers	Real-Time	Human	-
[97]	Unsupervised	CNN, Conv-LSTM	People littering, skateboard, Discarding items, loitering	Industrial	Real-Time	Human	CUHK Avenue UCSD Ped 1 UCSD Ped 2
[107]	Supervised	CNN, KNN	Injury	Public Places	Real-Time	Human	UMN
[108]	Supervised	Conv-LSTM	Violence	Public Places	Real-Time	Human	Standard crowd anomaly
[109]	Supervised	CNN, MII Optical Flow	Escape or panic situation	Public Places	Real-Time	Human	UMN PETS2009
[110]	Unsupervised	Vgg-16 and LSTM	Non-pedestrian	Public Places	Offline	Human Cars	UCSD Ped2 CUHK Avenue

Table 2. Cont.

Ref.	Type	Approach	Anomaly	Scope	Processing	Target	Dataset
[111]	Unsupervised	RNN, 2D CNN	Violence	Public Places	Offline	Human	Hockey, Violent-Flow, Real-Life Violence Situations
[112]	Supervised	VGGNet-19 BSVM	Running, Carts Bikers, Skateboarder	Public Places	Offline	Human	UMN, CSD-PED1
[113]	Supervised	FCNs	Car Skateboarder Wheelchair Bicycle, Wrong direction	Public Places	Offline	Human	UCSD, Subway
[114]	Supervised	2D CNN	-	Public Places	Offline	Vehicle, Human Animal, Bird Mixed	CVML Crowd Variety
[115]	Supervised	Optical Flow	Panics, loitering, running, throwing objects	Surveillance System	Offline	Human	UCSD, UMN CUHK Avenue ShanghaiTech

11. Publicly Available Datasets for Crowd Applications

Large-scale applications for crowd management have received significant attention over the last ten years. For the management and control of crowds, crowd analysis is crucial in intelligent video surveillance systems. The collection of crowd motion video data is not an easy task. During the past few years, more and more datasets have been created that focus on crowd density estimation, crowd analysis, and anomaly detection in crowded scenes. The use of these datasets allows for improving the quality of crowd applications. Table 3 presents a summary of publicly available crowd datasets.

Table 3. Summary of available crowd datasets.

Ref.	Year	Name	Scale	Train	Test	Total	Description
[116]	2021	CVCS	Medium	-	-	31	Multi-view crowd counting
[117]	2021	DroneCrowd	Large	-	-	112	Detection, tracking, and counting animal crowds with drones
[27]	2020	HAIJv1	Large	-	-	-	Human abnormal behavior in Hajj
[118]	2020	UCF-QNRF	Large	-	-	1535	Crowd counting and localization
[119]	2020	NWPU-Crowd	Large	-	-	5109	Crowd counting and localization
[120]	2019	DLR-ACD	Large	-	-	33	Crowd counting, density estimation, and localization
[121]	2019	JHU-CROWD	Large	-	-	-	Crowd counting dataset under different weather conditions
[122]	2020	JHU-CROWD++	Large	-	-	4372	
[123]	2018	CrowdFlow	Large	-	-	10	Crowd analysis, crowd flow, and movement estimation
[124]	2018	SCUT-HEAD	Large	-	-	4405	Head detection
[125]	2018	SmartCity	Large	-	-	50	Crowd counting
[126]	2017	Multi-Task Crowd	Large	-	-	100	Crowd counting, violence detection, and density level classification
[127]	2016	Shanghai Tech Part A Part B	Large	-	-	482 716	Crowd counting and density estimation
[128]	2015	WorldExpo '10	Large	-	-	3980	Crowd counting in a cross-scene
[129]	2015	WWW Crowd	Large	-	-	10,000	Crowd understanding
[130]	2015	SHOCK	Large	-	-	-	Analyze spectator crowd behavior at stadiums/theaters/events
[131]	2014	CUHK Crowd	Large	-	-	474	Analyze group behavior in crowd scenes.
[132]	2014	Crowd Saliency	Large	-	-	-	Crowd movement, counter flow, source, sink, and instability motion
[133]	2013	UCF-CC-50	Large	-	-	50	Extremely dense crowd dataset for crowd counting
[134]	2012	AGORASET	Large	-	-	-	Crowd motion simulation
[135]	2012	Violent flows	Large	-	-	246	Classify and detect violent and non-violent behavior
[136]	2012	Mall	Medium	-	-	2000	Crowd counting
[137]	2012	Grand Central	Medium	-	-	-	Crowd train station dataset
[138]	2009	PETS2009	Medium	-	-	875	Crowd counting, density estimation, tracking, and event detection
[139]	2009	UMN	Small	-	-	11	Abnormal crowd behavior detection
[140]	2008	UCSD Peds 1 UCSD Peds 2	Small	6800 2550	7200 2010	40 12	Abnormal crowd behavior detection

12. Discussion

A review examining recent research in crowd anomaly detection in automated surveillance systems has been presented in this paper, which includes the key aspects of the problem domain, approach, and method. Since video surveillance systems are widely

used in public places, crowd anomaly detection has become an increasingly critical part of the intelligent video surveillance system. In intelligent video surveillance, anomaly detection and localization remain challenging problems. The definition of the anomaly is significantly different from one situation to another, which means anomalies in a specific situation may not be an anomaly in another situation. Consequently, the type of event depends on the surrounding circumstances. Several types of abnormal behaviors depend on the environment and circumstances, making detecting them difficult. To discover these behaviors, it is necessary to understand the surrounding environment and the expected and unexpected behavior to facilitate the classification of any other behavior that may occur. The anomaly behavior varies according to the environment; specifying the behavior more precisely ensures a more accurate discovery of the events. According to Table 2, most of the studies in this area focus on one target, abnormal behavior detection, and few studies investigate multiple targets. However, addressing multiple targets is very common as a realistic scenario and can be challenging since each target must be addressed with a different feature extraction method.

Furthermore, most human anomaly detection applications cover a range of viewing fields from around 10 M to a medium size area of around 100 M. It is very rare to find some applications that examine the effects of human abnormalities detection in very small or very large fields of view, such as that seen in a satellite image. In addition, object occlusions, inconsistencies in perspective due to different angles and camera positions, as well as tilting and up-and-down motion of the camera, can occur in large-scale crowd scenes, making crowd analysis very challenging. The use of multiple CCTVs that cover complete angles to monitor the crowd and provide a full (360°) view of objects in that area to avoid occlusion. Furthermore, drones and satellite images will add more valuable results during crowd monitoring and abnormality detection. According to the literature and previous applications, classical machine learning techniques are often outperformed by deep learning techniques. Machine learning algorithms can still provide good results by training the model well, but deep learning algorithms provide more accurate results in large and complex applications for greater accuracy. However, simple, direct, and clear requirements projects can be performed with machine learning algorithms without wasting resources. Indeed, some of the literature on video surveillance systems does not provide an exact distinction between real-time anomalous event detection and offline anomalous detection. The two approaches are different from each other in the data processing. Offline video tracking involves analyzing sequential video frames and relating target objects based on their appearance in each frame. While real-time object tracking involves tracking moving objects in video from a camera over time. Moreover, many types of objects can be characterized as crowds. The term crowd does not refer to any single type of object. Identifying the nature of the crowd is a prerequisite for analyzing crowd scenarios. It could be a crowd of people, vehicles, animals, or any other object.

13. Trends and Future Works

Detecting abnormal crowd behavior in video scenes is a hot research topic currently. Additional developments and improvements will help to achieve effective crowd management and ensure human safety. Video anomaly detection has a wide range of potential application domains such as crime detection, pedestrian tracking systems, traffic violations, body motion tracking in crowds, abnormal crowd behavior, intrusion detection, fraud detection, abandoned objects, health monitoring, weapons at sensitive areas, and industrial protection. For future research, the ability to analyze crowd behavior at the macro and micro levels will generate valuable information to understand and manage crowd behavior. The research on large-scale crowd object detection still needs further development. In a dense crowd, it is not easy for the object detector to pinpoint the position where events happen in video frames. Additional improvement in this aspect will produce an accurate detection result. Most of the research has recently focused on supervised approaches while monitoring the real environment produces large streams of unannotated data. Therefore,

there is a need to improve the unsupervised real-time anomaly detection applications. New research areas have been identified, especially related to the crowds during pandemics the world is suffering.

Moreover, the use of adaptive deep networks is an advanced solution that utilizes real-time data to detect the anomaly. Exploring and determining the specific type of abnormal behavior for multiple people in the crowd under more diverse situations is another open challenge. Another important aspect of most deep learning projects, specifically for anomaly detection in crowds, is how to optimize and improve the model's performance during running time. This point opens another chance for achieving a competitive detection time in the crowd. Further improvements are required for low resolution, illumination variation, and occlusion in data. Finally, the researchers emphasize the importance of advancing the topic quickly and appropriately.

14. Conclusions

Over the last decade, CCTV surveillance has become more prevalent in crowded public places. This has led to more video data being produced than can be analyzed by an individual. Therefore, automated systems are necessary for analyzing large volumes of video streams in crowded areas to detect anomalies, ensure safety, and respond quickly. This paper reviews the recent development of automated anomaly detection systems from the perspective of computer vision. In addition, the taxonomic organization of existing works in crowd analysis and anomaly detection has been introduced. Previous applications provided great applications in detecting abnormal behavior. However, there is still a greater necessity to obtain higher performance and greater accuracy in detecting anomalies in crowded areas. Even though there have been numerous studies on detecting abnormal human behavior, more research is still needed to address numerous issues. Crowd abnormal behavior detection should be more accurate and robust against different situations in large-scale and heterogeneous crowds. Using advanced technology in monitoring the crowd, such as drones and satellites, will provide additional valuable insight.

Author Contributions: Conceptualization, A.A. and W.A.; methodology, A.A. and W.A.; validation, A.A. and W.A.; formal analysis, A.A. and W.A.; investigation, A.A.; resources, A.A. and W.A.; data curation, A.A. and W.A.; writing—original draft preparation, A.A.; writing—review and editing, W.A.; supervision, W.A.; project administration, W.A.; funding acquisition, W.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Deanship of Scientific Research, Qassim University.

Acknowledgments: The researchers would like to thank the Deanship of Scientific Research, Qassim University for funding the publication of this project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Owaidah, A.; Olaru, D.; Bennamoun, M.; Sohel, F.-D.; Khan, N. Review of Modelling and Simulating Crowds at Mass Gathering Events: Hajj as a Case Study. *J. Artif. Soc. Soc. Simul.* **2019**, *22*, 9. [[CrossRef](#)]
2. Bendali-Braham, M.; Weber, J.; Forestier, G.; Idoumghar, L.; Muller, P.-A. Recent Trends in Crowd Analysis: A Review. *Mach. Learn. Appl.* **2021**, *4*, 100023. [[CrossRef](#)]
3. Singh, V.; Singh, S.; Gupta, P. Real-Time Anomaly Recognition Through CCTV Using Neural Networks. *Procedia Comput. Sci.* **2020**, *173*, 254–263. [[CrossRef](#)]
4. Xu, M.; Yu, X.; Chen, D.; Wu, C.; Jiang, Y. An Efficient Anomaly Detection System for Crowded Scenes Using Variational Autoencoders. *Appl. Sci.* **2019**, *9*, 3337. [[CrossRef](#)]
5. Nguyen, T.-N.; Meunier, J. Anomaly Detection in Video Sequence with Appearance-Motion Correspondence. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27–28 October 2019; pp. 1273–1283.
6. Doshi, K.; Yilmaz, Y. Continual Learning for Anomaly Detection in Surveillance Videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Seattle, WA, USA, 13–19 June 2020; pp. 254–255.
7. Yamin, M.; Basahel, A.M.; Abi Sen, A.A. Managing Crowds with Wireless and Mobile Technologies. *Wirel. Commun. Mob. Comput.* **2018**, *2018*, 1–15. [[CrossRef](#)]

8. Zhang, X.; Yu, Q.; Yu, H. Physics Inspired Methods for Crowd Video Surveillance and Analysis: A Survey. *IEEE Access* **2018**, *6*, 66816–66830. [[CrossRef](#)]
9. Singh, U.; Determe, J.-F.; Horlin, F.; de Doncker, P. Crowd Monitoring: State-of-the-Art and Future Directions. *IETE Tech. Rev.* **2021**, *38*, 578–594. [[CrossRef](#)]
10. Elbishlawi, S.; Abdelpakey, M.H.; Eltantawy, A.; Shehata, M.S.; Mohamed, M.M. Deep Learning-Based Crowd Scene Analysis Survey. *J. Imaging* **2020**, *6*, 95. [[CrossRef](#)]
11. Sundararaman, R.; de Azevedo Braga, C.; Marchand, E.; Pettré, J. Tracking Pedestrian Heads in Dense Crowd. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 3865–3875.
12. Siddiqui, A.A.; Gwynne, S.M.V. Employing Pedestrian Observations in Engineering Analysis. *Saf. Sci.* **2012**, *50*, 478–493. [[CrossRef](#)]
13. Basahel, A.; Yamin, M. Measuring Success of E-Government of Saudi Arabia. *Int. J. Inf. Technol.* **2017**, *9*, 287–293. [[CrossRef](#)]
14. Bhat, G.; Danelljan, M.; van Gool, L.; Timofte, R. Know Your Surroundings: Exploiting Scene Information for Object Tracking. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; Springer: Cham, Switzerland; pp. 205–221.
15. Cheong, Y.Z.; Chew, W.J. The Application of Image Processing to Solve Occlusion Issue in Object Tracking. *MATEC Web Conf.* **2018**, *152*, 03001. [[CrossRef](#)]
16. Madhuri, G.S.; Rani, M.U. Anomaly Detection Techniques. *SSRN Electron. J.* **2018**, *7*, 449–453. [[CrossRef](#)]
17. Chalapathy, R.; Chawla, S. Deep Learning for Anomaly Detection: A Survey. *arXiv* **2019**, arXiv:1901.03407.
18. Ullah, F.U.M.; Ullah, A.; Muhammad, K.; Haq, I.U.; Baik, S.W. Violence Detection Using Spatiotemporal Features with 3D Convolutional Neural Network. *Sensors* **2019**, *19*, 2472. [[CrossRef](#)] [[PubMed](#)]
19. Sabih, M.; Vishwakarma, D.K. Crowd Anomaly Detection with LSTMs Using Optical Features and Domain Knowledge for Improved Inferring. *Vis. Comput.* **2021**, *38*, 1719–1730. [[CrossRef](#)]
20. Zhang, X.; Yang, S.; Tang, Y.Y.; Zhang, W. A Thermodynamics-Inspired Feature for Anomaly Detection on Crowd Motions in Surveillance Videos. *Multimed. Tools Appl.* **2016**, *75*, 8799–8826. [[CrossRef](#)]
21. Jeremy, M.; Leach, V. Automatic Human Behaviour Anomaly Detection in Surveillance Video. Ph.D. Thesis, Heriot-Watt University, Edinburgh, Scotland, 2015.
22. Yang, J.; Xu, R.; Qi, Z.; Shi, Y. Visual Anomaly Detection for Images: A Survey. *arXiv* **2021**, arXiv:2109.13157.
23. Lin, W.; Gao, J.; Wang, Q.; Li, X. Learning to Detect Anomaly Events in Crowd Scenes from Synthetic Data. *Neurocomputing* **2021**, *436*, 248–259. [[CrossRef](#)]
24. Usman, I.; Albeshier, A.A. Abnormal Crowd Behavior Detection Using Heuristic Search and Motion Awareness. *Int. J. Comput. Sci. Netw. Secur.* **2021**, *21*, 131–139. [[CrossRef](#)]
25. Doshi, K.; Yilmaz, Y. A Modular and Unified Framework for Detecting and Localizing Video Anomalies. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 4–8 January 2022; pp. 3982–3991.
26. Li, A.; Miao, Z.; Cen, Y.; Zhang, X.-P.; Zhang, L.; Chen, S. Abnormal Event Detection in Surveillance Videos Based on Low-Rank and Compact Coefficient Dictionary Learning. *Pattern Recognit.* **2020**, *108*, 107355. [[CrossRef](#)]
27. Alafif, T.; Alzahrani, B.; Cao, Y.; Alotaibi, R.; Barnawi, A.; Chen, M. Generative Adversarial Network Based Abnormal Behavior Detection in Massive Crowd Videos: A Hajj Case Study. *J. Ambient. Intell. Humaniz. Comput.* **2022**, *1*, 3. [[CrossRef](#)]
28. Gao, J.; Gan, L.; Buschendorf, F.; Zhang, L.; Liu, H.; Li, P.; Dong, X.; Lu, T. Omni SCADA Intrusion Detection Using Deep Learning Algorithms. *IEEE Internet Things J.* **2021**, *8*, 951–961. [[CrossRef](#)]
29. Hassan, M.M.; Gumaei, A.; Alsanad, A.; Alrubaiyan, M.; Fortino, G. A Hybrid Deep Learning Model for Efficient Intrusion Detection in Big Data Environment. *Inf. Sci.* **2020**, *513*, 386–396. [[CrossRef](#)]
30. Yin, C.; Zhu, Y.; Fei, J.; He, X. A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks. *IEEE Access* **2017**, *5*, 21954–21961. [[CrossRef](#)]
31. Voican, O. Credit Card Fraud Detection Using Deep Learning Techniques. *Inform. Econ.* **2021**, *25*, 70–85. [[CrossRef](#)]
32. Zhang, X.; Han, Y.; Xu, W.; Wang, Q. HOBA: A Novel Feature Engineering Methodology for Credit Card Fraud Detection with a Deep Learning Architecture. *Inf. Sci.* **2021**, *557*, 302–316. [[CrossRef](#)]
33. Tahura, S.; Hasnat Samiul, S.M.; Shamim Kaiser, M.; Mahmud, M. Anomaly Detection in Electroencephalography Signal Using Deep Learning Model. In Proceedings of the International Conference on Trends in Computational and Cognitive Engineering, Istanbul, Turkey, 21–22 October 2021; Springer: Singapore; pp. 205–217.
34. Yang, J.; Yang, F.; Zhang, L.; Li, R.; Jiang, S.; Wang, G.; Zhang, L.; Zeng, Z. Bridge Health Anomaly Detection Using Deep Support Vector Data Description. *Neurocomputing* **2021**, *444*, 170–178. [[CrossRef](#)]
35. Bao, Y.; Tang, Z.; Li, H.; Zhang, Y. Computer Vision and Deep Learning-Based Data Anomaly Detection Method for Structural Health Monitoring. *Struct. Health Monit.* **2019**, *18*, 401–421. [[CrossRef](#)]
36. Saini, D.K.; Ahir, D.; Ganatra, A. Techniques and Challenges in Building Intelligent Systems: Anomaly Detection in Camera Surveillance. In Proceedings of the First International Conference on Information and Communication Technology for Intelligent Systems, Ahmedabad, India, 28–29 November 2015; Springer: Cham, Switzerland, 2016; Volume 51, pp. 11–21.
37. Basora, L.; Olive, X.; Dubot, T. Recent Advances in Anomaly Detection Methods Applied to Aviation. *Aerospace* **2019**, *6*, 117. [[CrossRef](#)]

38. Shaukat, K.; Alam, T.M.; Luo, S.; Shabbir, S.; Hameed, I.A.; Li, J.; Abbas, S.K.; Javed, U. A Review of Time-Series Anomaly Detection Techniques: A Step to Future Perspectives. In *Advances in Intelligent Systems and Computing*; Springer: Cham, Switzerland, 2021; pp. 865–877.
39. Canizo, M.; Triguero, I.; Conde, A.; Onieva, E. Multi-Head CNN–RNN for Multi-Time Series Anomaly Detection: An Industrial Case Study. *Neurocomputing* **2019**, *363*, 246–260. [[CrossRef](#)]
40. Lu, Y. Industry 4.0: A Survey on Technologies, Applications and Open Research Issues. *J. Ind. Inf. Integr.* **2017**, *6*, 1–10. [[CrossRef](#)]
41. Vidhya, J.V.; Uthra, R.A. Violence Detection in Videos Using Conv2D VGG-19 Architecture and LSTM Network. In Proceedings of the Algorithms, Computing and Mathematics Conference, Chennai, India, 19–20 August 2021.
42. Ren, J.; Xia, F.; Liu, Y.; Lee, I. Deep Video Anomaly Detection: Opportunities and Challenges. In Proceedings of the 2021 International Conference on Data Mining Workshops (ICDMW), Auckland, New Zealand, 7–10 December 2021; IEEE: Piscataway Township, NJ, USA, 2021; pp. 959–966.
43. Buch, N.; Velastin, S.A.; Orwell, J. A Review of Computer Vision Techniques for the Analysis of Urban Traffic. *IEEE Trans. Intell. Transp. Syst.* **2011**, *12*, 920–939. [[CrossRef](#)]
44. Sodemann, A.A.; Ross, M.P.; Borghetti, B.J. A Review of Anomaly Detection in Automated Surveillance. *IEEE Trans. Syst. Man Cybern. Part C* **2012**, *42*, 1257–1272. [[CrossRef](#)]
45. Popoola, O.P.; Wang, K. Video-Based Abnormal Human Behavior Recognition—A Review. *IEEE Trans. Syst. Man Cybern. Part C* **2012**, *42*, 865–878. [[CrossRef](#)]
46. Vishwakarma, S.; Agrawal, A. A Survey on Activity Recognition and Behavior Understanding in Video Surveillance. *Vis. Comput.* **2013**, *29*, 983–1009. [[CrossRef](#)]
47. Gowsikhaa, D.; Abirami, S.; Baskaran, R. Automated Human Behavior Analysis from Surveillance Videos: A Survey. *Artif. Intell. Rev.* **2014**, *42*, 747–765. [[CrossRef](#)]
48. Wang, X. Intelligent Multi-Camera Video Surveillance: A Review. *Pattern Recognit. Lett.* **2013**, *34*, 3–19. [[CrossRef](#)]
49. Pimentel, M.A.F.; Clifton, D.A.; Clifton, L.; Tarassenko, L. A Review of Novelty Detection. *Signal Process.* **2014**, *99*, 215–249. [[CrossRef](#)]
50. Chong, Y.S.; Tay, Y.H. Modeling Video-Based Anomaly Detection Using Deep Architectures: Challenges and Possibilities. In Proceedings of the 2015 10th Asian Control Conference (ASCC), Kota Kinabalu, Malaysia, 31 May–3 June 2015; IEEE: Piscataway Township, NJ, USA, 2015; pp. 1–8.
51. Patil, N.; Biswas, P.K. A Survey of Video Datasets for Anomaly Detection in Automated Surveillance. In Proceedings of the 2016 Sixth International Symposium on Embedded Computing and System Design (ISED), Patna, India, 15–17 December 2016; IEEE: Piscataway Township, NJ, USA, 2016; pp. 43–48.
52. Yogameena, B.; Nagananthini, C. Computer Vision Based Crowd Disaster Avoidance System: A Survey. *Int. J. Disaster Risk Reduct.* **2017**, *22*, 95–129. [[CrossRef](#)]
53. Ma, J.; Dai, Y.; Hirota, K. A Survey of Video-Based Crowd Anomaly Detection in Dense Scenes. *J. Adv. Comput. Intell. Intell. Inform.* **2017**, *21*, 235–246. [[CrossRef](#)]
54. Swathi, H.Y.; Shivakumar, G.; Mohana, H.S. Crowd Behavior Analysis: A Survey. In Proceedings of the 2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT), Bangalore, India, 16–17 March 2017; IEEE: Piscataway Township, NJ, USA, 2017; pp. 169–178.
55. Grant, J.M.; Flynn, P.J. Crowd Scene Understanding from Video. *ACM Trans. Multimed. Comput. Commun. Appl.* **2017**, *13*, 1–23. [[CrossRef](#)]
56. Pawar, K.; Attar, V. Deep Learning Approaches for Video-Based Anomalous Activity Detection. *World Wide Web* **2019**, *22*, 571–601. [[CrossRef](#)]
57. Tripathi, R.K.; Jalal, A.S.; Agrawal, S.C. Suspicious Human Activity Recognition: A Review. *Artif. Intell. Rev.* **2018**, *50*, 283–339. [[CrossRef](#)]
58. Kiran, B.; Thomas, D.; Parakkal, R. An Overview of Deep Learning Based Methods for Unsupervised and Semi-Supervised Anomaly Detection in Videos. *J. Imaging* **2018**, *4*, 36. [[CrossRef](#)]
59. ben Mabrouk, A.; Zagrouba, E. Abnormal Behavior Recognition for Intelligent Video Surveillance Systems: A Review. *Expert Syst. Appl.* **2018**, *91*, 480–491. [[CrossRef](#)]
60. Ahmed, S.A.; Dogra, D.P.; Kar, S.; Roy, P.P. Trajectory-Based Surveillance Analysis: A Survey. *IEEE Trans. Circuits Syst. Video Technol.* **2019**, *29*, 1985–1997. [[CrossRef](#)]
61. Santhosh, K.K.; Dogra, D.P.; Roy, P.P. Anomaly Detection in Road Traffic Using Visual Surveillance. *ACM Comput. Surv.* **2021**, *53*, 1–26. [[CrossRef](#)]
62. Khan, A.; Ali Shah, J.; Kadir, K.; Albattah, W.; Khan, F. Crowd Monitoring and Localization Using Deep Convolutional Neural Network: A Review. *Appl. Sci.* **2020**, *10*, 4781. [[CrossRef](#)]
63. Pang, G.; Shen, C.; Cao, L.; Hengel, A. van den Deep Learning for Anomaly Detection. *ACM Comput. Surv.* **2021**, *54*, 1–38. [[CrossRef](#)]
64. Nayak, R.; Pati, U.C.; Das, S.K. A Comprehensive Review on Deep Learning-Based Methods for Video Anomaly Detection. *Image Vis. Comput.* **2021**, *106*, 104078. [[CrossRef](#)]
65. Rezaee, K.; Rezakhani, S.M.; Khosravi, M.R.; Moghimi, M.K. A Survey on Deep Learning-Based Real-Time Crowd Anomaly Detection for Secure Distributed Video Surveillance. *Pers. Ubiquitous Comput.* **2021**, 1–17. [[CrossRef](#)]

66. Sinha, A.; Padhi, S.; Shikalgar, S. A Survey and Analysis of Crowd Anomaly Detection Techniques. In Proceedings of the 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 4–6 February 2021; IEEE: Piscataway Township, NJ, USA, 2021; pp. 1500–1504.
67. Chaudhary, D.; Kumar, S.; Dhaka, V.S. Video Based Human Crowd Analysis Using Machine Learning: A Survey. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2022**, *10*, 113–131. [[CrossRef](#)]
68. Waqar, S.; Khan, U.G.; Waseem, M.H.; Qayyum, S. The Utility of Datasets in Crowd Modelling and Analysis: A Survey. *Multimed. Tools Appl.* **2022**, 1–32. [[CrossRef](#)]
69. Tyagi, B.; Nigam, S.; Singh, R. A Review of Deep Learning Techniques for Crowd Behavior Analysis. *Arch. Comput. Methods Eng.* **2022**, *1404*, 1–29. [[CrossRef](#)]
70. Roka, S.; Diwakar, M.; Karanwal, S. A Review in Anomalies Detection Using Deep Learning. In Proceedings of the Third International Conference on Sustainable Computing, Jaipur, India, 19–20 March 2020; Springer: Singapore; pp. 329–338.
71. Xia, X.; Pan, X.; Li, N.; He, X.; Ma, L.; Zhang, X.; Ding, N. GAN-Based Anomaly Detection: A Review. *Neurocomputing* **2022**, *493*, 497–535. [[CrossRef](#)]
72. Bhuiyan, R.; Abdullah, J.; Hashim, N.; Al Farid, F. Video Analytics Using Deep Learning for Crowd Analysis: A review. *Multimed. Tools Appl.* **2022**, *81*, 27895–27922. [[CrossRef](#)]
73. Patrikar, D.R.; Parate, M.R. Anomaly Detection Using Edge Computing in Video Surveillance System: Review. *Int. J. Multimed. Inf. Retr.* **2022**, *11*, 85–110. [[CrossRef](#)] [[PubMed](#)]
74. Bansod, S.D.; Nandedkar, A.V. Crowd Anomaly Detection and Localization Using Histogram of Magnitude and Momentum. *Vis. Comput.* **2020**, *36*, 609–620. [[CrossRef](#)]
75. Ullah, H.; Altamimi, A.B.; Uzair, M.; Ullah, M. Anomalous Entities Detection and Localization in Pedestrian Flows. *Neurocomputing* **2018**, *290*, 74–86. [[CrossRef](#)]
76. Khan, M.U.K.; Park, H.S.; Kyung, C.M. Rejecting Motion Outliers for Efficient Crowd Anomaly Detection. *IEEE Trans. Inf. Forensics Secur.* **2018**, *14*, 541–556. [[CrossRef](#)]
77. Cui, X.; Liu, Q.; Gao, M.; Metaxas, D.N. Abnormal Detection Using Interaction Energy Potentials. In Proceedings of the CVPR 2011, Colorado Springs, CO, USA, 20–25 June 2011; IEEE: Piscataway Township, NJ, USA, 2011; pp. 3161–3167.
78. Marsden, M.; Mcguinness, K.; Little, S.; O’connor, N.E. Holistic features for real-time Crowd behaviour Anomaly detection. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; IEEE: Piscataway Township, NJ, USA, 2016; pp. 918–922.
79. Guo, B.; Shi, J.; Zhu, L.; Yu, Z. High-Speed Railway Clearance Intrusion Detection with Improved SSD Network. *Appl. Sci.* **2019**, *9*, 2981. [[CrossRef](#)]
80. Amrutha, C.V.; Jyotsna, C.; Amudha, J. Deep Learning Approach for Suspicious Activity Detection from Surveillance Video. In Proceedings of the 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2020), Bangalore, India, 5–7 March 2020. [[CrossRef](#)]
81. Tay, N.C.; Ong, K.; Goh, M. A Robust Abnormal Behavior Detection Method Using Convolutional Neural Network. In *Computational Science and Technology*; Springer: Singapore, 2019; pp. 37–47.
82. Ko, K.E.; Sim, K.B. Deep Convolutional Framework for Abnormal Behavior Detection in a Smart Surveillance System. *Eng. Appl. Artif. Intell.* **2018**, *67*, 226–234. [[CrossRef](#)]
83. Ma, J.; Xu, Y.; Zhang, Y.; Jiang, Y. An Abnormal Behavior Detection Method of Video Crowds and Vehicles Based on Deep Learning. In Proceedings of the 2nd International Conference on Artificial Intelligence and Pattern Recognition, Beijing, China, 16–18 August 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 10–12.
84. Ullah, W.; Ullah, A.; Hussain, T.; Khan, Z.A.; Baik, S.W. An Efficient Anomaly Recognition Framework Using an Attention Residual Lstm in Surveillance Videos. *Sensors* **2021**, *21*, 2811. [[CrossRef](#)]
85. Mansour, R.F.; Escorcía-Gutierrez, J.; Gamarra, M.; Villanueva, J.A.; Leal, N. Intelligent Video Anomaly Detection and Classification Using Faster RCNN with Deep Reinforcement Learning Model. *Image Vis. Comput.* **2021**, *112*, 104229. [[CrossRef](#)]
86. O’ Mahony, N.; Campbell, S.; Carvalho, A.; Harapanahalli, S.; Velasco Hernandez, G.; Krpalkova, L.; Riordan, D.; Walsh, J. Deep Learning vs. Traditional Computer Vision. In Proceedings of the Science and Information Conference, Las Vegas, NV, USA, 25–26 April 2019; Springer: Cham, Switzerland, 2019; pp. 128–144.
87. Guan, Y.; Hu, W.; Hu, X. Abnormal Behavior Recognition Using 3D-CNN Combined with LSTM. *Multimed. Tools Appl.* **2021**, *80*, 18787–18801. [[CrossRef](#)]
88. Hao, Y.; Li, J.; Wang, N.; Wang, X.; Gao, X. Spatiotemporal Consistency-Enhanced Network for Video Anomaly Detection. *Pattern Recognit.* **2022**, *121*, 108232. [[CrossRef](#)]
89. Biswas, S.; Gupta, V. Abnormality Detection in Crowd Videos by Tracking Sparse Components. *Mach. Vis. Appl.* **2017**, *28*, 35–48. [[CrossRef](#)]
90. Li, Y.; Dong, H.; Li, H.; Zhang, X.; Zhang, B.; Xiao, Z. Multi-Block SSD Based on Small Object Detection for UAV Railway Scene Surveillance. *Chin. J. Aeronaut.* **2020**, *33*, 1747–1755. [[CrossRef](#)]
91. Xu, D.; Song, R.; Wu, X.; Li, N.; Feng, W.; Qian, H. Video Anomaly Detection Based on a Hierarchical Activity Discovery within Spatio-Temporal Contexts. *Neurocomputing* **2014**, *143*, 144–152. [[CrossRef](#)]
92. Li, N.; Wu, X.; Xu, D.; Guo, H.; Feng, W. Spatio-Temporal Context Analysis within Video Volumes for Anomalous-Event Detection and Localization. *Neurocomputing* **2015**, *155*, 309–319. [[CrossRef](#)]

93. Hu, X.; Huang, Y.; Duan, Q.; Ci, W.; Dai, J.; Yang, H. Abnormal Event Detection in Crowded Scenes Using Histogram of Oriented Contextual Gradient Descriptor. *Eurasip J. Adv. Signal Process.* **2018**, *2018*, 54. [[CrossRef](#)]
94. Song, W.; Zhang, D.; Zhao, X.; Yu, J.; Zheng, R.; Wang, A. A Novel Violent Video Detection Scheme Based on Modified 3D Convolutional Neural Networks. *IEEE Access* **2019**, *7*, 39172–39179. [[CrossRef](#)]
95. Bhuiyan, R.; Abdullah, J.; Hashim, N.; Al Farid, F.; Samsudin, M.A.; Abdullah, N.; Uddin, J. Hajj Pilgrimage Video Analytics Using CNN. *Bull. Electr. Eng. Inform.* **2021**, *10*, 2598–2606. [[CrossRef](#)]
96. Durán-Polanco, L.; Siller, M. Crowd Management COVID-19. *Annu. Rev. Control* **2021**, *52*, 465–478. [[CrossRef](#)]
97. Nawaratne, R.; Alahakoon, D.; de Silva, D.; Yu, X. Spatiotemporal Anomaly Detection Using Deep Learning for Real-Time Video Surveillance. *IEEE Trans. Ind. Inform.* **2020**, *16*, 393–402. [[CrossRef](#)]
98. Nguyen, V.-T.; Nguyen, A.-T.; Nguyen, V.-T.; Bui, H.-A. A Real-Time Human Tracking System Using Convolutional Neural Network and Particle Filter. In Proceedings of the International Conference on Intelligent Systems & Networks, Hanoi, Vietnam, 19–20 March 2021; Springer: Singapore, 2021; pp. 411–417.
99. Abdullah, F.; Ghadi, Y.Y.; Gochoo, M.; Jalal, A.; Kim, K. Multi-Person Tracking and Crowd Behavior Detection via Particles Gradient Motion Descriptor and Improved Entropy Classifier. *Entropy* **2021**, *23*, 628. [[CrossRef](#)] [[PubMed](#)]
100. Bouhleb, F.; Mliki, H.; Hammami, M. Crowd Behavior Analysis Based on Convolutional Neural Network: Social Distancing Control COVID-19. In Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Vienna, Austria, 8–10 February 2021; SciTePress—Science and Technology Publications: Setúbal, Portugal, 2021; pp. 273–280.
101. Pouw, C.; Willems, J.; van Schadewijk, F.; Thurau, J.; Toschi, F.; Corbetta, A. Benchmarking High-Fidelity Pedestrian Tracking Systems for Research, Real-Time Monitoring and Crowd Control. *Collect. Dyn.* **2022**, *6*, A134. [[CrossRef](#)]
102. Boltz, M.; Adrian, J.; Raytarowski, A.-K. A Hybrid Tracking System of Full-Body Motion Inside Crowds. *Sensors* **2021**, *21*, 2108. [[CrossRef](#)]
103. Priyadharsini, N.K.; Chitra, D. A Kernel Support Vector Machine Based Anomaly Detection Using Spatio-Temporal Motion Pattern Models in Extremely Crowded Scenes. *J. Ambient. Intell. Humaniz. Comput.* **2021**, *12*, 5225–5234. [[CrossRef](#)]
104. Jain, N.; Bansal, H. Anomaly Detection in Crowded Places: Review. In Proceedings of the 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 3–4 September 2021; IEEE: Piscataway Township, NJ, USA, 2021; pp. 1–6.
105. Azimjonov, J.; Özmen, A. A Real-Time Vehicle Detection and a Novel Vehicle Tracking Systems for Estimating and Monitoring Traffic Flow on Highways. *Adv. Eng. Inform.* **2021**, *50*, 101393. [[CrossRef](#)]
106. Ma, C.; Yang, F.; Li, Y.; Jia, H.; Xie, X.; Gao, W. Deep Trajectory Post-Processing and Position Projection for Single & Multiple Camera Multiple Object Tracking. *Int. J. Comput. Vis.* **2021**, *129*, 3255–3278. [[CrossRef](#)]
107. Sonkar, R.; Rathod, S.; Jadhav, R.; Patil, D. Crowd Abnormal behaviour detection using deep learning. *Itm Web Conf.* **2020**, *32*, 03040. [[CrossRef](#)]
108. Saba, T. Real Time Anomalies Detection in Crowd Using Convolutional Long Short-Term Memory Network. *J. Inf. Sci.* **2021**. [[CrossRef](#)]
109. Direkoglu, C. Abnormal Crowd Behavior Detection Using Motion Information Images and Convolutional Neural Networks. *IEEE Access* **2020**, *8*, 80408–80416. [[CrossRef](#)]
110. Xia, L.; Li, Z. A New Method of Abnormal Behavior Detection Using LSTM Network with Temporal Attention Mechanism. *J. Supercomput.* **2021**, *77*, 3223–3241. [[CrossRef](#)]
111. Traore, A.; Akhloufi, M.A. Violence Detection in Videos Using Deep Recurrent and Convolutional Neural Networks. In Proceedings of the IEEE Transactions on Systems, Man, and Cybernetics: Systems, Toronto, ON, Canada, 11–14 October 2020; Institute of Electrical and Electronics Engineers Inc.: Piscataway Township, NJ, USA, 2020; Volume 2020, pp. 154–159.
112. Al-Dhamari, A.; Sudirman, R.; Mahmood, N.H. Transfer Deep Learning along with Binary Support Vector Machine for Abnormal Behavior Detection. *IEEE Access* **2020**, *8*, 61085–61095. [[CrossRef](#)]
113. Sabokrou, M.; Fayyaz, M.; Fathy, M.; Moayed, Z.; Klette, R. Deep-Anomaly: Fully Convolutional Neural Network for Fast Anomaly Detection in Crowded Scenes. *Comput. Vis. Image Underst.* **2018**, *172*, 88–97. [[CrossRef](#)]
114. Khan, G.; Farooq, M.A.; Hussain, J.; Tariq, Z.; Khan, M.U.G. Categorization of Crowd Varieties Using Deep Concurrent Convolution Neural Network. In Proceedings of the 2019 2nd International Conference on Advancements in Computational Sciences, ICACS 2019, Lahore, Pakistan, 18–20 February 2019; Institute of Electrical and Electronics Engineers Inc.: Piscataway Township, NJ, USA, 2019.
115. Sikdar, A.; Chowdhury, A.S. An Adaptive Training-Less Framework for Anomaly Detection in Crowd Scenes. *Neurocomputing* **2020**, *415*, 317–331. [[CrossRef](#)]
116. Zhang, Q.; Lin, W.; Chan, A.B. Cross-View Cross-Scene Multi-View Crowd Counting. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 557–567.
117. Zhu, P.; Peng, T.; Du, D.; Yu, H.; Zhang, L.; Hu, Q. Graph Regularized Flow Attention Network for Video Animal Counting From Drones. *IEEE Trans. Image Process.* **2021**, *30*, 5339–5351. [[CrossRef](#)]
118. Idrees, H.; Tayyab, M.; Athrey, K.; Zhang, D.; Al-Maadeed, S.; Rajpoot, N.; Shah, M. Composition Loss for Counting, Density Map Estimation and Localization in Dense Crowds. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 532–546.

119. Wang, Q.; Gao, J.; Lin, W.; Li, X. NWPU-Crowd: A Large-Scale Benchmark for Crowd Counting and Localization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 2141–2149. [[CrossRef](#)]
120. Bahmanyar, R.; Vig, E.; Reinartz, P. MRCNet: Crowd Counting and Density Map Estimation in Aerial and Ground Imagery. *arXiv* 2019, arXiv:1909.12743.
121. Sindagi, V.A.; Yasarla, R.; Patel, V.M. Pushing the Frontiers of Unconstrained Crowd Counting: New Dataset and Benchmark Method. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 1221–1231.
122. Sindagi, V.A.; Member, S.; Yasarla, R.; Patel, V.M.; Member, S. JHU-CROWD++: Large-Scale Crowd Counting Dataset and A Benchmark Method. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *44*, 2594–2609. [[CrossRef](#)] [[PubMed](#)]
123. Schröder, G.; Senst, T.; Bochinski, E.; Sikora, T. Optical Flow Dataset and Benchmark for Visual Crowd Analysis. In Proceedings of the 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Auckland, New Zealand, 27–30 November 2018; IEEE: Piscataway Township, NJ, USA, 2018; pp. 1–6.
124. Peng, D.; Sun, Z.; Chen, Z.; Cai, Z.; Xie, L.; Jin, L. Detecting Heads Using Feature Refine Net and Cascaded Multi-Scale Architecture. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; IEEE: Piscataway Township, NJ, USA, 2018; pp. 2528–2533.
125. Zhang, L.; Youtu, T.; Shi, M.; Chen, Q. Crowd Counting via Scale-Adaptive Convolutional Neural Network. In Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018; IEEE: Piscataway Township, NJ, USA, 2018; pp. 1113–1121.
126. Marsden, M.; Mcguinness, K.; Little, S.; O’connor, N.E. ResnetCrowd: A Residual Deep Learning Architecture for Crowd Counting, Violent Behaviour Detection and Crowd Density Level Classification. In Proceedings of the 2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS), Lecce, Italy, 29 August–1 September 2017; IEEE: Piscataway Township, NJ, USA, 2017; pp. 1–7.
127. Zhang, Y.; Zhou, D.; Chen, S.; Gao, S.; Ma, Y. Single-Image Crowd Counting via Multi-Column Convolutional Neural Network. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway Township, NJ, USA, 2016; pp. 589–597.
128. Zhang, C.; Li, H.; Wang, X.; Yang, X. Cross-Scene Crowd Counting via Deep Convolutional Neural Networks. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; IEEE: Piscataway Township, NJ, USA, 2015; pp. 833–841.
129. Shao, J.; Kang, K.; Loy, C.C.; Wang, X. Deeply Learned Attributes for Crowded Scene Understanding. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; IEEE: Piscataway Township, NJ, USA, 2015; pp. 4657–4666.
130. Conigliaro, D.; Rota, P.; Setti, F.; Bassetti, C.; Conci, N.; Sebe, N.; Cristani, M. The S-HOCK Dataset: Analyzing Crowds at the Stadium. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; IEEE: Piscataway Township, NJ, USA, 2015; pp. 2039–2047.
131. Shao, J.; Loy, C.C.; Wang, X. Scene-Independent Group Profiling in Crowd. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; IEEE: Piscataway Township, NJ, USA, 2014; pp. 2227–2234.
132. Jiang, M.; Xu, J.; Zhao, Q. Saliency in Crowd. In Proceedings of the European conference on computer vision, Zurich, Switzerland, 6–12 September 2014; Springer: Cham, Switzerland, 2014; pp. 17–32.
133. Idrees, H.; Saleemi, I.; Seibert, C.; Shah, M. Multi-Source Multi-Scale Counting in Extremely Dense Crowd Images. In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; IEEE: Piscataway Township, NJ, USA, 2013; pp. 2547–2554.
134. Allain, P.; Courty, N.; Corpetti, T. AGORASET: A Dataset for Crowd Video Analysis. In Proceedings of the 1st ICPR international workshop on pattern recognition and crowd analysis, Tokyo, Japan, 11 November 2012; pp. 1–6.
135. Hassner, T.; Itcher, Y.; Kliper-Gross, O. Violent Flows: Real-Time Detection of Violent Crowd Behavior. In Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Providence, RI, USA, 16–21 June 2012; IEEE: Piscataway Township, NJ, USA, 2012; pp. 1–6.
136. Loy, C.C.; Gong, S.; Xiang, T. From Semi-Supervised to Transfer Counting of Crowds. In Proceedings of the 2013 IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013; IEEE: Piscataway Township, NJ, USA, 2013; pp. 2256–2263.
137. Zhou, B.; Wang, X.; Tang, X. Understanding Collective Crowd Behaviors: Learning a Mixture Model of Dynamic Pedestrian-Agents. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; IEEE: Piscataway Township, NJ, USA, 2012; pp. 2871–2878.
138. Ferryman, J.; Shahrokni, A. PETS2009: Dataset and Challenge. In Proceedings of the 2009 Twelfth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, Snowbird, UT, USA, 7–12 December 2009; IEEE: Piscataway Township, NJ, USA, 2009; pp. 1–6.

139. Mehran, R.; Oyama, A.; Shah, M. Abnormal Crowd Behavior Detection Using Social Force Model. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; IEEE: Piscataway Township, NJ, USA, 2009; pp. 935–942.
140. Chan, A.B.; Liang, Z.S.J.; Vasconcelos, N. Privacy Preserving Crowd Monitoring: Counting People without People Models or Tracking. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008; IEEE: Piscataway Township, NJ, USA, 2008; pp. 1–7.