A Survey on Video Anomaly Detection

Rajesh Kumar Yadav
Computer Science and Engineering
Delhi Technological University
Delhi, India

Rajiv Kumar
Computer Science and Engineering
Delhi Technological University
Delhi, India

***Abstract*—The number of surveillance cameras has increased considerably over the last decade, and so is the research in order to reduce the human intervention in surveillance in order to automate the process. Because human-based monitoring is time-consuming and difficult, there is an increasing demand for autonomous systems to detect abnormalities in crowded locations. Deep learning has shown to be a game-changing computing technology in the realm of computer vision.As a result, it's routinely utilised to perform complex cognitive tasks like spotting abnormalities in surveillance footage. Deep learning anomaly detection technologies beat conventional machine learning systems. Our research seeks to provide a complete analysis of video anomaly detection systems that employ deep learning that have been published since 2019.In this paper, we examined the learning methods as well as the dataset used for training and testing. Along with it, the model's accuracy is described.In addition, this work gives a brief review of video datasets for anomaly detection. Furthermore, present and emerging trends are emphasised.**

Keywords*—*Anomaly Detection, Surveillance Videos, Deep Learning

I. Introduction

Security cameras have become progressively common over the last few years to increase public safety.It takes a lot of work to keep a watch on the video footage collected by surveillance cameras in order to spot anomalies, which is why anomalies are not always recognised in a timely manner [1]. As the number of security cameras increases, more human observers are required and hence this scenario has created a key barrier in the exploitation of surveillance material. Because abnormal incidents are exceedingly rare, the monitoring effort needs complete focus. As a result, human monitors may miss security breaches. The technique of detecting anomalies in surveillance video is known as anomaly detection. In the context of security surveillance, security breaches or threats are classified as anomalies. Only a few among the most prevalent abnormalities include rioting, attack, gunfire, violence, explosions, abuse, theft, arming, stealing, shoplifting and graffiti[2]. Anomaly detection may be achieved using machine learning methods, which can be deep learning or non-deep learning approaches. Non-deep learning methods can be used both unsupervised and supervised. Labeled data is utilised for machine training in supervised learning systems. One example is teaching a child to recognise different creatures, such as a cat or a dog.For video classification, several supervised learning techniques which include Multiclass Support vector machine can be utilised. Non-labeled data is utilised for machine training in unsupervised learning approaches. It employs a number of clustering methodologies, including the use of the mean shift algorithm to give mob movement clustering in the input video sequence, to categorise on the basis of shifting velocity and location.The moving optical flow field is then further classified using the k-means approach, which combines small group categories with related motion parameters. Anomaly is considered in such algorithms only when more than k frames on a continuous basis have a clustering centre in an unexpected place.

This study looked at the deep learning models that have been published since 2019. Deep learning techniques for detecting abnormalities in surveillance videos were the only ones studied. Many papers were discussed and examined for a more detailed study.In order to compare, datasets utilised for training the model as well as for testing the model along with its accuracy and learning approaches are studied. Deep learning algorithms that are at the heart of the problem are determined and reviewed. Following anomaly detection, we go on to performance evaluation, for which we apply numerous curves, including the AUC-ROC (Area under curve - Receiver Operating Characteristic) curve. The curve that is obtained by AUC-ROC is used to monitor the efficiency of multi label classification problem at different threshold levels. AUC represents the degree of distinction and has the capability to distinguish groups in this scenario. A greater AUC means that it is capable of classifying anomalies more accurately. This curve depicts four possible results -

1) *Sensitivity / True positive rate*: It informs us what fraction of the true positive class has been appropriately categorised.
2) *False positive rate* : It informs us what fraction of the true negative class has been erroneously categorised.
3) *Specificity / True negative rate*:It tells us what fraction of the negative class got accurately classified.
4) *False negative rate* : It shows us how much of the positive class was misclassified by that of the algorithm.

978-1-6654-5883-2/22/$31.00 ©2022 IEEE

2022 IEEE Delhi Section Conference (DELCON) | 978-1-6654-5883-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/DELCON54057.2022.9753580

The equal error rate is defined as the intersection of true positive rate and the false positive rate (EER). The better the outcome, the lower the EER value.

This paper's structure is as follows: Section II discusses earlier articles on video anomaly detection. This research is compared to other surveys, as well as the gaps that have been identified and the contributions that have been made. Section III discusses the various datasets for identifying video abnormalities. Section IV investigates the researchers' various models and gives a tabular comparison of accuracies based on various datasets. It also examines evaluation criteria and trends. The use of deep learning models and their various techniques that are used to detect anomaly in surveillance videos is examined. Section V closes with a conclusion and projections for the future.

II. Literature Survey

Despite significant advances in deep learning approaches for several machine learning applications, deep learning approaches for anomaly detection remain rather limited. Chalapathy et al. (2019) undertake a survey in which they develop and assess numerous organised and comprehensive deep anomaly detection (DAD) methods[3]. There are various research and reviews on the issue since anomaly detection is strongly connected to outlier identification and novelty detection [4,5,6,7,8].

Jing Liu and Peng Wu[9] offer a four-module technique for exploiting temporal cue and feature discriminating influence.To improve features,local-range temporal correlations among variables are collected by the causal temporal relation (CTR). By using causal convolution, the classifier (CL) expands the temporal modelling range by projecting improved properties to the category space. The compactness (CP) and dispersion (DP) modules are intended to learn the discriminative power of features, with the compactness module assuring intraclass compactness of normal features and the dispersion module enhancing interclass dispersion.

Trajectories are also used to detect irregularities. The system [10] recognises anomalies in videos by learning consistency in skeletal motions. Tran et al. (2019) propose employing 2D human skeleton motions to detect anomalous human behaviour in surveillance footage.

Che Sun and others [11] describe a scene-aware context reasoning approach for unsupervised unusual event recognition in clips that uses context information from visual features to overcome the semantic gap among visual context and abnormal occurrence meaning. The efficacy of their technique is proved by assessments on three difficult datasets: UCF-Crime, Avenue, and ShanghaiTech.

J. Zhang and others [16] cited Multiple instance learning is constructed for semisupervised outlier detection, and a new inner bag loss is proposed to constrain the function space of the semisupervised issue by taking the minimum anomalous instance value and the greatest value in each case into consideration. In other words, the difference between the lowest and highest scores in a positive bag should be large, but in a negative bag, it should be trivial. The results of experiments on the Crime dataset reveal that a temporal convolutional network with complimentary inner bag loss exceeds the state-of-the-art.

Boyang Wan and colleagues 2020 [22] considered video anomaly detection to be a regression problem in terms of anomaly scores on video clips under inadequate supervision. As a result, they provide Anomaly Regression Net (ARNet), an anomaly detection system that requires just video-level labels during the training phase. In order to train discriminant features for anomaly detection for the proposed AR-Net, they also create a dynamic diverse learning loss and a centre loss. The former is proposed to increase the inter-class contrast between abnormal and normal cases, whilst the latter is proposed to reduce intra-class differentiation among normal examples. Extensive testing is being done on a challenging benchmark: ShanghaiTech. On the ShanghaiTech dataset, their method yields a new state-of-the-art performance for anomalous behavior detection in surveillance videos.

III. Widely utilised video anomaly datasets

In this section, we will look at the most notable aberrant video datasets that have been generated. The datasets we'll be discussing are the product of a variety of research challenges as well as experiments carried out by researchers in efforts to answer actual problems in this industry. The similarities between these datasets may be noticed in how they are used in a number of scenarios. Because of the availability of security cameras installed in various locations, the researchers produced a plethora of anomaly datasets. The large number of the erroneous video recordings are accessible to the general public for study purposes. Most individuals are unfamiliar with a few datasets since they were used in one or two scholarly studies. Such databases are deficient in information and are not easily accessible for research purposes. As a result, the community of computer vision researchers is still in the dark. Despite the fact that they are simple and have many restrictions, only a few datasets are widely used. This is due to the fact that a description of the dataset is supplied, that it is published in a prestigious conference, and that it is made freely available on their webpages. Researchers frequently use these databases to compare their conclusions to earlier findings published elsewhere. As a result, significant research progress may be made by utilising such datasets in a given application that is lacking in others. As a result, comprehension of the numerous anomalous video datasets accessible for investigation is necessary. This can help with research balance by concentrating on various aspects of the very same subject and solving challenges. The datasets mentioned and discussed below give information on a wide range of anomaly datasets that are available to the scientific community. Fig.1 shows several example frames from multiple anomalous video datasets.

UCSD: The UCSD anomaly detection dataset [12-13] is composed of surveillance video collected by a camera mounted in two separate scenarios of a busy pedestrian path: Peds. 1 and Peds. 2. It includes both conventional and unusual

activities on the paths, such as wandering pedestrians and the motion of motorcyclists, skaters, bicyclists, tiny carts, persons in wheelchairs, and so on. A pedestrian movement at an unexpected area is frequently referred to as an odd incident. The Peds 1 data set is made up of 34 and 26 training and testing video clips respectively. The Peds 2 dataset contains 16 and 12 training and testing video clips respectively. There is also some perspective distortion in the Peds 1 dataset. The UCSD dataset includes real data at the frame and pixel levels. The pixel-level ground truth serves in determining the accuracy of localisation in cutting-edge anomaly detection systems.

UMN Dataset: This dataset includes instances involving panic or crowd-escape [14]. The UMN dataset is separated into three 320240 resolution scenarios: grass (1450 frames), interior (4415 frames), and plaza (2145 frames). There are two competitions: walking and running (escape). Individuals travelling in different paths is a common phenomenon. These frames are used to select samples for training and normal testing. The hurrying individuals make for an interesting spectacle. These frames are used to locate and retrieve erroneous testing samples. The ground truth is included inside the video frames that must be recovered in order to evaluate the quality.

UCF Crime Dataset: There are 128 hours of video footage in the UCF-Crime collection. It is made up of 1900 hours of unedited authentic surveillance video. It is distinguished by extended unedited surveillance tapes containing 13 real-world anomalies such as Abuse, Prosecution, Property damage, Attack, Road Accident, Robbery, Explosion, Violence, Thievery, Killing, Theft, Petty theft, and Damaging Property. These anomalies were picked because they have a significant impact on societal safety.

Shanghai Tech Dataset: The ShanghaiTech Campus dataset includes 13 different themes with variable illumination and camera angles. It consists of around 270, 000 training frames and 130 anomalous events. Furthermore, the pixel level ground truth of anomalous occurrences is noted in this dataset. It is better if the trained anomaly detection model can be deployed in a variety of settings with varying view angles right away. However, almost all accessible datasets only include movies filmed with a single fixed angle camera, resulting in a lack of scene and view angle variety. It adds anomalies caused by rapid movements in this dataset, such as chases and fist fights, that do not exist in other datasets. Due to these characteristics, this dataset is more suited for usage in practical uses.

Avenue Dataset: The Avenue dataset [15] was built to document activity on CUHK Campus Avenue.It is made up of 16 training and 21 testing short films totaling 30652 frames, with 15328 frames dedicated to training and the remainder to testing. In the training films, normal settings are presented. Testing film depicts both normal and atypical occurrences. To add additional intricacy, a little camera shaking test video frame is supplied. Ground truth is offered for anomalous events marked by rectangles. The evaluation may be done at the frame and pixel levels.

*XD Dataset:*XD-Violence contains 217 hours of films, including 4754 unedited clips with audio signals and inadequate labelling. It is done on a large scale, which is beneficial for developing generalizable algorithms for recognizing violence. It encompasses a wide range of scenarios, enabling violence monitoring systems to dynamically evolve to diverse and complex environments, hence increasing their robustness. It also includes audio signals, allowing computers to utilise multimodal data with greater precision. Because of the frequency of violent events, each violent video is labelled with numerous violent labels (1 labels 3). The order of labels for each video reflects the significance of numerous violent episodes in the film.

Fig. 1 Sample images from various anomaly datasets Source:Adapted from [40]

IV. Evaluation Metrics And Trends

A. Evaluation Metrics

This section focuses on the evaluation methods that were most frequently utilised during the evaluation. In the majority of models, the Receiver Operating Characteristic Curve (ROC) and its related Areas Under the Curve (AUC) have been utilised (AUC). The ROC curve represents successful True Positives vs False Positives [30]. This statistic evaluates the specificity and accuracy of the model.

Many articles used the equal Error rate (ERR) metric to examine the incidence of mistakes in model predictions. It defines the criterion for balancing Falsified Acceptance rate along with the False Rejection [30]. If the ERR number is small, the model's accuracy is regarded high, and vice versa.

In addition to ERR, F1 Score is another method for determining the model's accuracy. This statistic is used to assess the precision of binary categorization. It has been optimised by researchers to check the validity of outlier detection, whereas outlier detection has been optimised as two class classification at the object level.It is computed by taking the ratio of product of recall and precision and the total sum of the recall and precision.In contrast to the Receiver Characteristic Operator curve, it takes into account outcomes that are both false positives and false negatives [30]. The F1 score typically varies between 0 and 1. The greater the magnitude of the F1 number, the greater the accuracy of the model.

TABLE I: Performance Analysis on UCF-Crime Dataset

Publication	Reported on	Supervision	AUC
J. Zhang, L. Qing and J. Miao 2019 [16]	ICIP 19	Weakly	78.66
Jia-Xing Zhong and others 2019 [17]	CVPR 19	Weakly	82.12
Kun Liu and Huadong Ma. 2019 [18]	ACM MM 19	Fully	82.0
Zaigham Zaheer and others 2020 [19]	ECCV 20	Weakly	83.03
Jia-Chang Feng and others 2021[20]	CVPR 21	Weakly	82.30
P. Wu and J. Liu 2021[9]	TIP 21	Weakly	84.89

TABLE II: Performance Analysis on SHANGHAI Tech Dataset

Publication	Reported on	Supervision	AUC
Antonio Bărbălău and others 2021[21]	CVPR 21	Unsupervised	90.2
Boyang Wan and others 2020 [22]	ICME 2020	Weakly	86.3
Jia-Xing Zhong and others 2019 [17]	CVPR 19	Weakly	84.44
Zaigham Zaheer and others 2020 [19]	ECCV 20	Weakly	89.67
Tian, Yu et al 2021[23]	ICCV 21	Weakly	97.21
P. Wu and J. Liu 2021[9]	TIP 21	Weakly	97.48

B. Trends

It should be noted that the bulk of the research evaluated used transfer learning and autoencoders deep learning methodologies. The ability to transfer information has simplified the design and deployment of models.The usefulness of transfer learning has grown as a result of its ability to improve the efficiency of the base method.Researchers have leveraged autoencoders' capacity to train with no or little supervision to build a variety of deep learning models that outperformed other models in the same domain.

Generative adversarial networks can also be used to prevent anomalies before they occur. To anticipate future frames, Generative adversarial networks are trained using multiple datasets that contain various behaviour. The normal/abnormal classification is subsequently applied to subsequent frames.

Because anomalies are highly subjective, establishing patterns will be dependent on the model's ability to continually acquire fresh knowledge. Security cameras must be capable of detecting anomalies in real time and progressively learning distinct observations. As a result, the new approach involves the incorporation of both continuous and reinforcement learning. Currently, only a small amount of work has been completed, and owing to the structure of anomaly detection, additional attention is required to be spent in that area in order to enable realtime anomaly detection and sequential model changes.

TABLE III: Performance Analysis on AVENUE Tech Dataset

Publication	Reported on	Supervision	AUC
T. N. Nguyen and J. Meunier 2019 [24]	ICCV 19	Unsupervised	86.9
Liu, Wen et al. 2019 [25]	ICME 2020	Weakly	89.2
Guang Yu and others 2020 [26]	ACM MM 20	Unsupervised	90.2
Ruichu Cai and others 2021 [27]	AAAI 21	Unsupervised	86.6
Wang, X and others 2021 [28]	TNNLS 21	Unsupervised	88.3
Ziming Wang and others 2020[29]	ACM MM 20	Unsupervised	87.0

V. Conclusion and Future Work

We looked at deep learning approaches used to detect abnormalities in videos in this study. The most extensively used datasets for testing and training models were examined, and the performance of many models working on these datasets was compared. It was also said that in this domain, online decision making is a significant but often overlooked component. Instead of just detecting the crime, more work should be done to prevent it. Predicting anomalies in future frames is one such way; this will allow us to detect abnormalities before they occur. In-depth study and improvements in the field of anomaly detection using predicted frames will be the focus of future research.

References

[1] R. Yadav and M. Rai, "Advanced Intelligent Video Surveillance System (AIVSS): A Future Aspect," Research Gate, 2018.

[2] W. Sultani, C. Chen and M. Shah, "Real-World Anomaly Detection in Surveillance Videos," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6479-6488, 2018.

[3] Chalapathy Raghavendra, Chawla Sanjay (2019)."Deep learning for anomaly detection:A survey", A Preprint, ResearchGate, available online: https://www.researchgate.net/publication/330357393

[4] Chandola V., Banerjee A., and Kumar V.(2009). "Anomaly detection: A survey," ACM Computing Surveys, vol. 41, no. 3.

[5] Omar S.,Ngadi A. and Jebur H. (2013)."Machine learning techniques for anomaly detection: An overview," in International Journal of Computer Applications, vol. 79/2, pp. 33–41.

[6] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko (2014) "A review of novelty detection," Signal Processing, vol. 99, pp. 215–249.

[7] Agrawal S. and Agrawal J. (2015). "Survey on anomaly detection using data mining techniques," Procedia Computer Science, vol. 60, pp. 708–713.

[8] D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim (2017). "A survey of deep learning-based network anomaly detection,"ClusterComputing,vol.9,no.5,p.205.

[9] P. Wu and J. Liu, "Learning Causal Temporal Relation and Feature Discrimination for Anomaly Detection," in IEEE Transactions on Image Processing, vol. 30, pp. 3513-3527, 2021, doi: 10.1109/TIP.2021.3062192.

[10] Tran Truyen, Morais Romero and Venkatesh Svetha (2019). "Learning Regularity in Skeleton Trajectories for Anomaly Detection in Videos", Researchgate, available online: https://www.researchgate.net/publication/33851180

[11] Che Sun, Yunde Jia, Yao Hu, and Yuwei Wu. 2020. Scene-Aware Context Reasoning for Unsupervised Abnormal Event Detection in Videos. In Proceedings of the 28th ACM International Conference on Multimedia (MM '20). Association for Computing Machinery, New York, NY, USA, 184–192. DOI:https://doi.org/10.1145/3394171.3413887

[12] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos, "Anomaly detection in crowded scenes", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recog- nition,2010, pp. 1975–1981.

[13] S. V. C. Lab, "UCSD anomaly data set", 2014. "http://www.svcl. ucs d.edu/projects/anomaly/dataset.html"

[14] UMN, Minneapolis, MN, USA. Unusual Crowd Activity Dataset of University of Minnesota, Department of Computer Science and Engineering, 2006."http://mha.cs.umn.edu/movies/crowdactivity-all.avi"

[15] Lu, Cewu, Jianping Shi, and Jiaya Jia. "Abnormal event detection at 150 fps in matlab." Proceedings of the IEEE International Conference on Computer Vision. 2013. "http://www.cse.cuhk.edu.hk/leojia/projects/detectabnormal/dataset.html"

[16] J. Zhang, L. Qing and J. Miao, "Temporal Convolutional Network with Complementary Inner Bag Loss for Weakly Supervised Anomaly Detection," 2019 IEEE International Conference on Image Processing (ICIP), 2019, pp. 4030-4034, doi: 10.1109/ICIP.2019.8803657.

[17] J. Zhong, N. Li, W. Kong, S. Liu, T. H. Li and G. Li, "Graph Convolutional Label Noise Cleaner: Train a Plug-And-Play Action Classifier for Anomaly Detection," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 1237-1246, doi: 10.1109/CVPR.2019.00133.

[18] Kun Liu and Huadong Ma. 2019. Exploring Background-bias for Anomaly Detection in Surveillance Videos. In Proceedings of the 27th ACM International Conference on Multimedia (MM '19). Association for Computing Machinery, New York, NY, USA, 1490–1499. DOI:https://doi.org/10.1145/3343031.3350998

[19] Zaheer, Zaigham Mahmood, Arif Astrid, Marcella Lee, Seung-Ik. (2020). CLAWS: Clustering Assisted Weakly Supervised Learning with Normalcy Suppression for Anomalous Event Detection.

[20] J. -C. Feng, F. -T. Hong and W. -S. Zheng, "MIST: Multiple Instance Self-Training Framework for Video Anomaly Detection," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 14004-14013, doi: 10.1109/CVPR46437.2021.01379.

[21] M. -I. Georgescu, A. Bărbălău, R. T. Ionescu, F. Shahbaz Khan, M. Popescu and M. Shah, "Anomaly Detection in Video via Self-Supervised and Multi-Task Learning," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 12737-12747, doi: 10.1109/CVPR46437.2021.01255.

[22] B. Wan, Y. Fang, X. Xia and J. Mei, "Weakly Supervised Video Anomaly Detection via Center-Guided Discriminative Learning," 2020 IEEE International Conference on Multimedia and Expo (ICME), 2020, pp. 1-6, doi: 10.1109/ICME46284.2020.9102722.

[23] Tian, Yu et al. "Weakly-supervised Video Anomaly Detection with Contrastive Learning of Long and Short-range Temporal Features." ArXiv abs/2101.10030 (2021): n. Pag.

[24] T. N. Nguyen and J. Meunier, "Anomaly Detection in Video Sequence With Appearance-Motion Correspondence," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 1273-1283, doi: 10.1109/ICCV.2019.00136.

[25] Liu, W., Luo, W., Li, Z., Zhao, P., Gao, S. (2019). Margin Learning Embedded Prediction for Video Anomaly Detection with A Few Anomalies. IJCAI.

[26] Yu, Guang Wang, Siqi Cai, Zhiping Zhu, En Xu, Chuanfu Yin, Jianping Kloft, Marius. (2020). Cloze Test Helps: Effective Video Anomaly Detection via Learning to Complete Video Events. 10.1145/3394171.3413973.

[27] Ruichu Cai, Hao Zhang, Wen Liu, Shenghua Gao, Zhifeng Hao: Appearance-Motion Memory Consistency Network for Video Anomaly Detection. AAAI 2021: 938-946

[28] Wang, X., Che, Z., Yang, K., Jiang, B., Tang, J., Ye, J., Wang, J., Qi, Q. (2021). Robust Unsupervised Video Anomaly Detection by Multi-Path Frame Prediction. IEEE transactions on neural networks and learning systems, PP.

[29] Ziming Wang, Yuexian Zou, and Zeming Zhang. 2020. Cluster Attention Contrast for Video Anomaly Detection. In Proceedings of the 28th ACM International Conference on Multimedia. Association for Computing Machinery, New York, NY, USA, 2463–2471.

[30] T. Kanstren, "A Look at Precision, Recall, and F1-Score," Towards Data Science, 09 September 2012. [Online]. Available: https://towardsdatascience.com/a-look-at-precision-recall-and-f1-score-36b5fd0dd3ec.

[31] A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz, "Robust real-time unusual event detection using multiple fixed-location monitors," IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 3, pp. 555–560, Mar. 2008.

[32] Y. Cong, J. Yuan, and Y. Tang, "Video anomaly search in crowded scenes via spatio-temporal motion context," IEEE Trans. Inf. Forensics Secur., vol. 8, no. 10, pp. 1590–1599, 2013.

[33] R. Leyva, V. Sanchez, and C. T. Li, "Abnormal event detection in videos using binary features," 2017 40th Int. Conf. Telecommun. Signal Process. TSP 2017, vol. 2017-January, pp. 621–625, Oct. 2017.

[34] R. V. H. M. Colque, C. A. C. Junior, and W. R. Schwartz, "Histograms of Optical Flow Orientation and Magnitude to Detect Anomalous Events in Videos," Brazilian Symp. Comput. Graph. Image Process., vol. 2015-October, pp. 126–133, Oct. 2015.

[35] R. Chaudhry, A. Ravichandran, G. Hager, and R. Vidal, "Histograms of oriented optical flow and Binet-Cauchy kernels on nonlinear dynamical systems for the recognition of human actions," pp. 1932– 1939, Mar. 2010.

[36] H. Vu, T. D. Nguyen, A. Travers, S. Venkatesh, and D. Phung, "Energy-based localized anomaly detection in video surveillance," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 10234 LNAI, pp. 641–653, 2017.

[37] R. T. Ionescu, S. Smeureanu, B. Alexe, and M. Popescu, "Unmasking the abnormal events in video," Proc. IEEE Int. Conf. Comput. Vis., vol. 2017-October, pp. 2914–2922, May 2017.

[38] D. G. Lee, H. Il Suk, S. K. Park, and S. W. Lee, "Motion Influence Map for Unusual Human Activity Detection and Localization in Crowded Scenes," IEEE Trans. Circuits Syst. Video Technol., vol. 25, no. 10, pp. 1612–1623, Oct. 2015.

[39] N. Jain and H. Bansal, "Anomaly Detection in Crowded Places: Review," 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2021, pp. 1-6, doi: 10.1109/ICRITO51393.2021.9596229.

[40] N. Patil and P. K. Biswas, "A survey of video datasets for anomaly detection in automated surveillance," 2016 Sixth International Symposium on Embedded Computing and System Design (ISED), 2016, pp. 43-48, doi: 10.1109/ISED.2016.7977052.

[41] J. G. Munyua, G. M. Wambugu, and S. T. Njenga, "A Survey of Deep Learning Solutions for Anomaly Detection in Surveillance Videos", IJCIT, vol. 10, no. 5, Oct. 2021.

[42] Che Sun, Yunde Jia, Yao Hu, and Yuwei Wu. 2020. Scene-Aware Context Reasoning for Unsupervised Abnormal Event Detection in Videos. In Proceedings of the 28th ACM International Conference on Multimedia (MM '20). Association for Computing Machinery, New York, NY, USA, 184–192. DOI:https://doi.org/10.1145/3394171.3413887

[43] M. Emad, M. Ishack, M. Ahmed, M. Osama, M. Salah and G. Khoriba, "Early-Anomaly Prediction in Surveillance Cameras for Security Applications," 2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), 2021, pp. 124-128, doi: 10.1109/MIUCC52538.2021.9447668.

[44] Amraee, S., Vafaei, A., Jamshidi, K. et al. Abnormal event detection in crowded scenes using one-class SVM. SIViP 12, 1115–1123 (2018). https://doi.org/10.1007/s11760-018-1267-z