Detecting Criminal Activities From CCTV by using Object Detection and machine Learning Algorithms

Surbhi Singla
CSE AIML
University Institute of Engineering
Chandigarh University,
Mohali, India
surbhisngl06@gmail.com

Dr. Raman Chadha
Professor and Research Coordinator, CSE
University Institute of Engineering
Chandigarh University,
Mohali, India
raman.e11212@cumail.in

Abstract - **Now, a day's Crime in every country is increasing day by day. Generally, Every day we listen to the news of different crimes of different categories like rape, assault, Kidnapping ,Robbery ,ATM Theft, Murders etc happening in different states, cities , countries. Almost all the newspapers, TV channels', social media are filled with the news of Criminal activities happening all around the Whole World. In earlier times there is no method to detect Crime. After That the CCTV cameras were used to detect Crimes. But Watching these Videos manually by humans for detecting crimes is a very time Consuming process especially in today's world of Artificial Intelligence and Machine learning .So this crime detection in CCTV surveillance becomes an important area of research in the field of machine learning. So, there is a very urgent need of the intelligent system which will detect the crimes from the real time CCTV Feed and classify them and provides an alert system to the nearest police stations and ambulances etc. So, that system will help in reducing the crime rate in any country. This paper reviews all prior research in this area, including approaches for object recognition and finding priority frames, techniques and algorithms like Yolo used to detect crimes , various datasets used and algorithms used to analyze crime data and train the dataset .It covers the various recent trends in researches in this field and analyzing the challenges faced and various research gaps and this paper also discuss how we can overcome these gaps in research so as to develop a better intelligence surveillance system in ml field.**

Keywords: Artificial Intelligence, Machine learning, Crime, CCTV, Object detection, Anomaly detection, intelligent surveillance system , Yolo, Deep Learning, Nueral networks.

I. INTRODUCTION

Artificial Intelligence and Machine Learning are associated with human Intelligence and natural intelligence [1]. Intelligence means the ability to think, learn and understand something [2]. Human intelligence is the ability of the humans to think creative, understands something, learn from concepts and solve problems. Artificial intelligence is the ability of the machines to think creative, understand and analyze something or in other words ability of the machines to perform tasks like humans and think just like a human[1] does. Artificial intelligence uses machine learning algorithms in python to train the systems and machines to think and perform like humans. Machine Learning is a branch of AI. Machine learning is the ability of the systems or Machines to automatically learn and improve from past experiences without being explicitly programmed [2]. For Example alexa and Siri are AI based systems which we use in our daily based life. Other examples include google home and self driving tesla cars [3]. The main purpose of AI and ml based applications and systems is to develop various applications and intelligent systems that reduce human efforts and time by automating the system. AI and Machine based systems are almost developed in every sector such as aviation, healthcare, transportation, education, medical, electronic trading, remote sensing, education, banking and finance [4], crime and Object detection. But In this review paper we are going to discuss only about the Crime detection and object detection methods and the approach to develop the intelligent video surveillance system. But before discussing Crime and object detection methods firstly we will discuss why this topic is important for research and why we need to develop this intelligent surveillance system.

From past few years Crime and violence rate has increased all over the country in various categories like ATM Robbery , Murders, rape , assault etc specially the crime against women has been increased. In some states the crime has increased a lot like UP, Delhi, Assam etc. The states with the highest rates of violent crime are Bihar, Jharkhand, Odisha, West Bengal, Assam, Tripura, Arunachal Pradesh, Delhi, Haryana, Maharashtra, and Madhya Pradesh [5]. In 2022, UP has the highest rape and murders rate [6]. The per capita crime rate of UP is 7.4[6]. According to the National Crime Records Bureau, this ratio means that Uttar Pradesh has the highest number of crimes, and thus, the state is unsafe to travel alone[6]. So, in order to address these concerns and to reduce crime to some extent now days CCTV Cameras are now installed in almost every Crime prone area.

A. Importance of CCTV Cameras

CCTV does an excellent job in preventing and reducing criminal activity throughout the economy. To begin with, the cameras give evidence of crime, making it easier to track down culprits. Previously, proving crimes was one of the most difficult tasks for judges in many countries, and as a result, many criminals went unpunished, encouraging more crime. The cctv can not only assist in the detection of criminals, but it can also assist in the detection of missing children and persons from CCTV footage available at various locations such as railway stations [6].

Alarm notification is the most effective approach for CCTV cameras to prevent and reduce crime. According to Guh (2002), the areas covered in most establishments where surveillance security is used are explicit and clear ; security staff can secure the main item or focus area by attaching it to an alert notification system in the event of criminal activity [6]. In this type of security, the activities and movements in a certain area must fit into a schedule, with any deviation triggering an alarm Because any effort to break the glass sets off the alarm and crime is detected.

The psychological impact of surveillance is the most important stage in crime reduction and prevention. CCTV surveillance serves to remind people of their security precautions, which they would otherwise ignore most of the time. People understand that if security was not present, an area under monitoring would appear insecure. In some circumstances, persons travelling, working, take extra precautions to secure their property by installing CCTV in their houses. Robbery on moving trains in Brazil was [7] considered so common before the trains began to utilize CCTV cameras because criminals thought they had no recourse [7].

Most residents of insecure locations fear they could be attacked at any time .In order to dissuade crime, CCTV cameras are used to create fear among the public. The presence of monitoring services informs the public that these regions are more secure than those that are not under cctv coverage, hence more people visit the protected areas than those that are not .When the population of a place grows, criminal activity[8] decreases because there are more witnesses in the event.

B. Disadvantages of Manual Crime Detection System

The irony is that some factors actually make the cameras less useful. Several factors prevent high levels of crime, including surveillance. Following a crime scene or important incident, it takes a while to identify culprits or individuals in CCTV footage. Members of the Goa branch's cyber cell claim that since there is no automatic method for accomplishing this, they make several department employees to sit down with laptops and computers and really sift through CCTV tape to find and identify the guilty[8]. As a result, the current system is labor- and time-intensive. The solutions currently available rely on conventional facial recognition algorithms[7],which can be problematic in India due to conditions such as light, weather, and, most importantly. Some CCTV cameras are installed in inconvenient locations and might get slanted, resulting in a increase in inaccuracy [8]. Also if the CCTV feed is not clear in quality i.e. Resolution then it is of no use and its it will give inaccurate results.

So, there is a very urgent need of a intelligent system which will detect the crimes from the real time CCTV Feed and classify them and provides an alert system to the nearest police stations and ambulances[9] etc. So, that system will help in reducing the crime rate in any country. So this topic of research is very important to solve as if we are do some research on this topic and are able to propose a system which can predict crimes. This system basically relies on algorithms

for object detection and actions taken on those items to identify inappropriate human behaviour. Hence, the development of these intelligent video surveillance systems depends on object detection and activity detection [10]   systems. So we studied all the pervious researches in this field and then proposed a system that for developing this intelligent system.

II. LITERATURE SURVEY

In the subject of criminal detection, creating an intelligent surveillance system has been a significant study topic. Many studies have been conducted in this area, ranging from activity recognition to object detection.However, the majority of currently conducted research focuses on special high dimensionality or anomaly detection characteristics. As an example, a project examined various anomaly detection techniques in order to provide a basic understanding of the various approaches to anomaly detection. Numerous graph-based applications [10] in the actual world are examined in another study. To identify unresolved problems and difficulties, a fresh survey of hybrid intrusion detection systems [11] and anomaly detection systems was carried out. There has been some study linking anomaly detection to high dimensionality issues, either directly or indirectly To specifically address specific imaging issues, a number of recent techniques recommend CNN [12].

Deblurring has been suggested for a network that is similar to a one-iteration ODP network and uses a single, studied deconvolution step followed by a CNN. For describing such unforeseen events, such as anomaly in complex scenes[12], a real-time anomaly detection and localization technique [13] was developed in 2015. Our method uses two local and global descriptors to describe each video as a collection of non-overlapping cubic patches (13).. After studying a lot of research papers on it the previous research and studies suggests that the existing system of the automated CCTV monitoring system is not very good at giving decisions and correct responses that are appropriate for the circumstances. Videos produced by CCTV cameras located within the ATM [14] Or if the video required a lot of manual labour to watch it. Just in case there was an ATM heist, it became time-consuming to watch the entire, lengthy movie.

Weixin Luo and Wen Liu[17] suggested Temporally Coherent Sparse Coding (TSC) in one of their papers in 2017, motivated by the capacity to identify sparse data.

The suggested technique uses identical neighbouring frames with comparable [15] reconstruction coefficients to address the coding-based anomaly. The suggested method then uses a particular variety of stacked [16] recurrent neural network to map the TSC (sRNN). By using sRNN to learn all parameters simultaneously[16], it is possible to avoid the nontrivial selection of hyper-parameters for TSC, and The computation required for learning is reduced since the reconstruction coefficients can be deduced inside a forward passage with a shallow sRNN[17]. The CNN model's output effectively extracted suspicious activity frames from a lengthy movie, sending those suspicious frames into a neural network structure[18]. The main goal of this research is to create summaries of films in order to be able to extract the most important information from them and shorten their length[19]. The can be used to find strange events in videos

method that Yong Sheen Chong and Yong Haur Tay demonstrated.

Gaurav Kumar Singh and Vipin Shukla released a paper in 2020 titled "Automatic Alert of Security Threat utilising Video Surveillance System." [20][21]. This paper suggests a technique for using sensor systems that can alert on the existence of any suspicious activity. Sensors record events, but no information is provided about them. Analyzing the captured footage enables quick and accurate information gathering regarding the threat and its fundamental cause[22] in order to take preventative action. As a result, it's probable that deploying CCTV camera and sensor systems alone or in combination won't be sufficient to quickly identify harmful events. They consequently developed a system that employs cameras and sensor networks to swiftly detect threats in various lighting conditions.

Kooij et al[22] use of visual and audio data from surveillance videos allowed them to identify hostile behaviour. To evade tracking, several writers have proposed and employed a number of strategies. Motion patterns [23], histogram-based methods [23], social force models [23], topic modelling [23], and context-driven techniques [24] are all methods for learning global motion patterns. are a few examples of these difficulties in obtaining trustworthy tracks. To create a a surveillance infrastructure based on deep learning that uses object detection[25] a Deep Learning-based surveillance framework using object detection was implemented, according to a piece written by Bharath Raj[26].

The majority of studies have created strategies for instructing the distribution of typical movements through practise using already-existing recordings, according to the literature analysis carried out thus far.They have made an effort to recognise low partable patterns and classify them as anomalies. Some researchers have demonstrated that sparse matrices are more efficient in solving computer vision[26]-related tasks Moreover, certain patterns that result in a sizable restoration error are labelled as abnormal during testing[27]. Deep learning has shown to be the most efficient technique for classifying images, making it suitable for classifying video material.Deep convolutional neural networks [28] are now widely used in the literature for object detection and localization with class-specific bounding boxes. However, still there is not clear method of finding frames from videos in any of these previous researches.

According to this study, the uncertainty surrounding this field of study is the root of a lot of difficulties. There are many problems and most of the related concepts related to the problems are not explained clearly in previous researches for example:

- No particular algo or method was proposed to detect frames. So there is difficulty in getting time frames of the suspicious action in a video and it's also difficult to train very large amount of dataset to improve accuracy of predicting Crime.

- The system may face difficulty in detecting similar types of crime and non crime videos.

- The quality of the image and camera resolution has an impact on the system's ability to forecast outcomes.

- In scenarios with a lot of people, it can be challenging to predict any activity. In fact, it is unclear exactly what constitutes a crowd.

III. OBJECT DETECTION

The computer vision technique known as object detection makes it feasible to locate instances of objects in images or movies. Object detection algorithms usually make use of deep learning or machine learning to get relevant findings. It is commonly used in face detection, face and object co-segmentation, vehicle counting, activity recognition, picture annotation, and other computer vision applications. Moreover, it is used to track moving objects like a football during a game, a cricket bat during a match, or a person in a movie.

IV. YOLO

The abbreviation YOLO stands for You Only Look Once. This algorithm finds and recognises several elements in a photo (in real-time). The YOLO object identification method, which is done as a regression problem, gives the class probabilities of the detected photographs. The YOLO method recognises objects quickly by using convolutional neural networks (CNN). The method just requires one forward propagation through a neural network, as its name would suggest, to detect objects. The YOLO algorithm has several different iterations. The two most well-known ones are Small YOLO [28] and YOLOv3 [28]. This shows that the entire image is subjected to a single algorithm run for prediction. The CNN is used to predict multiple class probabilities and bounding boxes simultaneously [29]. The yolo algorithm generates boundary boxes. The term "bounding box" refers to an outline that highlights an object in a photograph. Each bounding box in the picture possesses the following characteristics: Height, Width, 30. Yolo's comparison to other algorithms is seen in this graph. It displays the speed of different algorithms in frames per second. [31] and yolo has highest speed , reliability and accuracy as compared to other algorithms.
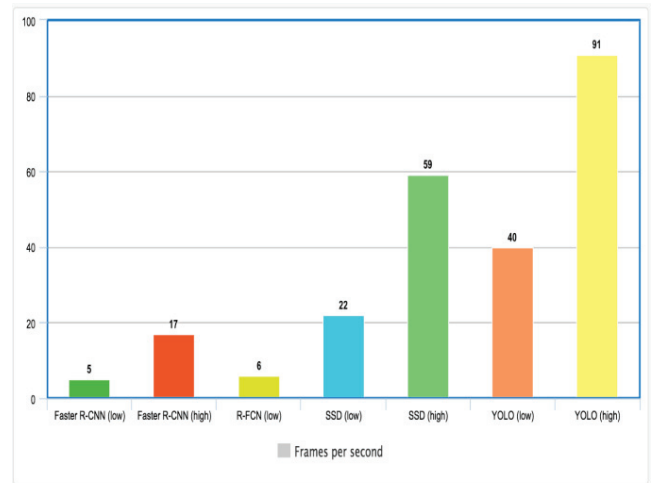

Fig. 1. FPS Capability of various object detectors

V. PROPOSED WORKFLOW

The suggested work tends to produce an application for the instantaneous detection of anomalies in public spaces. The suggested method is suited for surveillance in public areas like shopping centres, airports, train stations, etc. where there is a risk of theft or shooting [26]. Deep learning and neural networks [27] will be used in the proposed system to train it. Then, this model will be put into practise as a desktop application that will accept real-time CCTV footage as input and notify the administrator when a suspicious action is discovered. To downaload the dataset for our project. Go to the UCF Crime Website and download the large file of Crime Dataset of 1900 long real time videos of crime on 8 – 10 different Categories like shop lifting , abuse , shooting etc .The design and flow of the project can be shown in the below flow chart.

- The main approach is to detect objects and predict the anomalous behaviour humans in cctv feeds for threat detection .

- Initially, we will get the crime dataset from the UCF crime website. Next, we will pre-process it by extracting the priority frames where suspicious events, such as robberies, murders, and kidnappings, had occurred.

- And then we prepare one nueral network model using the various ml , CV algorithms to detect the objects in the different images of the video during that interval.

- With this training dataset, we trained our machine using the Keras model and Tensor flow. I've used yolo algo to detect objects.

- And the train the model accordingly the type of action performed in the video.

- At last of simulation, the performance parameters of proposed work will be calculated in terms of Error Rate, Accuracy.
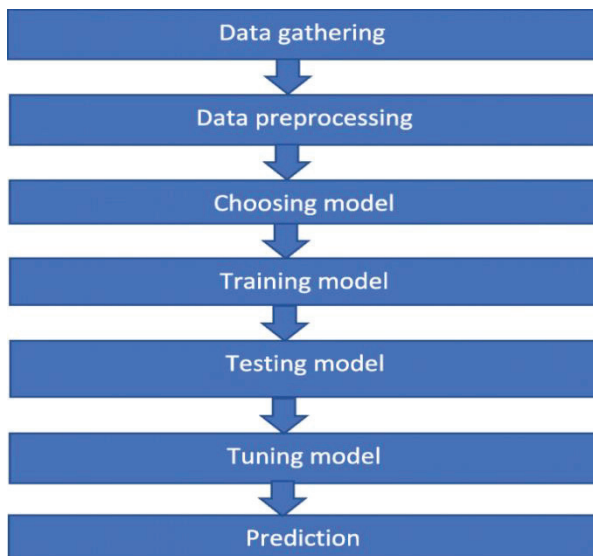

Fig. 2. Flowchart of the of the proposed work

A. Data Gathering

Firstly we have to Collect all the data about our topic this is the very first step in every research work. we started collecting data from all studying all the previous researches and then conducted literature survey and come to now about various datasets used. we needed a large dataset of different crime videos on various categories. So we used the UCF Crime dataset. We collected all the data from this dataset. It contains CCTV videos of crime. it is a very large dataset and containing real time crime videos.

B. Data Pre-Processing

In this step we will pre processes the data ie we clean the data and extract only the priority frames from the cctv videos. Priority frames extraction includes only that frames of video in which the actual crime has happened This can be done by using fps system we can set al limit of frames sequence that we want ex 10 frames per second.

C. Choosing Model

The model we create is a nueral network model with Convolutional Layers. The already-trained InceptionV3 model [30] is used to train our CNN utilizing the transfer learning technique. Transfer learning utilizes the feature extraction portion of the new model, while the classification component is retrained using the original dataset. The learning process as a whole requires less processing power and training time because the feature extraction component we don't need to learn is a particularly complex component of the model. The output of the Inception model is sent to CNN, which isn't the final classification model.

D. Training and testing Model

The generated neural network model is then trained and evaluated using several films..the model is trained based on object detection and activity detection. For this we used yolo. And then it will predict the activity in the videos. For ex. The two objects detected in the video are two cars and if we have to detect an accident we will train the model according to that if the two object overlap or the distance bw them is less than 0.1cm. The flowchart of the model trained is shown below.
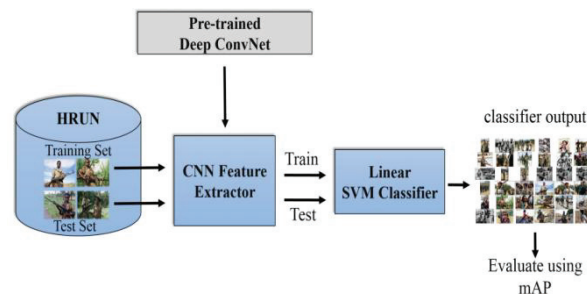

Fig. 3. Flowchart of the model

VI. CONCUSION AND FUTURE SCOPE

This paper gives an overview of why this topic building intelligent surveillance System is important for research in AI and ML Field. It also discusses the review of all the previous researches done in this Field from various object detection methods to finding priority Frames techniques used to detect

crimes, datasets used and algorithms used to analyze crime data and train the dataset .We discuss the recent researches in this field and analyze the various research gaps and then we proposed our approach and system of developing the intelligent surveillance system. In the previous researches the datasets used are not so large and does not contain the real time Crime Videos.But in the system we are proposing we are using the UCF Crime dataset .However, to enable our system to be trained on a variety of crime categories, we are leveraging the UCF Crime dataset, the largest crime dataset that contains real-time crime recordings on many different categories, such as fire, murder, robbery, abuse, and assault. And because of the YOLO algorithm's speed, high accuracy, and learning capabilities, we used it for object detection.

When compared to real-time object detectors, YOLO has the inherent advantage of speed in addition to better Intersection over Union in bounding boxes and increased forecast accuracy. YOLO is a far quicker algorithm than its rivals, operating at up to 45 frames per second. After that, we train the model, and the system not only detects crimes but also classifies them into distinct categories using the k-means approach in machine learning. All of this will aid in the creation of a monitoring system that is more effective and accurate.

The proposed System has a lot of scope for future research. The existing system can be further trained on more categories and its accuracy can be improved. This system proposed can be further used to predict the seriousness of the crime for example in case of fire the extent of fire can be predicted in that place and it also helps in finding criminals by matching it with previous databases.

REFERENCES

[1] N. Hnoohom, P. Chotivatunyu, S. Yuenyong, K. Wongpatikaseree, S. Mekruksavanich and A. Jitpattanakul, "Object Identification and Localization of Visual Explanation for Weapon Detection," 2022 Research, Invention, and Innovation Congress: Innovative Electricals and Electronics (RI2C), 2022, pp. 144-147

[2] J. Candamo, M. Shreve, D. B. Goldgof, D. B. Sapper and R. Kasturi, "Understanding Transit Scenes: A Survey on Human Behavior-Recognition Algorithms," in IEEE Transactions on Intelligent Transportation Systems, vol. 11, no. 1, pp. 206-224, March 2010

[3] N. Y. Katkar and V. K. Garg, "Detection and Tracking the Criminal Activity using Network of CCTV cameras," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), 2022, pp. 664-668, doi: 10.1109/ICOSEC54921.2022.9952104.

[4] Jebur SA, Hussein KA, Hoomod HK, Alzubaidi L, Santamaría J. Review on Deep Learning Approaches for Anomaly Event Detection in Video Surveillance. Electronics. 2022 Dec 22;12(1):29.

[5] Choudhari, S.J., Singh, S.A., Kumar, A.S. and Desai, K.A., 2022, June. Machine Setup Abnormality Detection Using Machine Vision and Deep Learning. In International Manufacturing Science and Engineering Conference (Vol. 85802, p. V001T04A009). American Society of Mechanical Engineers.

[6] S. W. Khan et al., "Anomaly Detection in Traffic Surveillance Videos Using Deep Learning," Sensors, vol. 22, no. 17, p. 6563, Aug. 2022, doi: 10.3390/s22176563.

[7] S. Shirsat, A. Naik, D. Tamse, J. Yadav, P. Shetgaonkar and S. Aswale, "Proposed System for Criminal Detection and Recognition on CCTV Data Using Cloud and Machine Learning," 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN),2019,pp.1-6/ViTECoN.2019.8899441.

[8] S. Ojha and S. Sakhare, "Image processing techniques for object tracking in video surveillance- A survey," 2015 InternationalConference on Pervasive Computing (ICPC), 2015, pp. 1-6, doi: 10.1109/PERVASIVE.2015.7087180. Conference on Pervasive Computing (ICPC), 2015, pp. 1-6, doi: 10.1109/PERVASIVE.2015.7087180.

[9] Chowdhary, Aastha, and Bhawana Rudra. "Video surveillance for the crime detection using features." In International Conference on Advanced Machine Learning Technologies and Applications, pp. 61-71. Springer, Singapore, 2021.

[10] Sikandar, T., Ghazali, K.H. and Rabbi, M.F., 2019. ATM crime detection using image processing integrated video surveillance: a systematic review. Multimedia Systems, 25(3), pp.229-251.

[11] Katkar, Nihanshu Yuvraj, and Vijay Kumar Garg. "Detection and Tracking the Criminal Activity using Network of CCTV cameras." 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC). IEEE, 2022.

[12] Dib, J., Sirlantzis, K. and Howells, G., 2022, October. Application of Deep Learning Techniques in Negative Road Anomalies Detection. In Proceedings of the 14th International Joint Conference on Computational Intelligence (Vol. 1, pp. 475-482).

[13] Jalgaonkar, Drishti, et al. "Detecting Crime Scenes using ML." (2020).

[14] Benzir, Mousumi Yeasmin. "A Study of Crowd Abnormal Events Understanding in Surveillance Videos." PLOMS AI 2.1 (2022).

[15] Althoff, L., Farias, M.C. and Weigang, L., 2022, November. Once Learning for Looking and Identifying Based on YOLO-v5 Object Detection. In Proceedings of the Brazilian Symposium on Multimedia and the Web (pp. 298-304).

[16] She C, Zheng R, Yang Q, Liang S. Research of Intelligent Video Surveillance System based on Artificial Neural Network. InJournal of Physics: Conference Series 2022 (Vol. 2181, No. 1, p. 012057). IOP Publishing.

[17] R. Leyva, V. Sanchez and C. -T. Li, "Video Anomaly Detection With Compact Feature Sets for Online Performance," in IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3463-3478, July 2017.

[18] Gorave, Asmita, et al. "Suspicious Activity Detection Using Live Video Analysis." Proceeding of International Conference on Computational Science and Applications. Springer, Singapore, 2020.

[19] A. Warsi, M. Abdullah, M. N. Husen and M. Yahya, "Automatic Handgun and Knife Detection Algorithms: A Review," 2020 14th International Conference on Ubiquitous Information Management and Communication (IMCOM), 2020, pp. 1-9, doi: 10.1109/IMCOM48794.2020.9001725.

[20] Sung, C.S. and Park, J.Y., 2021. Design of an intelligent video surveillance system for crime prevention: applying deep learning technology. Multimedia Tools and Applications, 80(26), pp.34297-34309.

[21] F. Porikli et al., "Video surveillance: past, present, and now the future [DSP Forum]," in IEEE Signal Processing Magazine, vol. 30, no. 3, pp. 190-198, May 2013

[22] Zheng, Zhaohui, et al. "Distance-IoU loss: Faster and better learning for bounding box regression." Proceedings of the AAAI conference on artificial intelligence. Vol. 34. No. 07. 2020.

[23] Lu, Shengyu, Beizhan Wang, Hongji Wang, Lihao Chen, Ma Linjian, and Xiaoyan Zhang. "A real-time object detection algorithm for video." Computers & Electrical Engineering 77 (2019): 398-408.

[24] Kumar, Kadari Kishore, and Husnabad Venkateswara Reddy. "Crime activities prediction system in video surveillance by an optimized deep learning framework." Concurrency and Computation: Practice and Experience 34.11 (2022): e6852.

[25] Li, Liyuan, et al. "Statistical modeling of complex backgrounds for foreground object detection." IEEE Transactions on image processing 13.11 (2004): 1459-1472.

[26] Toshniwal, Akanksha, Kavi Mahesh, and R. Jayashree. "Overview of anomaly detection techniques in machine learning." In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), pp. 808-815. IEEE, 2020.

[27] C. Zhan, X. Duan, S. Xu, Z. Song and M. Luo, "An Improved Moving Object Detection Algorithm Based on Frame Difference and Edge

Detection," Fourth International Conference on Image and Graphics (ICIG 2007), 2007, pp. 519-523.

[28] Tripathi, R.K., Jalal, A.S. and Agrawal, S.C., 2018. Suspicious human activity recognition: a review. *Artificial Intelligence Review*, *50*(2), pp.283-339.

[29] Singh, Virender, Swati Singh, and Pooja Gupta. "Real-time anomaly recognition through CCTV using neural networks." *Procedia Computer Science* 173 (2020): 254-263.

[30] Sreenu, G., and Saleem Durai. "Intelligent video surveillance: a review through deep learning techniques for crowd analysis." *Journal of Big Data* 6, no. 1 (2019): 1-27.

[31] Pang, Guansong, et al. "Deep learning for anomaly detection: A review." *ACM Computing Surveys (CSUR)* 54.2 (2021): 1-38.