A Crowd Monitoring and Real-Time Tracking System using CNN

Ajay K Baiju
Dept. of CSE
College of Engineering Kidangoor
Kottayam,Kerala,India
ajaykbaiju@gmail.com

Akshara Manoj
Dept. of CSE
College of Engineering Kidangoor
Kottayam,Kerala,India
aksharamanoj02@gmail.com

Ganga Binukumar
Dept. of CSE
College of Engineering Kidangoor
Kottayam,Kerala,India
gangabinukumar03@gmail.com

Sandra Elizabeth Rony
Dept. of CSE
College of Engineering Kidangoor
Kottayam,Kerala,India
officialsandrarony@gmail.com

Jyothis Joseph
Dept. of CSE
College of Engineering Kidangoor
Kottayam,Kerala,India
jyothis@ce-kgr.org

***Abstract*—In response to the ever-evolving landscape of security concerns, the project endeavors to fortify public safety through an innovative surveillance system called SECURE SPHERE. The system strategically places cameras to continuously monitor crowd behavior, employing cutting-edge algorithms to detect abnormalities, such as attacking, fighting or the presence of weapons. Upon identification of anomalies, alerts are promptly transmitted to law enforcement agencies through an intuitive and user-friendly application. A distinctive feature of the project lies in its real-time tracking capability, allowing for the monitoring of a culprit's movement captured on any of the strategically positioned cameras. The proposed innovation significantly enhances the system's efficacy in providing precise and actionable information to law enforcement, thereby bolstering efforts to ensure public safety. The heart of this initiative lies in the Application Interface, providing law enforcement with an accessible and user friendly platform. The interface not only enables the viewing of alerts but also grants access to the invaluable real-time tracking feature. The integrated approach to public safety envisioned in the project ensures that law enforcement agencies are equipped with the essential tools and information needed to respond swiftly and effectively to incidents. By seamlessly combining continuous crowdmonitoring, anomaly detection, real-time tracking, and user-friendly interfaces, the project strives to create a safer environment for the public. The comprehensive surveillance system, with its advanced technology and integrated features, not only serves as a proactive measure but also contributes to the creation of a secure and resilient societal environment.**

***Keywords*—Abnormal Activity, Convolutional Neural Network**

I. INTRODUCTION

Public safety has emerged as a critical focus for law enforcement agencies globally, necessitating the deployment of sophisticated technologies for proactive and effective surveillance. Video Surveillance plays a pivotal role in today's world. The technologies have advanced too much when artificial intelligence and machine learning are pitched into the system. Using above combinations, different systems are in place which helps to differentiate various suspicious behaviors The ubiquity of Closed-Circuit Television (CCTV) cameras in public spaces reflects an earnest attempt to address these challenges. However, the sheer volume of visual data generated by these surveillance devices has given rise to a new set of hurdles. The efficacy of surveillance is no longer solely contingent on the ability to capture copious amounts of data; it hinges on the capacity to distill meaningful insights from this data swiftly and accurately. The paradigm shift towards smart cities and advanced security measures has led to an increased reliance on CCTV cameras in public spaces. While these cameras serve as vigilant eyes, the sheer volume of data they generate necessitates intelligent systems capable of discerning patterns indicative of abnormal behavior. The proposed framework represents a novel approach to this challenge, harnessing the capabilities of deep learning through CNNs to provide law enforcement agencies with a proactive and responsive surveillance solution.

II. MOTIVATION

The motivation behind this research is rooted in the imperative to address the multifaceted and evolving challenges in the realm of public safety and urban security. In the face of rapid urbanization, increased public gatherings, and the growing potential for unconventional security threats, the need for advanced surveillance systems has become more pressing than ever. Traditional surveillance methods, relying heavily on human monitoring, are confronting limitations such as fatigue, human error, and the inability to process vast amounts of real-time data. The ubiquity of CCTV cameras, while a testament to technological advancements, has introduced a new set of complexities. The sheer volume of visual data generated by these cameras necessitates a paradigm shift in surveillance capabilities. The challenge extends beyond data capture to discerning meaningful patterns within that data—identifying anomalies that may signify potential security risks. This research is motivated by the aspiration to bridge existing gaps in surveillance infrastructure. Traditional systems, often reliant on human monitoring, are prone to limitations that hinder their effectiveness. The envisioned framework, driven by advanced CNN algorithms, seeks to overcome these limitations and augment the surveillance capabilities of law enforcement agencies. Furthermore, as smart city initiatives become integral to urban planning, the integration of technology demands intelligent systems that not only respond to security threats but anticipate and prevent them. The proposed framework is motivated by the desire to contribute to the creation of a proactive and responsive surveillance ecosystem—one that can analyze patterns, identify abnormalities, and enable preemptive measures in the interest of public safety. In the context of evolving security landscapes marked by the rise of cyber-physical threats, terrorism, and public safety concerns, there is a compelling need for surveillance systems that transcend conventional means. By harnessing the capabilities of CNN algorithms, this research seeks to empower law enforcement with a toolset that goes beyond the limitations of traditional surveillance, ensuring thatcities and public spaces remain secure, resilient, and prepared to tackle emerging security challenges. In essence, the motivation behind this research lies in the pursuit of technological innovation to create a surveillance framework that not only keeps pace with societal changes but anticipates and addresses security threats in real-time, fostering safer and more secure urban environments.

III. LITERATURE REVIEW

The landscape of abnormal activity detection and tracking systems has witnessed substantial growth as technology continues to play a pivotal role in shaping the future of public safety. In this literature survey, we embark on a comprehensive exploration of existing research endeavors that have paved the way for the development of sophisticated surveillance systems. Focused on the amalgamation of advanced CNN-based algorithms and cross-camera tracking mechanisms, this survey seeks to unravel the nuances of key studies, methodologies, and advancements in the realm of abnormal activity monitoring. By scrutinizing seminal contributions in abnormal activity detection, delving into the intricacies of CNN-based approaches, and dissecting studies on cross-camera tracking, we aim to provide a panoramic view of the state-of-the-art in the field. Through this synthesis, we lay the groundwork for the subsequent exposition of our proposed framework, which aspires to bridge gaps, overcome challenges, and contribute to the ever-evolving landscape of public safety technology.

A. Crowd Monitoring and Localization

The paper [5] proposed a model for crowd monitoring, specifically focusing on person counting and detection using Deep Convolutional Neural Network (CNN) and Support Vector Machine (SVM) techniques. The paper emphasizes the importance of crowd monitoring in India due to its high population density, increased criminal activities, and the spread of diseases in crowded areas. It aims to address the challenges associated with crowd behavior analysis and proposes methodologies for crowd counting, localization, and detection using various real-world datasets from locations such as malls, Kumbh Mela, and UCFD. The Deep Convolutional Neural Network (CNN) is particularly well-suited for image analysis and computer vision tasks. It is used for feature extraction and identification of specific image characteristics, making it effective for tasks such as crowd counting, behavior analysis, and person detection in crowded areas. On the other hand, Support Vector Machine (SVM) is a popular supervised machine learning algorithm used for classification and regression tasks.The combination of CNN and SVM allows for a comprehensive approach to crowd monitoring, leveraging the strengths of each technique. RFID (Radio Frequency Identification) technology is utilized in conjunction with Deep Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for crowd monitoring. RFID is employed to identify individuals in crowded areas, both indoors and outdoors, using radio frequency signals and unique tag chip identification cards. This technology is particularly useful for real-time counting and tracking of individuals through wireless sensor networks and GPS tracking devices. In the context of this paper, RFID is used to identify crowded individuals in various locations such as malls, Kumbh Mela, and UCFD. The data collected through RFID technology is then integrated into the crowd management and monitoring model, which utilizes CNN and SVM techniques for crowd counting, localization, and behavior analysis. The combination of RFID with CNN and SVM allows for a comprehensive approach to crowd management, leveraging the strengths of each technology for accurate counting and detection of individuals in crowded areas. Regression guided detection network (RDNet) method is proposed for RGB-Datasets and can simultaneously estimate head counts and localize heads with bounding boxes.A density map is used to localize the heads in dense images with accurate results.Compressed Sensing based Output Encoding (CSOE)method has been proposed toimprove the efficiency of localization in highly dense crowded situations. Scale Driven Convolutional Neural Network (SD-CNN): SD-CNN has been proposed to count and localize the crowd, addressing the issue of scale variation and reducing classification time significantly while improving detection accuracy.DISAM (Density Independent and Scale Aware Model) model has the ability to precisely localize the heads in complex scenes and handle both counting and localizing people in dense crowds. [4]

B. Abnormal Activity Detection

The paper [1] introduces a deep learning approach for detecting suspicious activities from surveillance videos in an academic environment. It discusses the use of neural networks, specifically CNN and RNN, for feature extraction and classification, and the training of a pre-trained model called VGG-16 to predict human behavior from surveillance footage. Additionally, the paper includes the use of LSTM architecture for classification, and the implementation of the system in an academic environment. The accuracy achieved by the model in classifying videos as suspicious or normal behavior is reported as 87.15%. The methodologies used in the paper demonstrate the application of advanced deep learning techniques for detecting and classifying suspicious activities insurveillance videos, with a focus on enhancing security and safety measures in various environments. The author mentions the use of two datasets for training and testing the system. The first dataset is the KTH dataset, which contains sequences representing six actions, with each action class having 100 sequences. Each sequence consists of approximately 600 frames, and the videos are shot at 25 frames per second (fps). The KTH dataset is used for training the model to recognize normal behavior, specifically running and walking. The second dataset is a combination of the CAVIAR dataset, videos taken from the campus, and YouTube videos. This dataset is used for training the system to recognize suspicious behavior, such as using a mobile phone inside the campus, fighting, and fainting. The CAVIAR dataset, along with videos captured from the campus and YouTube, provides a diverse set of frames for training the system to detect suspicious activities. The entire dataset is manually labeled and separated into 80% for the training set and 20% for the validation set. The limitations of the approach include challenges related to accuracy, dataset limitations, pre-processing complexity, real-time processing, manual intervention, and model generalization. The paper also highlights the potential for the proposed model to prevent crimes before they occur and its applicability beyond academic environments. Additionally, the system's ability to send SMS alerts to the respective authority in case of suspicious behavior is emphasized.

The paper [8] includes an ECNN algorithm which is a mechanism for detecting abnormal or suspicious actions such as shooting and stealing from surveillance video datasets. It employs a Residual Neural Network (ResNet) CNN architecture to detect these actions by creating classes for CNN blocks. Video datasets are processed as image sequences, where RGB frames are converted to grayscale to capture intensity information. Optical flow is applied to identify patterns among objects, surfaces, and edges, enhancing the visual sense of the scenes. The algorithm's performance metrics including accuracy, precision, false-positive rate, and false-negative rate were evaluated, with reported figures of 98.38% accuracy, 98.54% precision, 1.25% false-positive rate, and 1.66% false-negative rate. Meanperformance measures, calculated using SPSS, showed ECNN achieving 97.050% accuracy, 96.743% precision, 2.957% false-positive rate, and 2.927% false-negative rate. The results indicate that ECNN outperformed traditional CNN in detecting suspicious activities, demonstrating its novelty and efficacy.

In paper [3] the author describes a deep learning assistive framework for efficiently recognizing violent activity in surveillance videos during the Hajj pilgrimage. The framework includes the training of a lightweight CNN object detector on pilgrims' datasets to select specific video frames for further processing, as well as the development of a lightweight sequential learning LSMT model for the extraction of spatial and temporal features. The proposed framework aims to accurately identify violent activity and trigger an alarm to notify law enforcement agencies to take appropriate action in the case of any violent activity. The paper also discusses the experimental setup and evaluation of the proposed model, as well as its potential for future investigation, including the development of a dataset for violent activity recognition during the Hajj pilgrimage and the exploration of different deep learning models for efficient learning of motion and spatio-temporal features. The framework involves several steps. First, a lightweight CNN model is trained on a pilgrim's dataset to detect pilgrims from surveillance cameras. These preprocessed frames are then passed to a lightweight CNN model for spatial feature extraction. Subsequently, a Long Short Term Memory (LSTM) network is developed to extract temporal features. Finally, in the event of violent activity or accidents, the system generates real-time alarms to inform law enforcement agencies to take appropriate action. The authors conducted experiments on publicly available violent activity datasets, achieving high accuracies of 81.05% and 98.00% on the Surveillance Fight and Hockey Fight datasets, respectively. The proposed framework aims to address the urgent need for an intelligent and automatic system to efficiently monitor crowds and identify abnormal activities in surveillance videos, particularlyin the context of religious gatherings such as Hajj and Umrah.

In paper [10], utilized different Deep Learning models for real-time anomaly recognition through CCTV. Here, two Different Neural Networks: CNN and RNN have been used. CNN is the basic neural network that is being used primarily for extracting advanced feature maps from the available recordings. This extraction of high-level feature maps alleviates the complexity of the input. To apply the technique of transfer learning, we use InceptionV3- a pre-trained model. The inceptionV3, pre-trained, is selected by keeping in view that the modern models used for object recognition consider loads of parameters and thus take an enormous amount of timeto completely train it. However, the approach of transfer learning would enhance this task by considering initially the previously learned model for some set of classified inputs e.g. ImageNet; which further can be re-trained based on the new weights assigned to various new classes. The output of CNN isfed to the RNN as input. RNN has one additional capability of predicting the next item in a sequence. Therefore, it essentiallyacts as a forecasting engine. Providing the sense to the captured sequence of actions/movements in the recordings is the motivation behind using this neural network in this work. This network has an LSTM cell in the primary layer, trailed bysome hidden layers with appropriate activation functions, and the output layer will give the final classification of the video into the 13 groups (12 anomalies and 1 normal). The output of this system is used to perform real-time surveillance on the CCTV cameras of different organisations to avoid and detect any suspicious activity. Hence, the time complexity is reducedto a great extent.

The UMN dataset incorporates five distinctively arranged recordings where people walk around and begin running in various directions after some time. The suspicious activity is characterized by running action. In the following dataset, the author has used UCSD Ped1 dataset which contains 70 and UCSD Ped2 dataset having 28 CCTV recordings.These videos have been recorded at a single place. The recorded suspicious activities are basic and do not add any significant value to CCTV surveillance. Avenue is the third dataset that contains 37 recordings. Although it consists of comparatively more oddities, they are recorded at a single place. Similarly , the time span of the recordings in this dataset is not large with few anomalies such as throwing paper are unrealistic. Subway Exit and Subway Entrance datasets form the fourth dataset that contains single long surveillance recording for each. Basic irregularities like moving in the incorrect direction and payment skip are the part of this dataset. Lastly, BOSS dataset is obtained from a surveillance CCTV installed on a train's roof. But these abnormalities are performed by actors containing peculiarities, for example, harassment, a person with a disease, panic circumstances, along with the normal recordings. All these previously utilized datasets are short in terms of quantity and time-span of the recording. Also, they cover limited anomalies and some abnormalities are not even realistic.

The document [10] outlines a proposal for detecting suspicious and anomaly activities using a Convolutional Neural Network (CNN) architecture, Haar Cascade classifier.The project utilizes open-source datasets available on Kaggle, forming a total of 6 classes, including activities such as running, jumping, and kicking in public places, as well as carrying guns, bats, and knives.Convolutional Neural Network (CNN) serves as a cornerstone for detecting dangerous objects and activities, presenting a sophisticated architecture tailored for image processing. With its specialized layers, including convolutional, pooling, dropout, and fully connected layers, the CNN scrutinizes input images, extracting intricate features crucial for identification. Convolutional layers meticulously apply filters or kernels across the input image, discerning edges, textures, and other patterns essential for recognition. Subsequently, pooling layers condense the spatial dimensions of these features while preserving vital information, optimizing computational efficiency. Dropout layers strategically deactivate neurons during training to curb overfitting, ensuring the network generalizes well to unseen data. Global average pooling further refines feature representation, paving the way for accurate classification through dense layers. Activation functions such as ReLU, sigmoid, and softmax introduce non linearity, vital for capturing complex relationships within the data. Through meticulous training with the Adam optimizer and categorical cross-entropy loss function, the CNN attains remarkable accuracy, laying the groundwork for detecting dangerous objects and activities. The Haar Cascade classifier is an object detection method characterized by its efficient hierarchical cascade of classifiers. Initially, Haar-like features are extracted from training images, capturing variations in intensity. These features are then used to train weak classifiers through the Adaboost algorithm, which iteratively focuses on misclassified samples. The trained classifiers are organized into a cascade, with each stage progressively increasing in complexity. Duringdetection, the cascade evaluates regions of interest in the inputimage using a sliding window approach, swiftly discarding non-object regions. This sequential evaluation, coupled with a mechanism to prioritize promising regions, enables rapid and accurate identification of objects while minimizing computational overhead and false positives. Overall, the Haar Cascade classifier provides a robust framework for object detection, adept at handling complex scenes with efficiency.

C. Human Tracking

The paper provides a comprehensive overview of the evolution and interplay of object detection algorithms, particularly focusing on YOLO (You Only Look Once), R-CNN (Region based Convolutional Neural Networks), and Deep SORT (Simple Online and Realtime Tracking). It delves into the technical intricacies and performance metrics of these algorithms, emphasizing their significance in human tracking applications, surveillance, security predictions, and autonomous vehicles. The review highlights the unique features and developments of YOLOand R-CNN, such as processing speed, object detection capabilities, and potential modifications for specific tracking applications.

The Deep SORT algorithm plays a pivotal role in enhancing object tracking, particularly in addressing challenges such as occlusion in object detection. It achieves this by predicting an object's trajectory during occlusions, providing more accurate predictions. Deep SORT leverages a combination of techniques, including recursive Kalman filtering and the Hungarian method, specifically designed to measure bounding box over laps. Additionally, the algorithm incorporates a convolutional neural network trained to distinguish individual pedestrians using a large-scale dataset. This approach allows Deep SORT to outperform other algorithms in handling occlusions and predicting object trajectories with high precision. The algorithm's efficiency, swift training, and modest computational footprint make it a valuable tool for developers and institutions with varying levels of technical resources. However, future work should focus on making it more versatile and ensuring consistent performance across a range of hardware setups, particularly in real-time scenarios and situations with limited GPU capabilities. This pursuit of optimal performance paired with broad accessibility will continue to drive innovations in the field of object detection and tracking. [7]

D. Face Recognition

The document [6] a real-time face recognition system for video surveillance using a Haar cascade classifier. It aims to automate the identification of individuals in surveillance videos, reducing the need for manual review. It utilizes a surveillance camera system, Raspberry Pi as a processor, and Pi Camera as a camera module. The system's successful output demonstrates human face recognition within specific parameters, such as facial angle, light conditions, and distance, with the potential to reduce manpower costs in real-time identification scenarios. The system utilizes a Raspberry Pi as the main processing unit and a Pi Camera module for image capture.

It follows three main phases: data gathering, training the recognizer, and face recognition process, all implemented using Python programming and OpenCV library on a Raspbian operating system.Data Gathering:The system captures the input image using the Pi Camera module.The captured images are then converted to grayscale to facilitate processing and reduce memory usage.The system collects various facial impressions and styles, including images with and without accessories such as glasses and different hair styles. The collected data is stored in the memory of the Raspberry Pi.Training Recognizer:Each collected face image is assigned a specific ID for training purposes. The system creates a dataset folder and a trainer folder to organize the collected data.A specific file named trainer. yml is created for training the recognizer with the collected data. The system uses the training data to learn and recognize different IDs for different persons.Face Recognition Process: The system executes the recognition code to perform the face detection and recognition process.The input image of the detected faceis compared with the set of images in the database automatically. The system displays the output result of the recognized human face, including the name of the person based on the set person ID.If the person is not in the database, the system displays an "UNKNOWN" result. The output of the system includes the recognition of known individuals based on the input image, as well as the ability to detect and classify unknown individuals. The system's performance is influenced by factors such as light conditions, facial angles, and the presence of accessories. The system has been tested and validated for its accuracy and limitations, with the goal of reducing the cost of manpower for real-time identification in security and application systems.The collected data includes various facial impressions and styles, with and without accessories such as glasses and different hair styles. The system collected 60 samples of data for each person, with the minimum requirement for the system to run face detection being at least 30 images for accuracy. The collected data is stored in the memory of the Raspberry Pi, utilizing a 32 GB memory capacity.

IV. CONCLUSION

The paper provides a comprehensive exploration ofadvanced surveillance systems, focusing on the development of innovative frameworks for abnormal activity detection, human tracking, and face recognition in video surveillance. It discusses the integration of cutting-edge technologies such as Convolutional Neural Networks (CNNs), Recursive Neural Networks (RNNs), and Haar Cascade classifiers to enhance public safety and security measures.

Specifically, the paper examines the application of CNNs and SVMs for crowd monitoring and localization, highlighting their effectiveness in analyzing complex scenes and detecting anomalies. It also discusses the use of CNN architectures like Residual Neural Networks (ResNets) for detecting suspicious activities, demonstrating their superior performance in comparison to traditional methods. Furthermore, the paper addresses the development of lightweight CNN and LSTM models for efficient recognition of violent activity during large gatherings, such as the Hajj pilgrimage, showcasing the potential of deep learning in safeguarding public events.

In the realm of object detection and human tracking, the paper reviews the evolution of algorithms like YOLO, R CNN, and Deep SORT, emphasizing their significance in surveillance applications. It highlights the unique capabilities of Deep SORT in handling occlusions and predicting object trajectories accurately, thus enhancing the overall efficiency of surveillance systems.

Additionally, the paper discusses the implementation of real time face recognition systems using Haar cascade classifiers, outlining the process of data gathering, training, and recognition. It illustrates the practical application of these systems in automating the identification of individuals in surveillance videos, thereby reducing manual review efforts and enhancing operational efficiency.

Overall, the paper underscores the importance of leveraging advanced technologies such as deep learning and object detection algorithms to address the evolving challenges in public safety and security. By integrating these technologies into surveillance systems, authorities can proactively identify and respond to potential threats, fostering safer and more secure urban environments.

REFERENCES

[1] Amrutha C.V, C. Jyotsna, Amudha J, ”Deep Learning Approach for Suspicious Activity Detection from Surveillance Video”, Proceedings ofthe Second International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2020), IEEE Xplore Part Number: CFP20K58-ART; ISBN: 978-1-7281-4167-1

[2] P. Kuppusamy, V.C. Bharathi,Human abnormal behavior detection usingCNNs in crowded and uncrowded surveillance– A survey, Measurement: Sensors, Volume 24, 2022, 100510, ISSN 2665-9174, https://doi.org/10.1016/j.measen.2022.100510.

[3] Habib, S.; Hussain, A.; Albattah, W.; Islam, M.; Khan, S.; Khan, R.U.; Khan, K. Abnormal Activity Recognition from Surveillance Videos Using Convolutional Neural Network. Sensors 2021, 21, 8291. https://doi.org/10.3390/s21248291

[4] Khan, Akbar Shah, Jawad Kadir, Kushsairy Albattah, Waleed Khan, Faizullah. (2020). Crowd Monitoring and Localization Using Deep Convolutional Neural Network: A Review. Applied Sciences. 10. 10.3390/app10144781.

[5] Raju Pal and Praveen K. Shukla (eds.), SCRS Conference Proceedings on Intelligent Systems, 171–179. Computing and Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-93-91842-08-6-15

[6] Ahmad, A. H., Saon, S., Mahamad, A. K., Darujati, C., Mudjanarko, S. W., Nugroho, S. B., Hariadi, M. (2021). ”Real time face recognition of video surveillance system using haar cascade classifier”. Indonesian Journal of Electrical Engineering and Computer Science, 21(3), 1389. https://doi.org/10.11591/ijeecs.v21.i3.pp1389-1399

[7] Chen, Yancheng. (2024). Evolution of CNNs in human tracking applications. Applied and Computational Engineering. 36. 135-139. 10.54254/2755-2721/36/20230434.

[8] Citation: Selvi, E.; Adimoolam, M.; Karthi, G.; Thinakaran, K.; Balamu rugan, N.M.; Kannadasan, R.; Wechtaisong, C.; Khan, A.A. Suspicious Actions Detection System Using Enhanced CNN and Surveillance Video. Electronics 2022, 11, 4210. https://doi.org/10.3390/ electronics11244210

[9] Monali Ahire, Devarshi Borse, Amey Chavan, Shubham Deshmukh, Favin Fernandes,”Suspicious and Anomaly Detection”,Department of Electronics and Telecommunication, Vishwakarma Institute of Technology,Pune,2020,https://arxiv.org/ftp/arxiv/papers/2209/2209.03576.pdf

[10] Virender Singha, Swati Singha,Dr. Pooja Guptaa, (2020). ”Real-Time anomaly recognition through CCTV using neural networks” Procedia Computer Science, 173, 254–263. https://doi.org/10.1016/j.procs.2020.06.030

[11] Design of a real-time crime monitoring system using deep learn ing techniques. (2023, December 7). Design of a Real-time Crime Monitoring System Using Deep Learning Techniques- ScienceDirect. https://doi.org/10.1016/j.iswa.2023.200311 didula.c@sliit.lk

[12] Al-Nawashi, M. M., Al-Hazaimeh, O. M., Saraee, M. (2016). A novel framework for intelligent surveillance system based on abnormal human activity detection in academic environments. Neural Computing and Ap plications, 28(S1), 565–572. https://doi.org/10.1007/s00521-016-2363-z