International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

COPY RIGHT

ELSEVIER
SSRN

IJIEMR Transactions, online available on 18th October 2024. Link

https://ijiemr.org/downloads.php?vol=Volume-13 &issue= Issue 10

DOI:10.48047/IJIEMR/V13/ ISSUE10/49

Title: " INTRODUCING THE FIRST AI APPROACH FOR PREDICTING ROBBERY BEHAVIOR POTENTIAL USING INDOOR SECURITY CAMERAS"

Volume 13, ISSUE 10, Pages: 334- 344

Paper Authors
SAVALLA ARUN KUMAR, SYED ABDUL JALEEL

International Journal for Innovative Engineering and Management Research

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

INTRODUCING THE FIRST AI APPROACH FOR PREDICTING ROBBERY BEHAVIOR POTENTIAL USING INDOOR SECURITY CAMERAS

[1]SAVALLA ARUN KUMAR, [2]SYED ABDUL JALEEL
[1]Assistant Professor, Department of CSE, Sree Chaitanya College of Engineering, Karimnagar
[2]Student, Department of CSE, Sree Chaitanya College of Engineering, Karimnagar

ABSTRACT

For video surveillance systems to stop incidents and safeguard assets, crime prediction is necessary. In this regard, our paper presents the first artificial intelligence method for indoor camera Robbery Behaviour Potential (RBP) identification and prediction. Three detection modules—the head cover, crowd, and loitering detection modules—are the foundation of our approach, which enables us to take prompt action and stop robberies. Using our collected, manually annotated dataset, the YOLOV5 model is retrained for the first two modules. Additionally, we provide a brand-new definition for the DeepSORT-based loitering detection module. A fuzzy inference computer converts expert information into rules before making a final determination about the likelihood of a heist. This is difficult because the thief uses a different technique, the security camera is angled differently, and the video pictures have poor quality. Using actual video surveillance footage, we successfully completed our experiment and obtained an F1-score of 0.537. Thus, we develop threshold value for RBP to assess video pictures as a robbery detection issue and compare experimentally with other similar research. Assuming this, the experimental findings clearly reflect an F1-score of 0.607, indicating that the suggested technique outperforms other robbery detection methods in terms of identifying robberies. We firmly think that by anticipating and averting robbery incidents, the suggested method's implementation might reduce the harm caused by robberies at a surveillance camera control centre. However, the human operator's situational awareness improves and additional cameras may be controlled.

INDEX TERMS: Deep Learning Approach, RBP Prediction, Low Resolution, Fuzzy Inference Machine, and Deep Learning Method.

I. INTRODUCTION:

These days, surveillance cameras are often installed in a variety of locations, including houses, banks, shops, and airports, in order to improve public safety and deter crime. Alternatively, by examining these films and trying to identify the offender, it is possible to determine the date, time, and location of the incident as well as the identity of the perpetrator. A person is required to monitor the recordings from behind the scenes and identify any anomalies that may occur. However, the individual becomes exhausted since an anomaly occurs extremely seldom, and if it does, sometimes he is unaware that it has occurred. Stated otherwise, he loses the oddity [1]. Moreover, the process of detecting anomalies is grounded on human intuition that is acquired over time. However, there are additional issues with nonautomated crime prediction and detection systems that rely on examining surveillance films, such as the skill level of the individual and the expense of hiring him.

Certain visual cues need to be retrieved using machine learning and deep learning methods in order to automate anomaly detection [2], [3]. Certain properties for various anomaly classes [4], such as vandalism [5], violence detection [6], and robbery [7], may be helpful for these algorithms to operate more efficiently. Reducing the amount of devastation is predicting the place and timing of the act. However, security personnel are also there on schedule. In one such trial, cops in Santa Cruz, California, get daily crime projections every morning. They use this predicting to guide them as they monitor certain areas. According to a Santa Cruz representative, during the first six months of the trial, thirteen wrongdoers were apprehended in the designated regions [8]. According to papers [8] through [10], there are many key indicators that indicate the importance of using predictive policing for federal funding and security systems, such as reduced crime and cost. According to a 2021 Seattle Police Department (SPD) study, violent crimes in Washington, USA, rose by 20%, making them more hazardous owing to the likelihood of victimisation. [11]. Robbery is one of the top five most frequent crimes in the US, using data obtained from the FBI's Uniform Crime Reporting System (FBI-UCR) [12]. One of the numerous reasons security cameras are installed is to detect robberies. According to the Oxford Dictionary, robbery is the act of stealing or trying to take any property by force, threat, or weaponry [13]. It is distinguished from other types of theft like shoplifting, pickpocketing, or burglary by its inherently violent nature [14], [15]. Robbery is usually classified as a felony in jurisdictions, even when many lesser forms of theft are only penalised as misdemeanours. The time and place of the robbery, whether it is armed or not, the kind of weapon used, and the degree of force used are all factors that criminologists consider when classifying various sorts of robberies. Thus, street and commercial robberies are two examples of classic scissor [16]. Robberies on the street often occur in impoverished, congested areas without closed-circuit television systems (CCTV). There are two ways that commercial robberies happen: either the perpetrator dresses up as a client and enters the store, hiding his face with a mask or helmet, or he just walks up and starts pulling a gun and threatening the staff. The other is when violent criminals arrive, usually in groups, and most often cover their faces or heads [17]. Both kinds of business robberies took place in indoor locations that are often equipped with CCTVs to enable the discovery of criminals or even the anticipation of such crimes. Furthermore, criminals who carry a knife or weapon often threaten people physically. However, a large force is more likely to be used against criminals who are unarmed or who are carrying any kind of stick [16], [17]. Hence, forceful commercial robberies with or without weapons result in harm, suffering, and even fatalities.

Predicting the behaviour of commercial robberies, whether by humans, machines, or a mix of these, is crucial in averting their occurrence and associated risks [18].

Generally speaking, there are techniques for automating the detection or prediction of crimes that are based on taking various criminal situations and applying them to various sectors. However, no technique has been able to foresee the possibility of robbery behaviour. Consequently, an algorithm for RBP prediction in video pictures has to be developed. It is evident that in order to make predictions, characteristics and evidence from surveillance footage must be extracted. Investigating the possibility of robbery behaviour in video footage is necessary to do this. The circumstances of each location chosen for robbing and the diverse cultures of the nations involved cause scenarios of robbery occurrence to change from one context to another [19]. As a result, assurance does not accompany strong feature extraction.

Even though there are many different robbery occurrence scenarios, scenario-based techniques [20], [21] allow for the consideration of a common scenario with key aspects for commercial robbery videos. In particular, someone or anything selecting a poorly frequented location, often hiding their face or head with a helmet, mask, glasses, or any other article of clothing to avoid recognition, and waiting for a chance to brandish a weapon, threaten, or use force. This situation perfectly aligns with the concept of the first kind of commercial robbery behaviour as well as the expertise of an individual [17], [22]. We take into consideration similar traits present in the majority of robbery instances under three modules, namely head cover, crowd, and loitering detection, in order to construct a system based on this common scenario abstracted from other situations inferred from robbery footage. An inference machine is required to conclude on the RBP once these characteristics have been extracted for module implementation. For possible derivation, the conclusion method ought to be as competent as human decision making. Fuzzy measurement suggests that fuzzy set theory's capacity to imitate human inference [23] allows experience to be expressed as fuzzy rules, which in turn helps with complicated decision diagnostics and reasoning [24], [25]. However, deep learning techniques lack this flexibility and may not be up to par with the subtleties and fluctuations of ambiguous data [25]. For these reasons, this study proposes a fuzzy inference engine.

In summary, our paper's primary contributions are listed below: 1. The suggested algorithm is predicated on a cutting-edge technique that forecasts RBP and guards against damages brought on by its occurrence inside.As far as we are aware, this is the first study to concentrate on predicting robbery behaviour. It is based on three primary modules: the head cover, crowd, and loitering detection modules.

2. A dataset with and without head covers has been manually tagged and produced for our system. We total the outcomes of the two states that the head cover detecting module reports in order to count the crowd. The technique outperforms security video limitations including poor quality and single camera footage.

3. Our concept of the lingering point offers a fresh approach to the loitering computation. The length of time each individual spends loitering has been calculated using the Deep Simple Online Real-time Tracking (DeepSORT) algorithm in relation to the

tracking techniques. Each one has been given a point by evaluating the quantity of loitering that was discovered using the Euclidean distance computation.

4. Our algorithm's main innovation is the use of a fuzzy inference engine with phases for fuzzification and defuzzification as well as optimised rules. The data from these three modules were analysed using an inference engine and an expert's understanding of robbery behaviour.

The remaining text is formatted as follows: Section II examines some of the literature that is relevant to our work, such as studies that discuss our modules and the prediction or detection of suspicious behaviour. In Section III, the suggested method is explained together with the principles of RBP prediction, along with the suggested modules and how they improve YOLOV5 to produce low-resolution video pictures. Section IV presents and discusses the experimental outcomes. The last part wraps up the findings and outlines next projects.

II. RELATED WORKS

Anomalies are sporadic observations, unusual occurrences, or suspicious behaviours that deviate markedly from the norm. Any behaviour that deviates from a usual activity is considered an anomaly, and that is what crime is [2]. It is possible to argue that the rise in crimes in public spaces is the reason for the widespread deployment of CCTV. Based on the identification of suspicious behaviour, crimes may be anticipated. Prediction requires incomplete, imprecise, and uncertain data [26]. Our suggested methodology focusses on indoor RBP prediction. A kind of crime is robbery, and the suggested algorithm requires the identification of loitering, crowds, and head cover. Providing a general RBP prediction framework—a subject that no other study addresses—is a key component of our method. This part will include a few related works that are pertinent to the detection or prediction of suspicious behaviour, the detection or prediction of crime, and papers that deal with the detection of loitering or head coverings.

In order to update items meant in each frame, Mohannad and Levine [27] suggested a semantics-based suspicious behaviour identification system based on object tracking via blob matching with colour histograms and geographical information. To specify the similarity between blobs and objects, the value of the intersection of the histogram is computed and compared to a predetermined threshold. It then allocates the proper classes, which include objects for lifeless items and humans for living ones. Behaviours are semantically defined by computing their 3D motion attributes and capturing it in the form of historical records. The following suspicious behaviours have been identified: fighting by computing merge, split, and simultaneously movement of blob's centroid; abandoned luggage by background subtraction methods; fainting by comparing assumed 2D and actual 3D location of person's feet and also head coordinates of that person; and eventually loitering by aggregating presence time of a person in an area.

An automatic normal method for problematic pedestrian identification via lingering detection was presented by Ishikawa and Zin [18]. [18] claims that a dubious individual spends a considerable

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

amount of time walking, stopping, and circled the area with an increased number of route changes. His distance value is higher than that of an average individual, and his acceleration changes significantly. In order to apply these characteristics, [18] counts the frequency of block numbers indicating how many feet a person is in each of the 25 blocks that make up the video frames. If the frequency exceeded the threshold, the pedestrian was viewed with suspicion. It computes the angles of motion in order to compute direction changes. All necessary characteristics are extracted by computing the changing distance and acceleration. Ultimately, a decision fusion procedure aggregates the results of every phase to identify pedestrians who seem suspicious. A two-part E-police system—crime prediction and a video surveillance monitoring system—was proposed by Rajapakshe et al. [28].[28] use human activity recognition techniques to identify suspicious behaviours, such as violence and vandalism, and divides them into normal and abnormal categories. They identify suspicious people who hide their faces by using Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) for feature extraction. The crime prediction method of [28] is based on the use of public information resources to anticipate the location and timing of criminal activity using classification algorithms like SVM and Decision tree.

An expert real-time suspicious behaviour detection system for shopping malls was suggested by Arroyo et al. [22]. They use an image segmentation and background removal method to detect things in the foreground. The blobs of each subdivided section are then gathered using a blob fusion technique in order to recognise humans. A novel two-step procedure is employed in conjunction with a tracking algorithm: one using two SVM kernels for occlusion control and one Kalman filter for human identification and tracking. The behaviour of individuals is then examined using the acquired trajectories. The entry or existence alarm is activated when an individual flees or there are too many individuals entering, and it is picked up by trajectory analysis. Additionally, interiors with more costly items that are selected by human security agents are designated danger zones. People's loitering is assessed based on their paths and the amount of time they spend in specific areas. Security experts have stipulated that if the period exceeds 30 seconds, their system would sound an alert. To keep the cash counter safe, they installed a camera on it. An alarm will sound if someone is lingering near it and no store employees are present in the area where the cash desk is. Furthermore, a naturalistic dataset supplied by many cameras at a store's entrance, inside, and pay desk is evaluated.

III. PROPOSED METHODOLOGY

In this section, the proposed method is introduced in detail for prediction of RBP as illustrated in Fig. 1. At the beginning, three main blocks defined, including:

• Head cover detection module.

• Crowd detection module to check number of humans attended in environments.

• Loitering detection module.

Dataset gathered by our group with regards to intention of proposed robbery scenario, to implement head and crowd detection modules. Next, data prepared by manually

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

annotating and convolving to decrease their resolution. By retraining YOLOV5s to customize it, two first modules are completely provided. An Euclidean method is used to calculate distance traveled by human and the DeepSORT algorithm is employed to track him. By defining our individual thresholds, the label of loitering allocated to each person and based on this innovating definition, our particular loitering detection module is presented. Finally, a fuzzy inference machine is used for potential prediction of robbery behavior. The RBP prediction can be decomposed into three main parts: feature extraction, feature analysis and RBP prediction.

TABLE 1. Literature survey based on relevance to our proposed approach.

Reference	Proposed Scheme	Implemented Methods
[18]	suspicious persons detection system by loitering detection	To implement this system, they calculate traveled distance, acceleration and direction changes. Finally, they compare calculated parameters with a threshold they have defined to decide if the person is suspicious.
[22]	An expert video-surveillance system is utilized for detection of real-time suspicious behaviors like loitering around interior or cash desk and running, in shopping malls.	The proposed system detects humans by blob fusion algorithm and track them by Kalman filter and SVM kernels. Analysing of behaviours done by obtained trajectories.
[27]	Real-time detection of suspicious behaviors like loitering in public transport areas	The proposed solution recognize behaviour semantically based on object tracking and assigning a label as object or human for each one. Next, they calculate 3D motion features to detect suspicious behavior
[28]	Suspicious behavior detection and spatio-temporal crime prediction	CNN and RNN methods are used for feature extraction to recognize human activity and detect suspicious behavior. they also use public information resources belong to place and time of crime occurrence for crime prediction.
[29]	Automatic early detection of a suspecious behavior with track-based features in a shopping mall. they detect behaviour of pickpocket by it's scenario implementation.	Feature extraction and Fisher linear discriminant classifier is used to early detect pickpocket. Their features are based on an expert knowledge including: speed of walking, changing of orientation and crowd merging and splitting.
[30]	Suspicious action detection	An ECNN algorithm proposed for suspicious action detection.
[31]	Snatch theft detection model	Their model uses GMM based on common theft scenarios. it also uses HOF and MBH to extract dense trajectory feature sets and create action-vectors to train GMM. .
[32] – [37]	Deep learning algorithms are used for face mask detection.	Retraining CNN, YOLOV5, MobilenetV2, RESNET50, InceptionV3, YOLOV3 and faster R-CNN with two-class images: with or without mask.
[38] – [40]	Approaches based on deep learning algorithms are used to detect helmet.	Proposed solutions detect helmet by retraining YOLOV3, YOLOV5 and SSD MobileNet V2.
[41]	Loitering detection method based on human detection and track.	A YOLOV3 based detector is used for human detection and a DeepSORT algorithm for tracking him. By comparing time duration and displacement of individuals with a defined threshold, loitering is detected.
[42]	Spatio-temporal based detection system for loitering.	A YOLOV3 based detector is used for human detection and a DeepSORT algorithm for tracking him. By comparing time duration and displacement of individuals with a defined threshold, loitering is detected.

FEATURE EXTRACTION

Features extracted from each video, are organized as three main modules based on common scenario of robbery occurrence, including: head cover, crowd and loitering detection modules. In this part we will explain about each one.

1) Head Cover Detection

In our proposed method, head cover is defined as any hat, helmet, mask, glasses or clothes that conceals a person's head and face and disturb their identification. Most of robbers prefer not to be recognized when robbing and head covering is their choice. In this paper, we propose a method which can detect head of human intended in a store with or without head cover in low-resolution single frames and label them as masked or no-mask. The proposed algorithm uses a deep learning method for head cover module. You Only Look Once (YOLO) [43] is a deep learning algorithm that uses to detect objects and views entire image as a regression problem [44]. Within training process of YOLO, it looks over the whole image to extract global information of target and divides the image into S × S grids. This division makes some grid cells and the position of object's center should fall into that grid cell to be labeled as detected object. Resolution of surveillance videos are usually low and head cover detection module should also overcome this challenge.The YOLO algorithms can be retrained by low-resolution images to make an accurate detection result [45]. YOLO v1, v2 and v3 can detect type of an object and its position at an image [46]. YOLOv4 [47] is an improvement on the YOLOv3 in mean average precision(map) and the number of frames per second [48]. YOLOv5 is based on YOLOv4 with higher processing speed and smaller size, even 90% smaller than YOLOV4. Therefore, YOLOv5 is a better choice to be used in an embedded device. Moreover, YOLOv5 has better capability in small objects detection [49]. Due to small size of YOLOV5 and it's ability for small objects detection, we choose YOLOV5 for our proposed method and retrain it with our prepared low resolution images. The YOLOv5 algorithm regulates the width and depth of the backbone network. Beside these reasons, high human and object detection

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

accuracy of YOLOv5, make it a good choice to be used in head cover detection for our proposed method. Therefore, it has four versions of the model, which are YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x. YOLOv5s is the simplest, smallest and fastest version [50] and we use this version in our method.

2) Crowd Detection

People attended in the store, either have head cover or not. As a consequence, by considering the results obtained from the head cover detection module, the number of people can be obtained and the amount of crowded environment can be commented on. Therefore, crowd detection module is actually human's head part detection acquired from head cover detection module. Number of humans attended in store changes the potential of robbery in stores. Low attendance makes higher risk for robbery. As human presence increases, the robbery potential decreases as defined in Eq. 1 and 2.

$$C = \left\{ 0.1, \frac{2}{\{N\}} \right\} \quad (1)$$

$$C^s = \begin{cases} 100 \times MinC & \text{N=1 or N} \geq 7 \\ 100 \times MaxC & 1 < N < 7 \end{cases} \quad (2)$$

Let C be crowd corresponding set, N be the number of people and C s be crowd score. As a result, C s is between 0-100 and showing the number of people present in the stores. Based on crowd score we defined, the effect of the presence of people can be considered on RBP.

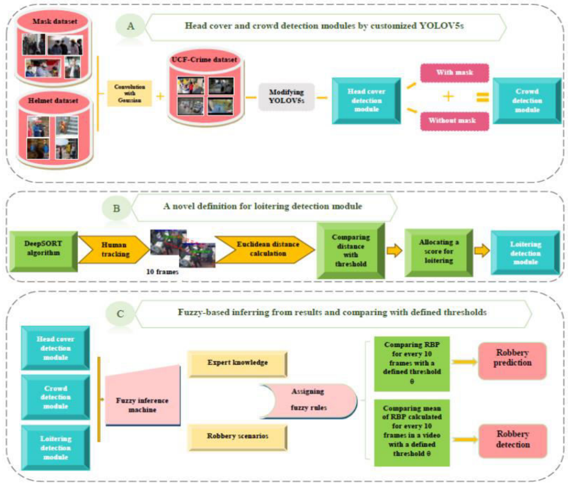

FIGURE 1. The block diagram of robbery behavior prediction and detection system. After collecting 7622 images from three different datasets and preparing them for YOLOV5s retraining, head cover and crowd detection modules are implemented (block A). Loitering detection module uses an Euclidean method to calculate distance traveled by human and DeepSORT algorithm is employed to tracck the human.Eventually, the label of loitering allocated to each person by defining our individual thresholds (block B). A fuzzy inference machine concludes modules results with the help of fuzzy rules to compute robbery potential for prediction and detection (block C).

IV. RESULTS AND DISCUSSION

Our empirical evaluation is on two cases predicting robbery potential and detecting robbery. Preparing proper dataset for retraining deep learning algorithms is for modules created to extract individuals and their head cover beside their loitering. To test our proposed system, 45 videos are eligible from Robbery-UCF-Crime dataset for prediction process. In videos falling into predicting state, there is a period of time that robber loiters for a appropriate moment for

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

robbery. It is crucial to accentuate that prediction signifies for behavior which is not happened yet. For robbery behavior, prediction is feasible up to the moment that the robber shows his weapon, threat or force. While individuals with head cover loiter to get a proper opportunity for threatening or forcing with robbery aim in low crowded shopping malls. On the other hand, 70 videos of Robbery-UCF-Crime dataset are used for robbery detection algorithm. It is noticeable that the detection is the process of finding out the behavior after its occurrence which for robbery behavior it means that the robber shows his weapon, threat or force and maybe leaves the place after robbing something. The 70 videos have condition of potential calculation such as proper camera angle for cash desk sight and visible human. For robbery potential calculation, a fuzzy inference machine is used instead of human brain to allocate a proper potential to each scene of video based on three modules outputs. This part renders experimental results to evaluate our proposed algorithm.

DATA PREPARATION

To develop improved YOLOv5s for head cover detection in low-resolution images, we prepared image dataset which gathered from three groups: video image sequences collected from anomaly folders of UCF-Crime dataset except robbery folder [1], Bikes Helmets Dataset 1 , and Mask Dataset 2 .

Thereupon images with proper view angle of CCTV cameras, with more differs in human position and images in which human can be detached from background, are selected. Prepared dataset includes 7621 images comprises 5129 images from Bikes Helmets Dataset, 254 images from Mask Dataset and 2238 images are from UCF-Crime dataset videos except videos from robbery folder, which converted to image sequences. These images divided into training, validation and test sets with approximate ratio of 7:2:1 respectively. Table 2 represents dataset in detail.

TABLE 2. Images from different dataset used for retraining YOLOV5 in our proposed approach.

Dataset	UCF-Crime	Bikes Helmets	Mask
Number of images for training	1639	3762	187
Number of images for validation	449	1025	51
Number of images for test	150	342	16

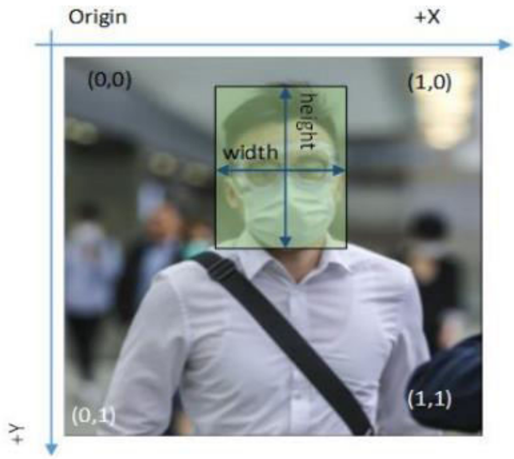

FIGURE 2. Coordinates position belongs to bounding box of human head part.

FIGURE 3. A sample image from prepared dataset, Left: main image - Right: after convolving with Guassian filter of Eq.

TABLE 3. Comparison of different model complexity met

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Model	Parameters (M)	FLOPs (G)
YOLOV5s	7.57	15.98
Retrained YOLOV5s	7.22	0.601

TABLE 4. System configuration for retraining YOLOV5s in our proposed method.

Hardware and software	NVIDIA GeForce RTX 2060
Operating system	Windows 10 Pro 64-bit
CPU	Intel® Core™ i7-9750H, CPU@2.60GHz
GPU	CUDA 10.2
RAM	16GB
Deep Learning algorithm	YOLOV5

V. CONCLUSION

An technique for RBP prediction in video surveillance pictures is presented in this research effort. CCTV footage presents a number of difficulties, including varying methods for robbery incidents, variations in camera angles installed in various locations, and poor quality video pictures captured by CCTVs. By removing these barriers, prompt action is taken and robberies that are entirely or partly visible on security camera footage are avoided. This study is being done because, despite the importance of preventing robberies from happening, no RBP prediction has ever been made, according to our thorough literature assessment. With the assistance of professional commentary and after viewing several CCTV robbery recordings, we are able to extract some typical situations of robbery occurrence. We examine these situations in order to identify additional characteristics that they have in common and to put into practice a useful method for RBP prediction. Our work suggests a deep learning-based method to determine the likelihood for robbery with the use of a fuzzy inference machine. This method collects appropriate datasets of people wearing or not wearing head coverings, which retrains the YOLOV5 algorithm. The crowd and head cover detection modules are implemented effectively by means of this deep learning based approach. This work also implements our described approach for the loitering module, which uses the DeepSORT technique to compute the Euclidean travelled distance of people. Based on three module outcomes, a fuzzy inference machine is constructed to infer the robbery potential of films every ten frames and average them for each clip. The F1-score of the suggested system is 0.537, and it is applied to the Robbery folder of the UCF-Crime dataset. This result demonstrates that for over half of the movies, our suggested technique can accurately forecast the likelihood of a heist.

As such, we shift the prediction issue to the robbery detection problem. As a result, we can contrast it with earlier research that focused on anomaly detection, particularly robbery detection, using the UCF-Crime dataset. The detection method's F1-score is 0.607, which is the highest of all the approaches. The outcome demonstrates that our suggested scenario-based method operates accurately and has a high degree of proficiency in both identifying and forecasting robbery behaviour. Anywhere that wants to stop robbery crimes and has security cameras may apply our suggested strategy. They don't need to hire someone to carefully examine the in-the-moment footage from these cameras and determine if a heist is likely to occur. To avoid making a mistake, however, this individual should watch the films through to the end. Furthermore, anybody may modify our methods in private by altering the threshold values based on cultural differences.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

By increasing the accuracy of loitering detection, we can raise the F1-score. By enhancing the DeepSORT technique, we want to produce a tracking algorithm for low-resolution video pictures in the future. Low-resolution movies do not allow for exact human detection and tracking. This is as a result of FrRCNN serving as the DeepSORT algorithm's detector. Consequently, we will modify the DeepSORT algorithm's detection architecture in order to retrain YOLOV5 using low-resolution human photos. Low resolution photos will be the sole object type in the proposed YOLOV5.

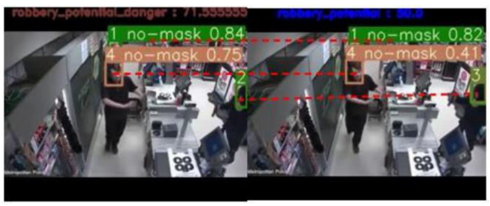

FIGURE 4 . Missing ID number sample two frames. ID= 2 and ID=3 are for the same person.

REFERENCES

1) Sultani, Waqas, Chen Chen, and Mubarak Shah. "Real-world anomaly detection in surveillance videos." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 6479-6488. 2018.

2) Suarez, Jessie James P., and Prospero C. Naval Jr. "A survey on deep learning techniques for video anomaly detection." arXiv preprint arXiv:2009.14146 (2020).

3) Mei, Tao, and Cha Zhang. "Deep learning for intelligent video analysis." In Proceedings of the 25th ACM international conference on Multimedia, pp. 1955-1956. 2017.

4) Yan, He, Xueliang Liu, and Richang Hong. "Image classification via fusing the latent deep cnn feature." In Proceedings of the International Conference on Internet Multimedia Computing and Service, pp. 110-113. 2016.

5) Ghazal, Mohammed, Carlos Vázquez, and Aishy Amer. "Real-time automatic detection of vandalism behavior in video sequences." In 2007 IEEE International Conference on Systems, Man and Cybernetics, pp. 1056- 1060. IEEE, 2007.

6) Febin, I. P., K. Jayasree, and Preetha Theresa Joy. "Violence detection in videos for an intelligent surveillance system using MoBSIFT and movement filtering algorithm." Pattern Analysis and Applications 23, no. 2 (2020): 611-623.

7) Lao, Weilun, Jungong Han, and Peter Hn De With. "Automatic videobased human motion analyzer for consumer surveillance system." IEEE Transactions on Consumer Electronics 55, no. 2 (2009): 591-598.

8) Ferguson, Andrew Guthrie. "Predictive policing and reasonable suspicion." Emory LJ 62 (2012): 259.

9) Beck, Charlie, and Colleen McCue. "Predictive policing: what can we learn from Wal-Mart and Amazon about fighting crime in a recession?." Police Chief 76, no. 11 (2009): 18.

10) Bowers, Kate J., and Shane D. Johnson. "Who Commits Near Repeats? A Test of the Boost Explanation." Western Criminology Review 5, no. 3 (2004).

11) "Seattle Police Department," SPD progress report, 2021. [Online]. Available: https://www.seattle.gov/Documents/Depart

ments/Police/Reports/2021_SPD_ CRIME_ REPORT_FINAL.pdf.

12) ”FBI,” 2019. [Online]. Available: https://ucr.fbi.gov/crime-in-theu.s/2019/crime- n-the-u.s.-2019/topic-pages/robbery.

13) Fawei, Biralatei, Jeff Z. Pan, Martin Kollingbaum, and Adam Z. Wyner. "A semi-automated ontology construction for legal question answering." New Generation Computing 37, no. 4 (2019): 453-478.

14) Thompson, Rebecca. Understanding theft from the person and robbery of personal property victimisation trends in England and Wales, 1994–2010/11. Nottingham Trent University (United Kingdom), 2014.

15) Cook, Philip J. "Robbery violence." J. Crim. L. and Criminology 78 (1987): 357.

16) McCluskey, John D. "A comparison of robbers’ use of physical coercion in commercial and street robberies." Crime Delinquency 59.3 (2013): 419- 442.

17) Luckenbill, David F. "Patterns of force in robbery." Deviant Behavior 1.3-4 (1980): 361-378.

18) Ishikawa, Tatsuya, and Thi Thi Zin. "A study on detection of suspicious persons for intelligent monitoring system." In International Conference on Big Data Analysis and Deep Learning Applications, pp. 292-301. Springer, Singapore, 2018.

19) Medel, Jefferson Ryan, and Andreas Savakis. "Anomaly detection in video using predictive convolutional long short-term memory networks." arXiv preprint arXiv:1612.00390 (2016).

20) Shah, Mubarak, Omar Javed, and Khurram Shafique. "Automated visual surveillance in realistic scenarios." IEEE MultiMedia 14, no. 1 (2007): 30- 39.