**PROPOUNDING FIRST ARTIFICIAL INTELLIGENCE APPROACH FOR PREDICTING ROBBERY BEHAVIOR POTENTIAL IN AN INDOOR SECURITY CAMERA**

**ABSTRACT**

Crime prediction in video-surveillance systems is required to prevent incident and protect assets. In this sense, our article proposes first artificial intelligence approach for Robbery Behavior Potential (RBP) prediction and detection in an indoor camera. Our method is based on three detection modules including head cover, crowd and loitering detection modules for timely actions and preventing robbery. The two first modules are implemented by retraining YOLOV5 model with our gathered dataset which is annotated manually. In addition, we innovate a novel definition for loitering detection module which is based on Deep SORT algorithm. A fuzzy inference machine renders an expert knowledge as rules and then makes final decision about predicted robbery potential. This is laborious due to: different manner of robber, different angle of surveillance camera and low resolution of video images. We accomplished our experiment on real world video surveillance images and reaching the F1-score of 0.537. Hence, to make an experimental comparison with the other related works, we define threshold value for RBP to evaluate video images as a robbery detection problem. Under this assumption, the experimental results show that the proposed method performs significantly better in detecting the robbery as compared to the robbery detection methods by distinctly report with F1-score of 0.607. We strongly believe that the application of the proposed method could cause reduction of robbery detriment in a control center of surveillance cameras by predicting and preventing incident of robbery. On the other hand, situational awareness of human operator enhances and more cameras can be managed.

**INTRODUCTION**

Today, surveillance cameras are widely used in various places such as stores, banks, airports and homes, to increase public safety and prevent the occurrence of crime. Alternatively, the time and place of the crime and specifically the wrongdoer, can be achieved by analyzing these videos and aiming to identify the delinquent. Meanwhile, someone is needed behind the scene, watching the videos and noticing whenever something anomaly is happening. However, due to very rare occurrence of an anomaly, the person becomes tired and if an anomaly happens, sometimes he cannot realize its occurrence. In other words, he loses the anomaly [1]. Furthermore, the anomaly-detection process is based on human common feeling which is learned during years. On the other hand, skill amount of the person for signs of crime occurrence understanding ability and the cost of employing him are other problems of non - automated crime prediction and detection systems which are based on watching surveillance videos.

To automate anomaly detection, some visual features must be extracted using machine learning and deep learning algorithms [2], [3]. For better performance of these algorithms, specific features for different anomaly classes [4] like vandalism [5], violence detection [6] and robbery [7] can be useful.Predicting the location and time of the crime reducing the destruction. On the other hand, security forces are also present on time such as an experiment, manufactured in Santa Cruz, California, where officers benefit from daily crime forecasts every morning. This forecasting navigates them to patrol determined regions. A Santa Cruz spokesperson declared that thirteen wrongdoers have been stopped in the determined areas during first the six month of experiment [8].Due to paper [8], [9], and [10], some main symptoms prove that predictive policing is significant to be used for federal financing and security systems including: cost saving and crime reduction. Violent crimes are more dangerous because of their victimization probability and they increased by 20%due to Seattle Police Department (SPD) report during 2021 in Washington, USA [11]. According to statics acquired from Federal Bureau of Investigations-Uniform Crime Reporting System (FBI-UCR), robbery is one of five common crimes in the United States [12]. The detection of robbery is one of the purposes of installing surveillance cameras in many places. Robbery is the crime of taking or attempting to grab any property by force, threat or weapon [13] based on oxford dictionary definition and differentiated from other forms of theft such as shoplifting, pickpocket or burglary, by its intrinsically violent essence [14], [15]. Whiles many lesser types of theft are punished as misdemeanors, robbery is always a felony in jurisdictions. Criminologists distinguish different types of robbery with regards to time and space of occurrence, armed or unarmed robberies, weapon types and force amount. Therefore, one typical scissor is commercial robberies and street robberies [16].Street robberies usually happens in poor crowded locations with no Closed-circuit televisions (CCTV). Commercial robberies occur in two ways: one where the offender enters the scene dressed up as a customer or conceal his face with normal covers like mask or helmets then suddenly out of the blue pulls a weapon and scares the employee. The other which offenders enter with force, typically in a group and probably conceal their face or head [17]. Both types of commercial robberies occurred in the indoor places which have customarily CCTVs so that detecting offenders or detection and even prediction of commercial robberies can be possible. Additionally, offenders who armed by weapon or knife usually threaten human with force. On the other hand, for offenders bearing any stick or be unarmed, a massive force is more probable [16], [17]. Hereupon, armed or unarmed commercial robberies force causes injury, pain and even death. Thus, predicting commercial robbery behavior by human, machine or combination of these two, plays an important role in preventing its occurrence and its arisen dangers [18].

In general, there are some methods to automate detection or prediction of crimes based on extracting different crime scenarios and implementing them in different fields. But none of these methods have predicted the potential of robbery behavior. Therefore, there is a need to develop an algorithm for RBP prediction in video images. One could easily notice that, extracting the evidences and features in the surveillance videos is needed for prediction. To do this, the potential of robbery behavior in video images should be investigated. Scenarios of robbery occurrence, vary from one context to another [19] due to different conditions of each place selected for robbing and different cultures of countries. Therefore, robust feature extraction is not accompanied with certainty.

Despite the variety of robbery incidence scenarios and due to scenario-based approaches [20], [21], a common scenario with main points can be considered for commercial robbery videos. Specifically, one or some person choosing a poorly attended place who are usually covering their face or head by helmet, mask, glasses or any garment to not be recognized and they are loitering to get an opportunity for showing their weapon, threat or force. This scenario is completely matched with the knowledge of an expert person and definition of first type of commercial robbery behavior [17], [22]. To implement a system based on this common scenario abstracted from different scenarios inferred from robbery videos, we consider common features found in most robbery cases under three modules including: head cover, crowd and loitering detection. After extracting these features, for modules implementation, an inference machine is needed to conclude on the RBP. The conclusion process must be as competence as a human decision making for potential derivation. Due to the ability of fuzzy set theory to mimic human inference [23], experience could be put in the form of fuzzy rules and according to fuzzy measurement, it facilitates the diagnosis and reasoning of a complex decision [24], [25]. Deep learning methods on the other hand, do not offer such adaptability and may not be able to deal with the nuances and variations of uncertain data well [25]. Owing to these reasons a fuzzy inference machine is proposed in this paper.

To sum up, main contributions of our paper are as below:

1. The proposed algorithm is based on a novel method

which can predict RBP and prevent damages resulted by its occurrence in indoor \ places. To the best of our knowledge, his is the first work focusing on robbery behavior prediction and grounded on three main modules: Head cover, crowd and loitering detection modules.

2. A dataset has been prepared for our system and annotated manually as two states: with or without head cover. For crowd counting, we sum the results of two states reported by head cover detection module. The method dominates the constraints of surveillance videos such as low resolution and single camera videos.

3. The loitering point we have defined, is a novel definition for loitering calculation. A Deep Simple Online Real-time Tracking (Deep SORT) algorithm has used with respect to the tracking methods to calculate amount of loitering for each person. By analyzing the obtained amount of loitering based on Euclidean distance calculation, a point has assigned to each one.

4. The key contribution of our algorithm is using a fuzzy inference machine with optimized rules, fuzzification and defuzzification steps. Obtained results of these three modules analyzed based on an expert person knowledge about robbery behavior and an inference machine. The rest of paper is arranged as follows: Section II reviews some literature related to our work including suspicious behavior prediction or detection and also papers related to our modules. Section III explains proposed algorithm and outlines concepts of RBP prediction, the proposed modules and outcome to low-resolution video images by improving YOLOV5. Experimental results are presented and discussed in section IV. The last section concludes the research work and presents future works.

**SYSTEM ANALYSIS**

**3.1 EXISTING SYSTEM**

Anomalies are infrequent observations, events or behaviors which are suspicious because they are significantly different from normal patterns. Crime is a kind of an anomaly which is any behavior deviating from a normal activity [2]. One could say that the proliferation usage of CCTV has been because of increasing crimes in public places. Crime can be predicted according to suspicious behavior detection. Prediction needs defective, vague and unsure information [26]. Our proposed approach concentrates on RBP prediction in indoor places. Robbery is a kind of crime and the proposed algorithm needs loitering, crowd and head cover detection. One important concept of our algorithm is providing a generic RBP prediction framework which is not addressed in any other paper. In this section, we will discuss about some related works relevant for suspicious behavior detection or prediction, crime detection or prediction and articles concerning with the loitering or head cover detection.

Elhamod and Levine [27] proposed a semantics-based suspicious behavior recognition algorithm based on object tracking by blob matching with color histograms and spatial information, for updating objects intended in each frame. For blob and objects similarity specification, intersection of histogram’s value is calculated and compared with the defined threshold. Next it assigns appropriate classes contains people for animated and objects for inanimated things. By calculating their 3D motion features and recording it in the form of historical records, behaviors are semantically determined. detected suspicious behaviors include: 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, fighting by computing merge, split and simultaneously movement of blob’s centroid and eventually loitering by aggregating presence time of a person in an area.

Ishikawa and Zin [18], introduced an automated normal system for questionable pedestrian detection by loitering detection. According to [18], a questionable person walks, stops and goes around the location repeatedly for a long time with enhancement of direction changing number. His distance value is greater than the normal person and changing in acceleration is so much. To implement these features, [18] divides the video frames into 25 blocks and counts the frequency of block numbers which feet of person are in that location. if this frequency was more than threshold value, that person descent as suspicious pedestrian. To compute changing of direction, it calculates angles of moving direction. Computing of distance and acceleration changing extracts all needed features. Finally, a decision fusion process detects suspicious pedestrians by aggregating the scores of each step.

Rajapakshe et al. [28] presented an E-police system which contains two components: video surveillance monitoring system and crime prediction.To detect suspicious behaviors such as violent and vandalism, [28] uses human activity recognition methods and classifies them into normal and abnormal categories. They use Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) for feature extraction and detect suspicious human who conceal their faces. Crime prediction process of [28] is based on public information resources for place and time of crime occurrence prediction with the help of classification algorithms such as SVM and Decision tree.

Arroyo et al. [22] proposed an expert real-time suspicious behaviors detection system in shopping malls. They locate foreground objects by an image segmentation and background subtraction algorithm. Next, a blob fusion algorithm is used to gather the blobs of each segmented parts to detect human. A tracking algorithm is used with the help of a new two-step method: a) using a Kalman filter for detection and tracking human, b) SVM kernels for occlusion management. Then, the obtained trajectories of people are used to analyze human behaviors. The entrance or existence alarm is for the time that too many people enter or a person runs away and it is detected by trajectory analyzing. Moreover, specific risk areas are interiors with more expensive articles and chosen by the human security officers. Loitering of people is evaluated according to their trajectories and the length of time they present in those zones. If the time be more than 30 seconds, which has specified by security experts, their system gives off an alarm. They mounted a camera on the cash desk to protect it. If someone loiters around it, and no shop personnel attended in the cash desk zone, an alarm gives off. Additionally, evaluation process is done on a naturalistic dataset they provided by multi cameras located on entrance, interior and cash-desk of a shop.

**DISADVANTAGES**

• The complexity of data: Most of the existing machine learning models must be able to accurately interpret large and complex datasets to detect Robbery Behavior.

• Data availability: Most machine learning models require large amounts of data to create accurate predictions. If data is unavailable in sufficient quantities, then model accuracy may suffer.

• Incorrect labeling: The existing machine learning models are only as accurate as the data trained using the input dataset. If the data has been incorrectly labeled, the model cannot make accurate predictions.

**3.2 PROPOSED SYSTEM**

1. The proposed algorithm is based on a novel method which can predict RBP and prevent damages resulted by its occurrence in indoor places.To the best of our knowledge, this is the first work focusing on robbery behavior prediction and grounded on three main modules: Head cover, crowd and loitering detection modules.

2. A dataset has been prepared for our system and annotated manually as two states: with or without head cover. For crowd counting, we sum the results of two states reported by head cover detection module. The method dominates the constraits of surveillance videos such as low resolution and single camera videos.

3. The loitering point we have defined, is a novel definition for loitering calculation. A Deep Simple Online Real-time Tracking (DeepSORT) algorithm has used with respect to thetracking methods to calculate amount of loitering for each person. By analyzing the obtained amount of loitering based on Euclidean distance calculation, a point has assigned to each one.

4. The key contribution of our algorithm is using a fuzzy inference machine with optimized rules, fuzzification and defuzzification steps. Obtained results of these three modules analyzed based on an expert person knowledge about robbery behavior and an inference machine.

**ADVANTAGES**

• 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 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.

**SYSTEM REQUIREMENTS**

**4.1 FUNCTIONAL REQUIREMENTS**

* Servicer
* Telecommuter

**4.2 NON FUNCTIONAL REQUIREMENTS**

**4.2.1 HARDWARE REQUIREMENTS**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**4.2.2 SOFTWARE REQUIREMENTS**

* Operating system : Windows 7 Ultimate.
* Coding Language : Python.
* Front-End : Python.
* Back-End : Django-ORM
* Designing : Html, css, javascript.
* Data Base : MySQL (WAMP Server).

**SYSTEM ARCHITECTURE**



**MODULES**

* Servicer
* Telecommuter