International Journal of Computer Sciences and Engineering
Vol.**11**, Special Issue.**1**, pp.**101-106**, November **2023**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org
ISSN: 2347-2693 (E)

Research Paper

Crime Scene classification of Digital Images Sourced from CCTV Footage

Saumitra Biswas[1], Sanchayan Bhaumik[2*], Tanay Bag[3]

[1]Digital Forensics/Faculty, ISOAH Data Securities, Kolkata, India
[2]Computer Application, Assistant Professor, MAKAUT, Kolkata, India
[1]Digital Forensics/Faculty, ISOAH Data Securities, Kolkata, India

Corresponding Author: sanchayan@gmail.com

Abstract: CCTV image classification using machine learning algorithm is a novel way of using machine learning for security. This paper investigates how classification can be done using statistical indicators of image pixels without employing extensive image processing techniques. It further performs comparative study of performances of some commonly used classification algorithms in classifying images. The highest performance was by the K-Nearest Neighbor classifier with accuracy, precision, and recall scores of 95\%, 90\%, and 100\% respectively.

Keywords: CCTV image classification, Crime Scene Identification, CCTV Footage Analysis, Machine Learning, CCTV Security, Classification Algorithm.

1. Introduction

Close Circuit TV or CCTV has become a common accessory for security in contemporary society. It provides a relatively low cost solution to the problem of holistic security by keeping an "eye" on things at factories, malls, schools, banks, offices, rail stations, airports, homes and even on public streets, twenty four hours a day, seven days a week. One shortcoming of it is that human intervention is needed to understand what is happening or what has happened. In these days of machine learning and artificial intelligence, it is natural that machine learning techniques will be used to analyze content captured by CCTV to understand what is happening and what can be done with that understanding.

In the current study, we aim to identify a frame that captures a crime scene from a pool of frames captured by a CCTV. For example, if a robbery occurs at a bank, law enforcement agencies have to go through hours of video to identify the perpetrators. Identifying the correct frame(s) is one of the vital steps in the process. Here our goal is to investigate the effectiveness of machine learning algorithms to identify the frame-of-interest from a pool of thousands or even hundreds of thousands of frames. If one such model can be built and deployed, it may increase the security provided by the CCTV mechanism by multiple folds in a cost effective manner.

Before a problem statement can be formulated, two questions needed to be answered. At first, we need to decide what type of machine learning it is going to be – Regression, Classification, or Clustering. Here we may consider a video footage as a series of frames, which are nothing but still images. Some of those frames are of interest, say, belonging to Class-A and some of them are not, say, Class-B. A machine learning algorithm's task will be to learn about these two categories of frames from a given dataset and predict an untagged, previously unseen frame as belonging to one category or the other. The observations made above implies that we can formulate the problem as a supervised machine learning problem of binary classification type. Secondly, we have to decide about the nature of data that will dictate whether a frame belongs to a class or not. For this we decided to pick crime as the differentiating factor. The definition of crime covers a wide range of human activities with fluid boundaries, because penal law changes from country to country and sometimes even from state to state within a country. To formulate our problem, we rigorously defined our classes – A. If some video frame captures incidences of murder, shooting, and physical assault then that frame belongs to Crimes (Class – A), and B. All frames that does not belong to Class – A, belongs to Non-Crimes (Class – B). We can now formulate our problem statement as follows – Classification of digital images extracted from CCTV video footages by using Machine Learning algorithm.

It is improbable, if not impossible to choose a particular machine learning algorithm for our purpose without performing a comparative analysis between multiple algorithms. We chose the following algorithms for comparisons. A. Naive Bayes Classifier B. Logistic Regression C. Support Vector Machine D. Decision Tree E. Random Forest F. AdaBoost.

2. Related Work

A plethora of research works exist in these domain. Busarin Eamthanakul, Mahasak Ketcham, Narumol Chumuang [1] investigated how traffic situation can be classified into categories like Flow, Heavy, and Jammed from CCTV data by digital image processing techniques and how the results can be used for traffic planning or traffic light control. Manu Y M, Ravikumar G K, Shashikala S V [2] investigated how CCTV can be connected to fire sensor made out of Raspberry Pi, how the images can be used to detect fire and how the whole system can be used as an early warning system. Pattana Intani, Teerapong Orachon [3] investigated processing of CCTV images and sound for the purpose of intrusion detection, and identify fighting scenarios and report the incidence to the authorities using some form of signal. Pratishtha Gupta, Manisha Rathore, G N Purohit [4] investigated how CCTV data can be processed to identify vehicles going straight (correct) or deviating (incorrect) at a street junction. C.O. Conaire; N.E. O'Connor; E. Cooke; A.F. Smeaton [5] created a dynamic vision system by combining data from thermal infrared camera and CCTV. The system can be used to detect and track objects. Agata Chmielewska, Parzych Marianna, Tomasz Marciniak, Adam Dabrowski, Przemyslaw [6].

Created automated heat map generation system from CCTV data, which can be used to track moving objects. Yan Yang, Brian C. Lovell, Farhad Dadgostar [7] researched how CCTV data can be stored efficiently on RDBMS based on the meaning of the images and videos so that CBIR (Content Based Image Retrieval) and CBVR (Content Based Video Retrieval) is possible. Maleerat Sodanil, Chalermpong Intarat [8] investigated uses of homomorphic filtering techniques for enhancing CCTV images. Naomi Harte, Andrew Rankin, Gary Baugh, Anil Kokaram [9] created automatic detection of illegal dumping at recycle centres, that uses Bayesian learning techniques for CCTV image segmentation. Mohamad Hanif Md Saad, Aini Hussain [10] designed and developed intelligent system for anomalous behavior of humans by identifying and tracking humans and non-humans from CCTV live streams. Rai Purnama Rizki, Eki Ahmad Zaki Hamidi, Lia Kamelia, Ramdani Wahyu Sururie [11] created a security system for smart homes. If a person tries to enter the house by using wrong password, the system detects the face and sounds an alarm. The system performs face detection by analyzing images captured by the camera using Principal Component Analysis (PCA). Hirofumi Noguchi, Takuma Isoda, Seisuke Arai [12] created a system for identifying a person across multiple camera captures, enabling real-time search for a person. Boris A. Alpatov, Pavel V. Babayan, Maksim D. Ershov [13] made CCTV intelligent by incorporating image and video processing analytical logic right into an embedded system that is part of the CCTV hardware platform. Aliia Khasanova, Alisa Makhmutova, Igor Anikin [14] researched the application of deep learning techniques to reduce noises from the digital images and videos. Guillermo Casanova, Daniel Yandún, Graciela Guerrero [15] researched and developed a system that can recognizes faces in images captured by surveillance camera. Muhammad Aseer Khan [16] developed a system for detecting movements in live video feed.

3. Theory

Before a problem statement can be formulated, two questions needed to be answered. At first, we need to decide what type of machine learning it is going to be – Regression, Classification, or Clustering. Here we may consider a video footage as a series of frames, which are nothing but still images. Some of those frames are of interest, say, belonging to Class-A and some of them are not, say, Class-B. A machine learning algorithm’s task will be to learn about these two categories of frames from a given dataset and predict an untagged, previously unseen frame as belonging to one category or the other. The observations made above implies that we can formulate the problem as a supervised machine learning problem of binary classification type. Secondly, we have to decide about the nature of data that will dictate whether a frame belongs to a class or not. For this we decided to pick crime as the differentiating factor. The definition of crime covers a wide range of human activities with fluid boundaries, because penal law changes from country to country and sometimes even from state to state within a country. To formulate our problem, we rigorously defined our classes – A. If some video frame captures incidences of murder, shooting, and physical assault then that frame belongs to Crimes (Class – A), and B. All frames that does not belong to Class – A, belongs to Non-Crimes (Class – B). We can now formulate our problem statement as follows – Classification of digital images extracted from CCTV video footage by using Machine Learning algorithm.

It is improbable, if not impossible to choose a particular machine learning algorithm for our purpose without performing a comparative analysis between multiple algorithms. We chose the following algorithms for comparisons. A. Naive Bayes Classifier B. Logistic Regression C. Support Vector Machine D. Decision Tree E. Random Forest F. AdaBoost

4. Experimental Method

Before any analysis can begin, there has to be data upon which the analysis can done. So, the initial task was to decide how the data can be acquired and what kind of preprocessing can be done with the data before sending them to any kind of Machine Learning algorithm. The quality of the data will determine the performance of the algorithm. In the current scenario, there were no ready-made dataset available. One approach that could have been taken was to use datasets like COCO which provides large number of images of various objects and train our model on that. This approach was discarded since our intent was to let algorithms learn and identify crimes from images available through CCTV cameras. Such images will vary in resolution, number of pixels and size in media, and many other factors. The goal was to present such heterogeneous set of images to the algorithms and train them. To create such an image dataset, CCTV footage available freely over the Internet was sourced. Every such footage was viewed and screenshots were taken when the crime is happening and when it was not happening. The screenshots were then divided into separate directories representing two classes. Random samples were taken from them to create training and testing set of the data. As for the preprocessing part, some of the common approaches are – 1) convert the image into gray scale, 2) use image addition to remove noises from the image (assuming the noises are from some Gaussian distribution that is independent of the source image), 3) use various masks (predefined) and image multiplication/division to extract only those parts of the image that are of interest, 4) use filters to extract the set(s) of pixels that are correlated to the class labels or stores most information about the class labels and use them as features in the later analysis. Although all these approaches are valid, time tested and novel, in the current study, for initial analysis, we have taken a different approach. All color channels were retained for every image. Three matrices representing three color channels were squashed into various representative statistics like mean, standard deviation, skewness, kurtosis and so on, and then use those as representative features. One obvious question arises that how valid these measurements will be, when used as features.

5. Results and Discussion

As mentioned previously, statistics were computed for every color channel of every image and stored in a Pandas dataset. Every row of the dataset represented measurements for a single image. The dataset was stored as a comma separated value file in the secondary media. For training and testing purposes two such files were created.

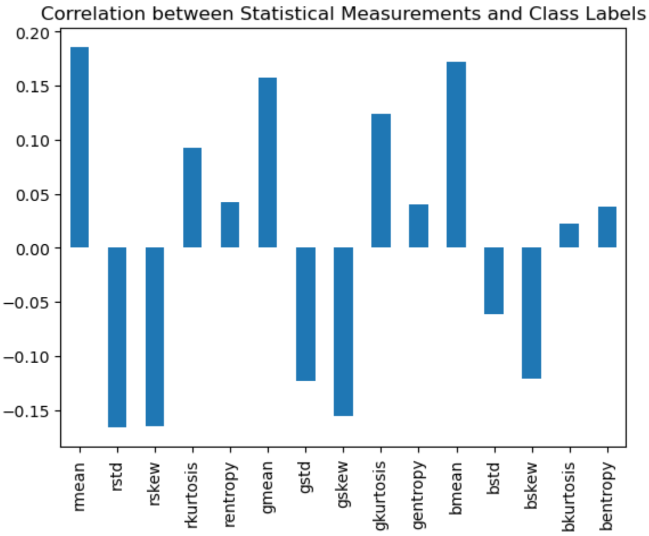

Figure 1. Correlation between statistical measurement and class labels

The next step was to investigate whether these measurements can be used as representative features. For this purpose, correlation coefficients and information gain of these statistics with respect to class labels were measured. The visualization of the results are as follows.

From the figure 1. It can be observed that each of these statistical measurements have some amount of correlation (positive/negative) with the class labels.

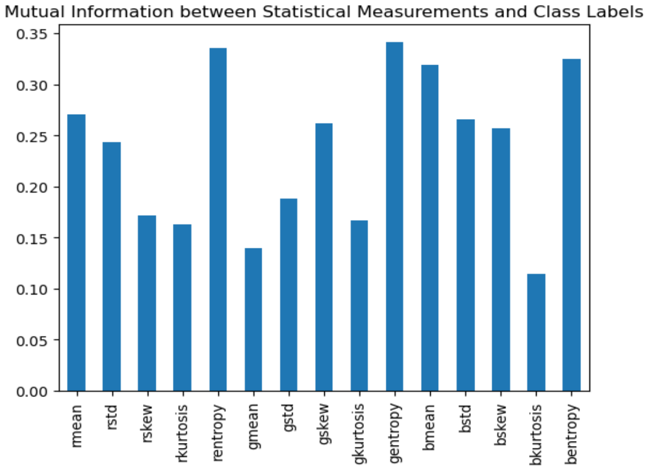

Figure 2. Mutual information between statistical measurements and class labels

From the figure 2. It can be observed that each of these statistical measurements store a fair amount of information about the class labels. From the observations made above, it is safe to say that the statistical measurements have the ability to capture the variability of the class labels. Hence, we proceed to present the data to machine learning algorithms. It is impossible to know which algorithm will provide optimal performance for the given dataset before applying the algorithms to the data. This directs us to choose from a bokeh of commonly used classification algorithm. The algorithms that were chosen for this problem statement are 1) Naive Bayes, 2) Logistic Regression, 3) Support Vector Machine (SVM), 4) K-Nearest Neighbors (KNN), 5) Decision Tree, 6) Random Forest, and 7) AdaBoost.

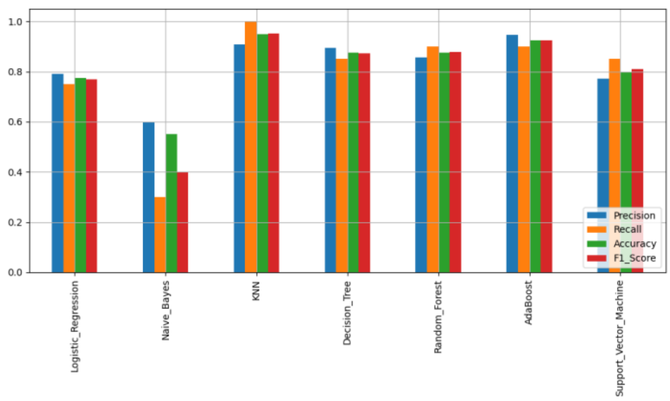

Figure 3. Precision Recall Accuracy and F1 Score of the different Machine Learning Models.

The strategy was to apply all these algorithms with their default hyper parameter settings to set a baseline for future analysis. The results that were received are after training these algorithms on training dataset and applying them to the test dataset are visualized in figure 3.

From the figure 3. It is clear that K-Nearest Neighbor along with ensemble methods like Random Forest and ADABoost have performed very well. All these algorithms have shown quite good balance between Precision, Accuracy and Recall measurements also, indicating balanced performances.

An inspection of the ROC curves produces the following figures 4.

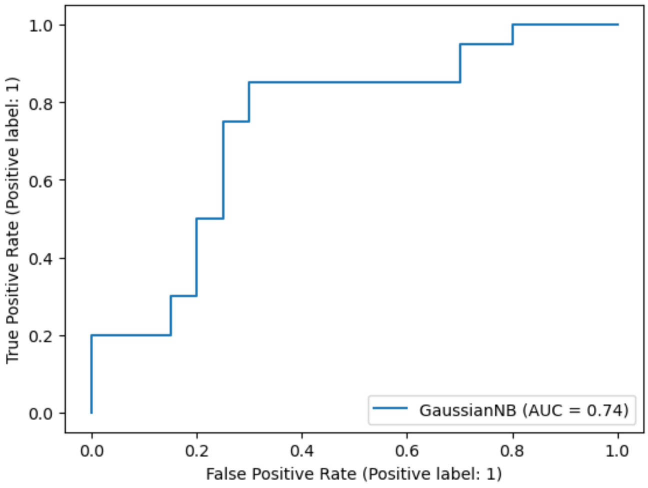

Figure 4. False Positive Rate of Naive Bayes Model

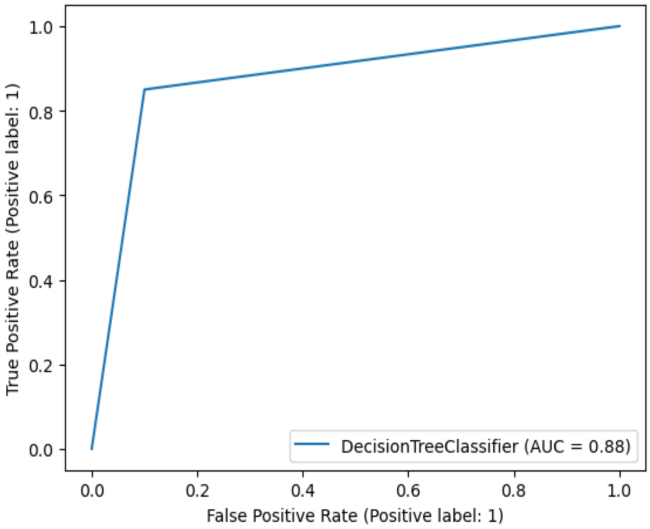

Figure 5. False Positive Rate of Decision Tree Model

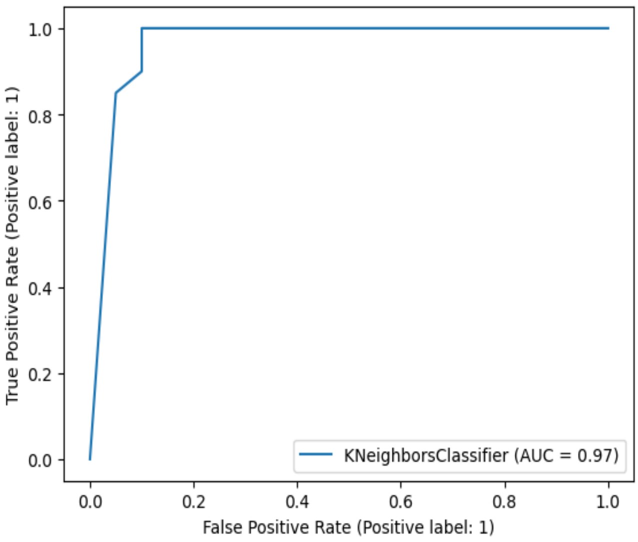

Figure 6. False Positive Rate of K-Nearest Neighbors Model

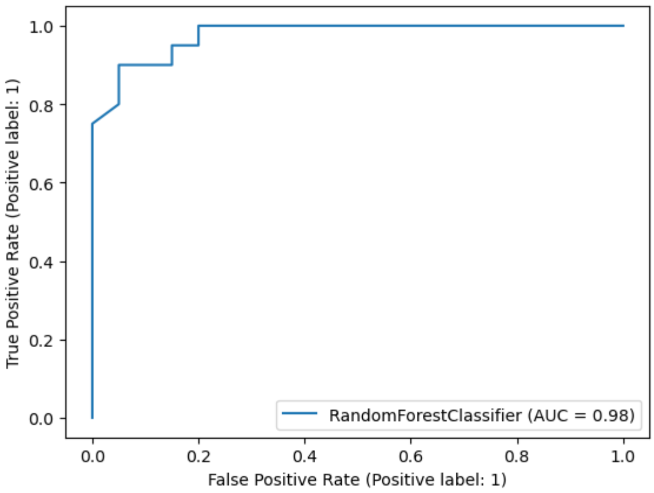

Figure 7. False Positive Rate of Random Forest Model

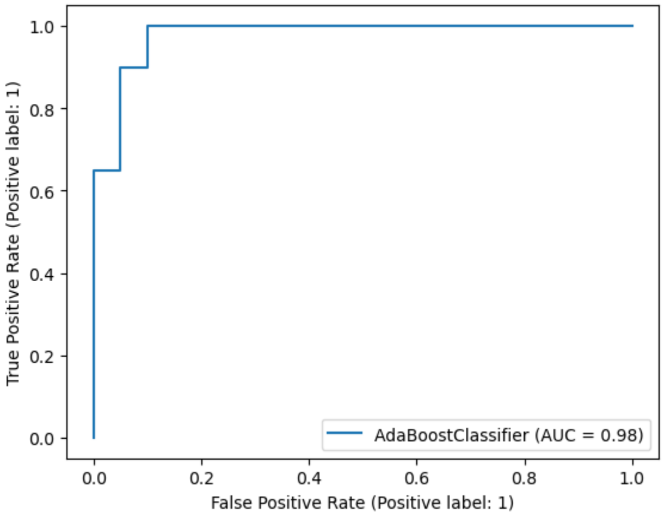

Figure 8. False Positive Rate of AdaBoost Model

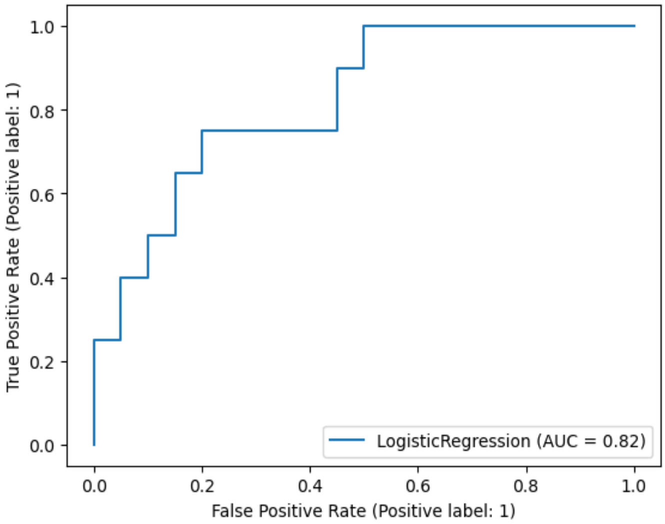

Figure 9. False Positive Rate of Logistic Regression Model

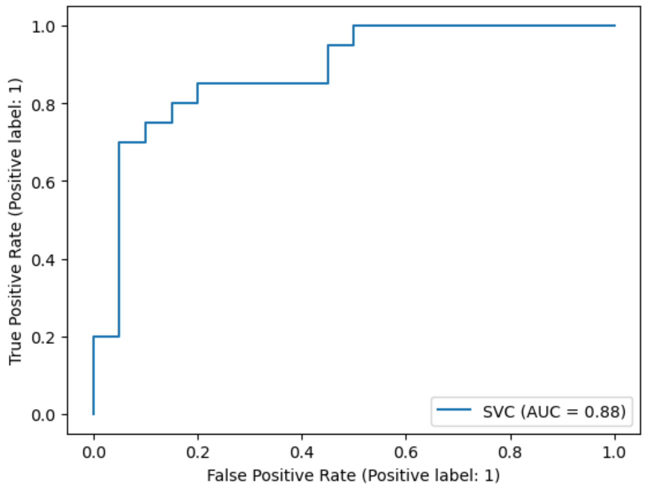

Figure 10. False Positive Rate of Support Vector Machine Model

Each one of Decision Tree, Random Forest, ADABoost, K-Nearest Neighbor classifiers has very good AUC score (90% or more, up to 98%), indicating, once more, very good performances of the algorithms on the dataset.

Table 1. Comparison between Accuracy Precision and Recall

Column1	Accuracy	Precision	Recall
Logistic Regression	0.775	0.789	0.750
Naïve Bayes	0.550	0.600	0.300
K-Nearest Neighbor	0.950	0.909	1.00
Decision Tree	0.875	0.895	0.850
Random Forest	0.875	0.875	0.900
AdaBoost	0.925	0.947	0.900
Support Vector Machine	0.800	0.773	0.850

In Table 1. A comparative study has been given between different Models like Naive Bayes, Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Decision Tree, Random Forest, and AdaBoost in terms of Accuracy, Precision and Recall.

Table 2. F1 Score comparison between different Models

Column1	F1-Score
Logistic Regression	0.769230769230769
Naïve Bayes	0.400000000000000
K-Nearest Neighbor	0.952380952380952
Decision Tree	0.871794871794872
Random Forest	0.878048780487805
AdaBoost	0.923076923076923
Support Vector Machine	0.80952380952381

In Table 2. A comparative study has been given between different Models like Naive Bayes, Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Decision Tree, Random Forest, and AdaBoost in terms of F1 score.

6. Conclusion and Future Scope

High level performances of the algorithms on a small dataset confirms the validity of our approach to data gathering, preprocessing and usage of statistical measurements as representative features for the class labels. It also indicates that color images of varying resolutions and sizes can be used to train Machine Learning algorithms and create models that will have the ability to distinguish and identify a crime from other, normal events that are happening around us.

Future work: The current study uses a small dataset. In the next phase of work, the dataset size will be increased. No optimization technique like n-fold cross validation, grid or random search to find out optimal values of hyper parameters were used in the current study. In the next phase of the study, our effort will be to select the best Machine Learning model from a set of optimized machine learning models trained on a larger dataset.

Data Availability

There is no publicly available according to the problem statement. Data has been gathered in such suitable way best fit with the problem statement which has been already discussed in introduction part. The data availability still remain closed because the data set and the work still in developing phase.

Conflict of Interest

It has been declared that there is no conflict of interest.

Funding Source

There is no funding source and supporting grant available for this research work.

Authors' Contributions

Author-1 researched literature and involved in Research methodology formulation and Data analysis. He also involved in writing the manuscript

Author-2 researched literature and conceived the study. He also involved in Research methodology formulation, edited the manuscript and approved the final version of the manuscript.

Author-3 perform the data collection and data cleaning part.

Acknowledgements

We want to extend our gratitude towards Mr. Sandeep Sengupta, Mr. Malay K Karanjai, Mr. Samyajit Mukherjee, and Ms. Amrita Acharya for their support in this endeavor.

References

[1] Eamthanakul, Busarin, Mahasak Ketcham, and Narumol Chumuang. "The traffic congestion investigating system by image processing from CCTV camera." In *2017 international conference on digital arts, media and technology (ICDAMT)*, pp. 240-245. IEEE, 2017.

[2] Manu, Y. M., G. K. Ravikumar, and S. V. Shashikala. "Anomaly Alert System using CCTV surveillance." In *2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon)*, pp. 1-5. IEEE, 2022.

[3] Intani, Pattana, and Teerapong Orachon. "Crime warning system using image and sound processing." In *2013 13th International Conference on Control, Automation and Systems (ICCAS 2013)*, pp. 1751-1753. IEEE, 2013.

[4] Gupta, Pratishtha, Manisha Rathore, and G. N. Purohit. "Detection of direction deviation of vehicle using CCTV cameras." In *International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)*, pp. 1-10. IEEE, 2014.

[5] Conaire, Ciarán O., Noel E. O'Connor, Eddie Cooke, and Alan F. Smeaton. "Multispectral object segmentation and retrieval in surveillance video." In *2006 International Conference on Image Processing*, pp. 2381-2384. IEEE, 2006.
[6] Chmielewska, Agata, Parzych Marianna, Tomasz Marciniak, Adam Dabrowski, and Przemyslaw Walkowiak. "Application of the projective geometry in the density mapping based on CCTV monitoring." In *2015 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, pp. 179-184. IEEE, 2015.
[7] Yang, Yan, Brian C. Lovell, and Farhad Dadgostar. "Content-based video retrieval (cbvr) system for cctv surveillance videos." In *2009 Digital Image Computing: Techniques and Applications*, pp. 183-187. IEEE, 2009.
[8] Sodanil, Maleerat, and Chalermpong Intarat. "A development of image enhancement for CCTV images." In *2015 5th International Conference on IT Convergence and Security (ICITCS)*, pp. 1-4. IEEE, 2015.
[9] Harte, Naomi, Andrew Rankin, Gary Baugh, and Anil Kokaram. "Detection of illegal dumping from CCTV at recycling centres." In *2007 International Machine Vision and Image Processing Conference*, pp. 204-204. IEEE Computer Society, 2007.
[10] Saad, Mohamad Hanif Md, and Aini Hussain. "Design and development of intelligent anomalous behaviour and event detection system." In *2010 6th International Colloquium on Signal Processing & its Applications*, pp. 1-5. IEEE, 2010.
[11] Rizki, Rai Purnama, Eki Ahmad Zaki Hamidi, Lia Kamelia, and Ramdani Wahyu Sururie. "Image processing technique for smart home security based on the principal component analysis (PCA) Methods." In *2020 6th International conference on wireless and telematics (ICWT)*, pp. 1-4. IEEE, 2020.
[12] Noguchi, Hirofumi, Takuma Isoda, and Seisuke Arai. "End-to-End Person Re-identification including Camera Zooming based on Meta-Analysis for Images." In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 3285-3290. IEEE, 2021.
[13] Alpatov, Boris A., Pavel V. Babayan, and Maksim D. Ershov. "Embedded image processing and video analysis in intelligent camera-based vision system." In *2020 9th Mediterranean Conference on Embedded Computing (MECO)*, pp. 1-4. IEEE, 2020.
[14] Khasanova, Aliia, Alisa Makhmutova, and Igor Anikin. "Image denoising for video surveillance cameras based on deep learning techniques." In *2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, pp. 713-718. IEEE, 2021.
[15] Casanova, Guillermo, Daniel Yandún, and Graciela Guerrero. "Analysis of video surveillance images using computer vision in a controlled security environment." In *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*, pp. 1-6. IEEE, 2020.
16] Khan, Muhammad Aseer. "Image Processing Based Real Time Motion Detection System Using GSM and Embedded System." In *2020 IEEE 23rd International Multitopic Conference (INMIC)*, pp. 1-6. IEEE, 2020.

AUTHORS PROFILE

Mr. Saumitra Biswas earned his M.Tech in computer science from West Bengal University of Technology, MSC in statistics from Kalyani University. He is working as Assistant professor at ISOAH Data Securities Pvt Ltd. with 22 years of teaching experience. His research interests include Machine Learning, Deep Learning, and Image Processing.

Mr. Sanchayan Bhaumik earned his ME degree from Jadavpur University in 2013. Mr. Bhaumik is currently working as Assistant Professor in Department Computer Application from Maulana Abul Kalam Azad University of Technology (MAKAUT). He is also Certified Ethical Hacker (CEH) and Certified Computer Hacking Forensic Investigator (CHFI) from EC-Council, USA. Previously Mr. Bhaumik worked as Sr. Information Security Analyst at ISOAH Data Securities. He has more than 7 years of industry experience as an Ethical Hacker and Digital Forensic Investigator. He has successfully conducted various information security audit and Forensic Investigation for Government and leading corporate houses. His research domain is Information Security, Image Processing, Machine Learning, Signal Processing, and Target Tracking.

Mr. Tanay Bag earned his MSC from Kalyani University. He is working as Assistant professor at ISOAH Data Securities Pvt Ltd. His research domain is information Security, Image Processing.

