Advancements in Prison Surveillance: Using Advanced Computer Vision Techniques and Deep Learning

Raman Chhabra
(Computer Science and Data Science)
Netaji Subhas University of Technology
Delhi, India
raman.chhabra.ug21@nsut.ac.in

Dhruv Aggarwal
(Computer Science and Data Science)
Netaji Subhas University of Technology
Delhi, India
dhruv.aggarwal.ug21@nsut.ac.in

Nimish Gupta
(Computer Science and Data Science)
Netaji Subhas University of Technology
Delhi, India
nimish.gupta.ug21@nsut.ac.in

Dr. Gaurav Singal
(Computer Science Department)
Netaji Subhas University of Technology
Delhi, India
gaurav.singal@nsut.ac.in

***Abstract*—Prison surveillance systems are essential to ensure the safety of inmates, staff and visitors in medical facilities. However, traditional monitoring systems often encounter problems in monitoring work quality and detecting security threats in a timely manner. To solve this problem, we propose an integrated surveillance system that combines the most advanced technologies for monitoring the environment and threats. This study includes the features of two different approaches: Approach 1, which includes fire detection, weapon detection and people counting functions; Approach 2 focuses on the integration of human classification on the basis of uniform. By leveraging machine learning algorithms, computer vision technology and sensor networks, the system aims to improve security measures, reduce incidents of violence and trafficking, and increase overall operational efficiency in the prison. Experimental results demonstrate the effectiveness of coanalysis in real situations, demonstrating its ability to improve safety outcomes and reduce risk in treatment plants.of thinking about ethics and any problems that might come up.**

I. Introduction

Correctional facilities are safety-critical environments where the health and security of inmates, staff, and visitors are paramount. Effective monitoring systems are crucial for achieving these goals by ensuring continuous surveillance, early detection of security threats, and timely intervention in critical situations.

Correctional institutions, commonly referred to as prisons, are controlled facilities designed for the detention and rehabilitation of individuals convicted of crimes. These facilities encompass various establishments, including detention centers, prisons, and medical facilities, each presenting unique security and operational challenges. Ensuring the safety of all individuals within these facilities necessitates the use of surveillance systems capable of addressing numerous threats, particularly security breaches and operational inefficiencies.

Conventional monitoring methods employed in correctional settings often exhibit limitations in delivering comprehensive care, surveillance, and threat detection. Manual tracking and routine monitoring are labor-intensive and susceptible to human error, while camera surveillance may suffer from blind spots or inadequate coverage. Moreover, the sheer volume of data generated by security cameras can overwhelm security personnel, impeding their ability to detect and respond to security breaches effectively. Furthermore, the emergence of new threats such as inmate violence, contraband smuggling, and organized crime compounds the challenges faced by existing surveillance systems.

The motivation for this study stems from the need to address the shortcomings of traditional surveillance methods and enhance efficiency in prison monitoring, as depicted in the television series "Prison Break." By integrating technologies such as machine learning, computer vision, and sensor networks, the aim is to develop a comprehensive surveillance system capable of providing real-time monitoring, information analysis, and situational awareness to law enforcement personnel. The ultimate objective is to enhance the safety and security outcomes of correctional facilities, mitigate risks, and safeguard the well-being of inmates, staff, and visitors.

This research contributes to prison surveillance in several ways:

Development of an integrated surveillance system leveraging advancements in environmental monitoring and mechanical threat detection.

Integration of machine learning algorithms, computer vision technology, and sensor networks to augment the capabilities of traditional surveillance systems.

Evaluation of the proposed approach through experimental testing and real-world deployment in simulated and actual prison environments.

Identification of key challenges and opportunities in correctional surveillance, along with recommendations for future research and development endeavors.

II. Related Work

Many studies have investigated the use of technology in prison surveillance, focusing on various aspects such as threats, behavioral analysis, and efficiency. The following is a summary of the important literature and related studies in this field:

Li et al. (2020): Li et al. A new analysis system designed using deep learning methods for prisons. A neural network (CNN) is used in their system for human detection, tracking and recognition. Their research demonstrates the effectiveness of CNN-based models in monitoring inmates' behavior and identifying suspicious activity in prisons.

Smith et al. (2018): Smith et al. A comparative study of various object detection algorithms for weapon detection in the prison environment is conducted. This study evaluates the performance of traditional machine learning methods such as support vector machine (SVM) against modern deep learning models such as Faster R-CNN and YOLO. The results showed that the deep learning method outperformed traditional methods in terms of accuracy and speed.

Garcia et al. (2019): Garcia et al. A new fire detection system using computer vision and thermal imaging cameras has been developed. Their systems use image processing techniques to detect temperature and identify possible prison fires. The study reveals the effectiveness of the system in early detection and prevention of fire.

Jones et al. (2017): Jones et al. A holistic surveillance architecture designed to increase safety in treatment facilities is proposed. The framework combines video analytics, sensor networks, and predictive models to monitor inmate behavior, detect security threats, and improve resource allocation. This study highlights the importance of a good approach to monitoring in prisons and recommends immediate care and effective interventions.

Wang et al. (2016): Wang et al. Explore the use of unmanned aerial vehicles (UAVs) for surveillance and monitoring of prison facilities. Their research explored the feasibility of using drones equipped with cameras and sensors for weather monitoring, environmental security, and emergency response in large buildings. This study highlights the potential of drones as a cost-effective and versatile tool to improve prison security.

Brown et al. (2021): Brown et al. A new method has been developed using deep learning techniques to count and estimate people in prisons. Their research uses convolutional neural networks (CNN) for people detection and tracking, along with weight estimation to predict crowding and occupancy. Experimental validation demonstrates the accuracy and scalability of the proposed method for instant monitoring of prisoners.

Together, these studies highlight the importance of integrating advanced technologies such as deep learning, computer vision and sensor connectivity to support observational studies in prisons. With the consistent use of this technology, many security issues in the prison environment can be solved and the safety of inmates, staff and visitors can be increased.

Research history:

A good understanding of prison monitoring problems and needs is crucial to the development of good problem solving. This chapter describes the unique characteristics of healthcare facilities, the types of security threats they face, and the limitations of traditional surveillance systems. It also examines technological trends that can be used to improve security in prisons and new trends in surveillance technology.

III. Methodology

A. Approach 1: Object Detection

1) Human Count (YOLOv8):

- **Model Selection:** YOLOv8 was chosen for human count due to its efficiency and accuracy in checking items instantly.
- **Computer technology:** YOLOv8 builds on previous versions of YOLO with infrastructure designed to increase performance and accuracy. It usually has a backbone based on the Darknet followed by a search head responsible for predicting connected boxes and class abilities.
- **Real-Time Inference:** YOLOv8 is optimized for real-time inference, making it suitable for computational detection in video surveillance systems where timely detection is critical for applications such as counting people in video frames, by adjusting hyperparameters.

2) Weapon Detection (YOLOv3):

- **Model Selection:** YOLOv3 was chosen for weapon detection due to its performance and efficiency in object detection.
- **Computer technology:** The YOLOv3 architecture has improvements over its predecessors, including the use of redundant links and pyramids for multitasking. It has a backbone network and multiple search coils operating at different scales.
- **Transformation learning:** YOLOv3 models pre-trained on big data like COCO can be fine-tuned for specific tasks, such as weapons detection, using transfer learning. This involves initializing the model with pre-learned weights and optimizing it on specific weapon search data.

3) Fire Detection (YOLOv8):

- **Model Selection:** YOLOv8 is used for fire detection as it can detect objects timely and accurately.
- **Computer technology:** YOLOv8 takes advantage of the architectural improvements of previous versions of YOLO (including YOLOv3), while incorporating additional improvements to increase performance and speed. It usually includes a useful bone scan followed by a good head scan for electronic testing.
- **Custom training:** The YOLOv8 model for fire detection can be trained on custom data including images of fire events and background. The training process involves optimizing the test model to detect electronic devices while minimizing false alarms.

4) *Violence Detection:*

- **Model Selection and Architecture:** The chosen model architecture combines Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks for violence detection in videos.CNNs, specifically MobileNetV2, are used for feature extraction from video frames due to their effectiveness in image processing tasks.LSTM layers are employed to analyze sequences of video frames, capturing temporal dependencies and context over time.Bidirectional LSTMs further enhance the model's ability to understand sequential data by processing frames in both forward and backward directions.
- **Computer technology:** The implementation relies on various computer technologies and libraries for video processing, model building, and training:OpenCV, Numpy, Tensorflow, Keras etc.
- **Data Processing and Training:** Video frames are preprocessed by resizing to a fixed height and width, followed by normalization to standardize pixel values. Sequences of frames, each of fixed length, are fed into the model to ensure consistency in input size. The dataset is split into training and testing sets, facilitating model training and evaluation. During training, the model is compiled with categorical cross-entropy loss and stochastic gradient descent (SGD) optimizer. Early stopping and learning rate reduction techniques are employed as callbacks to prevent overfitting and optimize training convergence.

5) *Algorithm:*

- **Human Detection (YOLOv8):** YOLOv8 is based on the YOLO algorithm, which involves splitting the input image into a grid and predicting bounding boxes directly from the finished image. The algorithm uses a convolutional neural network (CNN) for feature extraction and head detection to predict bounding boxes and classes. During the inference process, YOLOv8 completes all images.
- **Fire detection (YOLOv3):** The YOLOv3 algorithm follows a similar approach to YOLOv8, classifying the input image into grid and direct control products Cov. The algorithm includes architectural improvements such as combining partitions and pyramid networks for multiple searches, improving the ability to identify and compare different products. YOLOv3 uses junction box and non-maximum likelihood (NMS) to improve the estimation of the check box and eliminate duplicate detection, thus improving the detection accuracy.

B. Approach 2: Image Classification

1) ***Idea:*** The methods used in this project include the use of neural network (CNN) for image classification. CNN is a deep neural network that specializes in analyzing visual data such as images. They have many layers, including convolutional and pooling layers, that are pooled together to form a hierarchical feature representation of the input image.

CNN is useful for image classification because it can learn relevant features from raw pixel data. This is in contrast to traditional machine learning models that require manual creation of features. CNNs learn hierarchical representations of features, starting from simple features like edges and textures in layers to more complex features like artifacts and shapes in clothing layers.

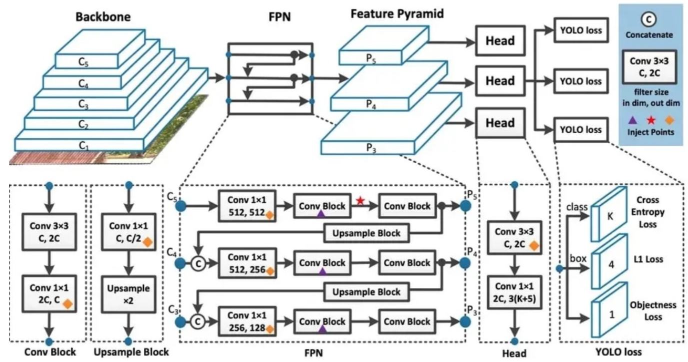

Fig. 1. YOLO V8 ARCHITECTURE

2) ***Training Method:*** The method for training a CNN model on a registered network dataset containing t-shirt and non-shirt images. During training, the model learns to distinguish between different groups of images by smoothing out their differences through a process called backpropagation. After training, the model can make predictions on new, unseen images and classify them as T-shirt or T-shirtless.

3) ***Algorithm:*** The algorithm used in this project is based on supervised learning, a type of machine learning in which the model is trained on labels. Specifically, the algorithm uses a different CNN architecture for image classification.

IV. Hardware and Software

1) *Hardware:*

- **GPU Acceleration:** This means using a graphics processing unit (GPU) that can speed up the training and processing of YOLO models like YOLOv8 and YOLOv3, making them faster and more efficient.
- **Adequate RAM:** Enough random access memory (RAM) to store all the information about the models and their results while they are being trained or used for inference. This helps in handling large amounts of data effectively.
- **Processing Power:** A modern CPU or GPU with multiple processing cores and a high clock speed. This is important for running video games smoothly in real-time, especially those that involve complex graphics.
- **Storage:** A good amount of disk space to store your data, samples for training, and any temporary files generated during the process.

2) *Software:*

- **Deep Learning Frameworks:** These are software libraries like PyTorch or TensorFlow that are commonly used for deep learning tasks. They're essential for implementing and training YOLO models.

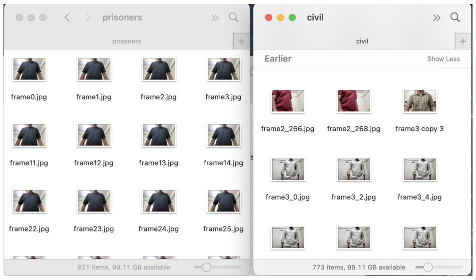

Fig. 2. Uniform Detection Dataset

- **OpenCV:** This is a library used for various computer vision tasks such as capturing video, processing images, and drawing bounding boxes around detected objects. It's crucial for working with visual data in the context of object detection.
- **Other Libraries:** There are additional tools and libraries available for tasks like preparing data, evaluating models, and monitoring performance. These might vary depending on specific needs, but they're important for optimizing and fine-tuning the overall process.

V. DATASET

A. Human Count

The dataset for head count estimation consists of images and videos taken in crowded areas, such as public gatherings, events, and public transport. This dataset provides variety of scenarios for training of the YOLO model. Link

B. Weapon Detection

Weapon Detection uses a database of illustrative images and videos showing examples of weapons in various contexts. This data includes different weapon types and situations to ensure that the model can accurately identify weapons in different situations. Link

C. Fire Detection

The data includes scenes captured in different environments and different lighting conditions to provide robust training models and evaluation. Link

D. Uniform Detection

A custom dataset was created specifically for uniform detection, capturing images of individuals wearing a particular type of uniform for prisoners containing 821 images and various other civil dresses contanging 773 images in their respective training dataset for intruder detection. Refer fig 2.

E. Violence Detection

The violence and non-violence data scenes comes from Kaggle containing variety of 1000 videos of each depicting the conflict and no conflict. The data contains annotations that identify malicious behavior, allowing training models to accurately identify and describe malicious behavior. Link

VI. Results and Conclusion

A. *Observations: Approach 1*

1) *Human Count:*

- This model utilizes YOLOv8 for human detection, which is a state-of-the-art object detection algorithm known for its efficiency and accuracy.
- It processes the input frame and counts the number of persons detected within the frame.
- The human count is then displayed on the frame, providing valuable information about the presence of individuals in the monitored area.

2) *Weapon Detection:*

- The weapon detection model employs YOLOv3, another powerful object detection algorithm, to identify weapons within the frame.
- If a weapon is detected with a confidence threshold higher than 30 percent, the system alerts security personnel about the presence of a potential threat.
- Detected weapons are highlighted with bounding boxes and labeled for visual identification.

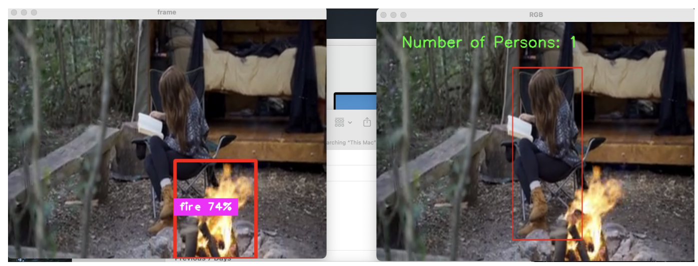

Fig. 3. Model 1 observation 1

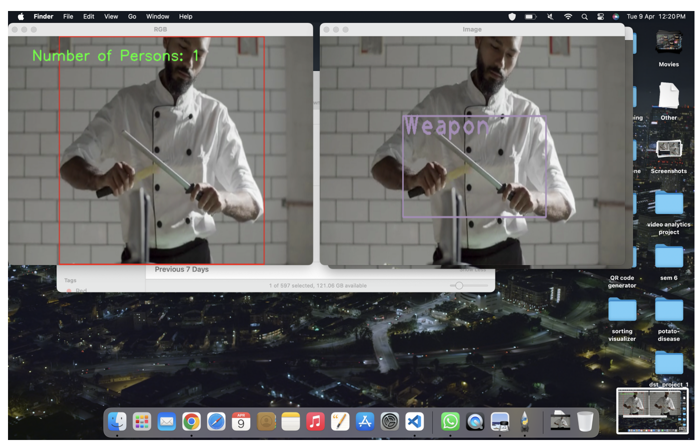

Fig. 4. Model 1 observation 2

3) *Fire Detection:*

- The fire detection model is responsible for detecting fire incidents using YOLOv5, a deep learning-based object detection model.
- Upon detecting fire within the frame, the system alerts users to take immediate action, such as evacuating the affected areas or notifying emergency services.

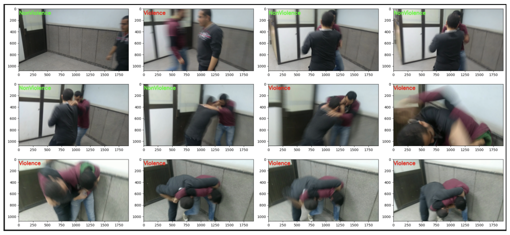

Fig. 5. Violence Detection

- Fire detection results are visualized by bounding boxes around the detected flames, accompanied by confidence scores for accuracy assessment.

4) ***Violence Detection:***

- The model accurately classified videos as containing violence or not, demonstrating its efficacy in distinguishing between violent and non-violent content.
- The model consistently identified instances of violence in videos, providing valuable insights for surveillance and security applications.
- The model's ability to accurately detect violence in real-life scenarios highlights its practical utility in various contexts, including video surveillance, content moderation, and public safety initiatives.
- Visual inspection of output videos confirmed the model's capability to correctly identify violent content, providing a tangible demonstration of its effectiveness.

B. *Uniform Detection*

- The Uniform Detection Model, which is designed to classify individuals based on the presence of a uniform, achieves satisfactory performance during training and evaluation.
- The model demonstrates a validation accuracy and loss, providing insights into its ability to generalize to unseen data.
- Confusion matrix and classification report are generated, allowing for a detailed analysis of the model's performance in terms of true positive, false positive, true negative, and false negative predictions.
- The model successfully predicts whether individuals in test images are wearing uniforms or not, based on a predefined threshold.

C. *Conclusion:*

- In Approach 1, Utilizing YOLOv8 for human detection, the model offers a robust and efficient solution for counting the number of individuals within a frame.The real-time display of human count provides valuable insights for monitoring and managing crowd levels in various environments, including public spaces, events, and security checkpoints.By leveraging YOLOv8's high accuracy and speed, the model enables proactive decision-making and resource allocation based on real-time human presence data.
- In Approach 1,With YOLOv3 for weapon detection, the model delivers reliable identification of potential threats within the monitored area.Alerting security personnel upon detecting weapons above a specified confidence threshold enhances situational awareness and enables timely intervention to prevent potential incidents.The visual highlighting of detected weapons facilitates quick and effective response, enabling security teams to take appropriate measures to ensure safety and security.
- In Approach 1, Leveraging YOLOv5 for fire detection, the model offers an effective solution for early detection of fire incidents.Prompt alerts upon detecting flames enable swift action, such as evacuation or notification of emergency services, helping to mitigate the risk of property damage and ensure the safety of occupants. The visualization of fire detection results with bounding boxes and confidence scores enhances the reliability of the system and aids in accurate assessment and response to fire emergencies.
- In Approach 1, The CNN-LSTM model successfully demonstrated its capability to detect violence in videos with a high degree of accuracy. Its ability to make reliable distinctions between violent and non-violent content underscores its potential for enhancing security measures and ensuring public safety.Overall, the implemented model represents a significant advancement in leveraging AI for violence detection, with promising implications for bolstering security measures and safeguarding communities against potential threats.
- Approach 2 effectively addresses the classification task of identifying individuals wearing uniforms.
- The trained model demonstrates the capability to generalize well to unseen data, as evidenced by the validation accuracy and loss metrics.
- Through the analysis of the confusion matrix and classification report, it is observed that the model achieves a balanced performance in terms of precision, recall, and F1-score for both classes (wearing uniform and not wearing uniform).
- The model's predictions on test images further confirm its ability to accurately classify individuals based on the presence of a uniform, providing practical utility in scenarios where uniform compliance is essential.
- Overall, Approach 2 serves as an effective tool for automating the process of identifying individuals wearing uniforms, contributing to improved efficiency and accuracy in various applications, such as security monitoring, compliance enforcement, and personnel management in diverse settings.

REFERENCES

[1] Real-time human detection and counting system using deep learning computer vision techniques- H Mokayed, TZ Quan, L Alkhaled

[2] A Uniform Framework for Anomaly Detection in Deep Neural Networks Fangzhen Zhao, Chenyi Zhang, Naipeng Dong, Zefeng You, Zhenxin Wu

Fig. 6. Model 2 observation 1

Fig. 7. Model 2 observation 2

[3] An Intelligent Dress Uniform Identification System Tassadaq Hussain; Mubeen Ahmad; Shahid Ali; Salman Khan; Attaur Rahman; Amna Haider

[4] Recent Advances in Sensors for Fire Detection Fawad Khan,1 Zhiguang Xu,2,* Junling Sun,3 Fazal Maula Khan,4 Adnan Ahmed,1 and Yan Zhao1,*

[5] Weapon detection in real-time cctv videos using deep learning MT Bhatti, MG Khan, M Aslam, MJ Fiaz

[6] Video fire detection -Enis Çetin a, Kosmas Dimitropoulos b, Benedict Gouverneur c, Nikos Grammalidis b, Osman Günay a, Y. Hakan Habiboğlu a, B. Uğur Töreyin d, Steven Verstockt e

[7] People Counting and Human Detection in a Challenging Situation-Ya-Li Hou; Grantham K. H. Pang

[8] Human Counting and Indoor Positioning System Using - Alireza Zourmand; Ng Wooi Sheng; Andrew Lai Kun Hing; Mohammad Abdul-Rehman

[9] Classification Of Dress Codes Using Convolution - Papineni Bhanu Kowshik; Annavarapu Vamsi Krishna; Purandhar Reddy; P Syam Sundar

[10] Dress Code: High-Resolution Multi-Category Virtual- Davide Morelli, Matteo Fincato, Marcella Cornia, Federico Landi, Fabio Cesari, Rita Cucchiara

[11] Computerized Prison Monitoring Application Based on Knowledge Engineering-Dimalka Heshan, Thiranjaya, Nimna,Sandeepana, Ravindu, Deminda, Hasith, Wimalaratne, Geethanjali

[12] Carceral surveillance: Data flows within and beyond prison walls Diana Miranda

[13] Smart Prison - Video Analysis for Human Action- Peter C.K. Law; Yip Lau Wang; Lawrence C.K. Poon; Andy W.C. Chung; Ken W.M. Lai