Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 10s (2025)

Optimized Deep Learning System for Criminal Activity Detection using Facial Recognition, Video Surveillance, and Weapon Detection

[1]Pranita Sanjay Waikar, [2]Waseem Ahmad Mir

[1]PG Student, Department of Computer Engineering, Department of Computer Engineering, JSPM Wagholi, Pune

pranitawaikar9595@gmail.com

[2]Professor, Department of Computer Engineering, Department of Computer Engineering, JSPM Wagholi, Pune

waseemmir78177@gmail.com

Article History:

Received: 12-01-2025

Revised: 15-02-2025

Accepted: 01-03-2025

Abstract: The evolving field of deep learning has automated intelligent surveillance and criminal activity detection systems. This paper provides refined optimized deep learning architecture systems aimed at improving crime detection using facial recognition, real-time video surveillance systems, and weapon detection. These systems utilize the latest neural network structures, chiefly CNNs, RNNs, and hybrid deep learning models, to derive accuracy, real-time analysis, and scalability. The automated facial recognition systems track users' behaviors, registers recognized users, recognizes suspicious behaviors, and detects weapons at video feed frames. Using spatial-temporal analysis helps detect and reduce response delays in public areas, transport systems, and other sensitive places. Important issues such as dataset deficiencies, real-time computing limits, aggressive attacks, and ethical concerns on surveillance AI are given attention. Transfer learning, model compression, edge computing, and AI transparency are some solutions put forward. The paper addresses these challenges and emphasizes the need for diverse, representative training datasets to strengthen system dependability and equity. Future research and development may benefit from the discussion highlighting transformer-based architectures, multimodal fusion, and other emerging trends which incorporate reinforcement learning. This review furthers the state of the art by integrating advanced technology development and documenting noted advancements, turning them into actionable frameworks by formulating real-world deployment strategies. The primary contribution of this paper is to assist researchers, developers, and even policymakers comprehend in-depth, the potential deep learning methods offer in the design of surveillance systems for criminal activity detection that are safe, smart, and ethically responsible.

Keywords: Deep Learning, Criminal Activity Detection, Facial Recognition, Video Surveillance, Weapon Detection, CNN, RNN.

Introduction

The application of artificial intelligence (AI) and deep learning (DL) technologies integrates seamless monitoring of public safety and crime detection. Law enforcement and security organizations are now able to recognize illegal activities with advanced precision and swiftness due to the implementation of AI technologies such as facial recognition, intelligent video monitoring, and automatic weapon detection systems. These systems employ algorithms to aid in the protection of metropolitan and

sensitive zones by instant threat analysis, thus increasing response agility and reducing man-made blunders. AI automation of surveillance systems and integration of control functions into the observation devices enhances the effectiveness and efficiency in the entire security sector. Innovations such as facial recognition technologies have proven to be useful in identifying individuals in sensitive areas using biometric characteristics, and it requires deep learning techniques such as CNNs on convolutional neural networks to recognize and compare faces using databases [1]. With the development of facial embedding models, systems can accurately carry out identity verification and enforce the law in airports, military-grade stadiums, and even consumer devices, presenting unmatched precision unnoticed before. These sophisticated systems identify and match suspects with criminal databases and even detect previously unrecognized individuals who potentially pose a threat. Widely publicized infrastructure systems help in forensic investigations while also preventing unlawful activities, integrating them serves to deter criminal activities. Regardless of the advancements made, the systems still require ongoing refinement due to challenges such as lighting variations, occlusion, and pose and dataset bias.

Just as Deep Learning (DL) techniques have advanced facial recognition technology, they have also enhanced video surveillance systems. Older surveillance technologies suffered from many drawbacks, including the need for intensive manual oversight, which made them prone to missing threats due to tiredness or simple human overlooking. Today, DL-enabled surveillance technologies can monitor video streams for behavioral anomalies and even suspicious activities in real-time. Visual and temporal patterns of video streams are analyzed using CNNs and RNNs (recurrent neural networks), making it possible to detect threats such as loitering (i.e., remaining in a particular space for longer than normal), unauthorized access, and aggressive behavior [2]. These systems allow instant notification to the security personnel, which in turn permits instant response and averts preventive action against criminal activities. Smart cities are employing behavioral analysis algorithms to monitor and analyze the movement of pedestrians and vehicles, identify unusual behaviors, and integrate multiple data sets to recognize advanced threats. In crowded places such as large public events or train stations, these systems provide extraordinary situational awareness that would be impossible to achieve manually. The public safety aspect of deep learning technology is the application of deep learning for automatic weapon detection. There is a growing need for concealed, and visible automated weapon detecting systems especially with the proliferation of mass violence and terrorist activities. Sophisticated object detection tools, especially CNN-based ones, are able to identify firearms, knives, and other dangerous materials in surveillance videos. These models are prepared with extensive weapon image datasets captured under varying lighting, occlusion, and resolution conditions. To enhance situational awareness, weapon detection systems can be used in schools, airports, and shopping malls enabling authorities to prevent violent actions by mitigating threats before they escalate. Automated weapon detection using deep learning is often executed with edge computing systems which require less cloud dependency and result in faster, real-time, low latency processing.

Despite the many advantages that come with implementing AI in the identification of criminal activities, these advancements are accompanied with limitations and ethical concerns. A critical technological concern is the bias within datasets, which may result in unjust forecasts and biased results. Models, for instance, may not perform accurately in real-life situations if specific demographic

groups or certain environmental factors are ignored or underrepresented in the training data. These kinds of biases can erode societal trust and have negative impacts, especially in the context of policing [4]. Fostering such biases entails creating more diverse and representative datasets alongside adopting fairness-aware training approaches, applying AI decision frameworks, and audit accountability systems that oversee and evaluate AI resolutions. Simultaneously, privacy issues have arisen as a significant barrier to the unrestricted use of AI surveillance technologies. The capturing of people's images and videos in public spaces and even in private places raises important dilemmas of consent, ownership of data, and individual freedoms. The face recognition technology has drawn much criticism for the likelihood of it being abused by the state or private organizations to spy on people without their knowledge or approval. The advantages of such systems of surveillance must also be balanced with their ethical considerations. Appropriate policies such as laws on data protection, guidelines on the ethics of AI, and others are critical in safeguarding the technologies from violations to the rights and the laws of people [4].

Addressing AI-powered surveillance systems, their implementation incurs considerable costs which stem from high-performance data servers capable of real-time video stream analytics and intricate neural network processing. Chronic underfunding from small agencies in developing regions renders high-performance GPUs economically unfeasible. Coupled with the need for multi-camera scalable systems with distributed architectural intelligence capable of functioning autonomously, the surveillance systems become much more advanced to build. Fulfilling these requirements without drastically increasing the system cost while retaining high performance marks poses additional obstacles to refine algorithms using techniques such as pruning or quantization alongside efficient deep learning models or lightweight ones. There is also a growing stress on utilizing edge solutions which tackle distant computing cores as close as possible to surveillance system sensors to increase bandwidth thresholds. The goal of this review is to provide an overview of modern technologies used to detect criminal activities, emphasizing the integration of facial recognition, video surveillance, and weapons identification systems.

The objectives of this research include analysing existing approaches, outlining unresolved problems, presenting actionable insights and recommendations, and advancing the discourse of the application of AI and deep learning for public safety. Regardless of the technological promise offered by such systems, the successful and ethical adoption of these systems will require concerted efforts of interdisciplinary research, engineering, policy and legal frameworks, and civil advocacy. As we look ahead, the area of AI-based criminal detection technology is expected to grow further. New considerations include the application of multimodal inputs like social media metadata, audio streams, and environmental sensors within the context of surveillance systems to make situationally aware systems that adapt to their surroundings. Performance enhancement of video analysis models using Transformer-based architectures and self-attention mechanisms is also under consideration. Reinforcement learning is being examined for its potential to enable systems that can learn optimal surveillance techniques in ever-changing environments. These developments will surpass current intelligent threat detection capabilities. Still, achieving the full promise of these technological advances will demand rigor in innovation balanced with ethical boundaries around governance, and transparency, and a balance of responsibility. Properly designed and responsibly deployed systems can

harness AI to enhance surveillance and empower targeted crime prevention while enabling community protection and the creation of safer shared public environments.

Literature Review

Y. Sahukar et al. [1] describes a more sophisticated intelligent surveillance system employing deep learning methods to improve security. The system uses Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) Networks, and Autoencoders for real-time object detection, tracking, and classification. The system achieves accurate object identification and tracking within a surveillance region through the integration of spatial and temporal data analysis. The authors underline the capability to efficiently process video feeds as a critical feature for prompt reaction to possible security issues. In essence, the implementation provides substantial enhancement in accuracy and responsiveness of surveillance, demonstrating the application of deep learning in contemporary security technology.

M. Qaraqe et al. [2] introduce PublicVision, an intelligent and secure surveillance system aimed at recognizing crowd behaviors. Powerful machine learning tools are embedded into the system to analyze crowd behaviors, detect and classify anomalies, and raise red flags in real time. Transmitting CCTV data to a centralized hub ensures PublicVision's integrity and privacy, safeguarding data. The system can be deployed in almost any public venue and exponentially increases situational awareness, allowing proactive security interventions or measures. The research demonstrates the impact of combining secure data transmission with intelligent behavior analysis for enhanced public safety. Yehia Ibrahim Alzoubi et al. [3] Any current open literature remains lacking around the subject of deep learning, machine learning, and cloud computing security, therefore Alzoubi has attributed a very comprehensive review to this intersection. Over four thousand publications were analyzed by the authors to identify particular trends, gaps, and other prominent issues in the fields of cloud infrastructure and security. The focus continues with learning systems that employ machine learning for strong and flexible threat intelligence automation, deep learning for complex situational anomaly detection, as well as cloud infrastructure security automation. It addresses the gap on adaptive and self-learning security mechanisms designed to operate in a multi-tiered, agile, responsive, distributed environment along with dynamically changing hostile frameworks. The review is extremely beneficial to practitioners and researchers looking for techniques to improve cloud security using advanced analytical computational strategies. The "Eye Spy" project, aimed at increasing security in sensitive areas through the enhancement of existing CCTV systems using ML and deep learning, was introduced by Sagar Rajebhosale et al. [4]. AI is used in the system to change passive surveillance systems into affirmative active intelligent networks that are capable of detecting and responding to real-time threats. The avoidance of large-scale hardware changes and overhauls allows the project to vastly improve situational awareness while remaining budget-friendly. Since CCTV systems are currently in widespread use, adapting them to incorporate AI enables real-time assessment of security risks, which enhances public safety and security measures.

As the number of IoT devices in operation continues to grow, their security has become a primary focus. This research analyzes some of the methods and techniques used for automating ML model training and evaluation in vulnerability detection, such as the IoT23 dataset. The authors outline the problems IoT environments face with security and present proactive ML-based solutions designed to

predict and prevent identified challenges, especially concerning security vulnerabilities focused on mitigable threat precursors. This study is useful to many researchers and practitioners who are looking to advance the precision of forensic analyses performed on IoT devices using sophisticated algorithms. In the work of Abdulraqeb Alhammadi et al.[6], I noted a thorough survey on the application of Artificial Intelligence (AI) technologies in crafting and managing 6G wireless networks. The authors looked at many AI applications that are targeted towards optimizing the performance, efficiency, and overall smartness of wireless networks. Some of the high-level items discussed include AI for network optimization and resource management as well as the provision of next-level services such as holographic communications and the tactile internet. The work also focuses on the relevant issues and gaps in this field explaining what they deem the future of this field should be. This document is useful in the development and analysis of 6G wireless networks driven by AI for researchers and practitioners.

V. Mandalapu et al., [7] reviewed more than 150 articles in relation to the application of machine learning (ML) and deep learning (DL) algorithms in the prediction of crimes. The authors show the application of decision trees, support vector machines, and neural networks parallel to crime prediction, their applications in pattern and trend detection, and also pay attention to their effectiveness. The study also reviewed the datasets utilized for predictive research and underscored the need for quality and diversity of the data used. Noting existing issues and recommending steps for future inquiries, such analytical studies become important not only to investigators, but also police departments aiming to refine predictive policing.

Yassine Himeur et al.[8] gives an account on the application of deep transfer learning (DTL) and deep domain adaptation (DDA) for video surveillance systems. They explain how DTL and DDA overcome such limitations as scarce labeled data pertaining to specific domains and shift in the domains themselves by improving the generalization capabilities of surveillance models. Through the review of some case studies and experiment analyses, the study shows the application of these methods for improved detection of important objects, activities, and even anomalies in their varied settings. This review also shows DTL and DDA as intelligent tools for video surveillance systems.

Olayemi Olaniyi et al. [9] examines the development of components of intelligent video surveillance systems (IVSS) which include the application of machine vision and video analytics. The authors analyze the development of IVSS and cover important areas like object detection, tracking, and behavior analysis. Within the scope of this study, the authors also describe the challenges associated with implementing IVSS, focusing on real-time processing and privacy issues. Through the review of different methodologies and technologies, this study provides an analysis of the current situation of IVSS and offers recommendations on how to improve system performance and operational efficiency through further research.

Rana Ayad et al. [10] proposes a CNN-based model developed specifically for mobile remote surveillance system as an addition to the existing home surveillance technologies to improve security. Using deep learning, the system accurately detects and classifies human motion, differentiating between members of the household and possible intruders. The model analyzes movement within a monitored area to determine when unauthorized person enters an area and has a very high accuracy of 99.8 percent. The active use of these intelligent surveillance systems provides a new dimension to

home security by minimizing instances of unnecessary alarm activations and enabling prompt alerts to users.

Athina Ilioudi et al [11] studies the application of deep learning techniques alongside domain knowledge for the enhancement of object detection and video data segmentation. For the review, the authors have analyzed existing methodologies alongside proposing a hybrid framework which includes a synergetic combination of domain-driven and data-driven techniques. Contextual information incorporation aids the system in improving interpretability as well as the accuracy of detection tasks. The study highlights the key focus of addressing the gap between the deep learning model and domain knowledge in order to achieve a more versatile and robust solution solutions in computer vision. Stephen Khor Wen Hwooi et al. [12] devises a system for continuous affect prediction using deep learning technique and facial expression images. Discrete emotion labels are mapped to a continuous valence-arousal space which allows for deepened comprehension of human emotions. Utilizing convolutional neural networks helps the system predict emotional states accurately and outperform the AffectNet dataset among others. This is particularly useful in Human-Computer Interaction, mental health evaluation, and affective computing, particularly during the assessment of subtle emotional cues.

Sathya G et al.[13] describes a face recognition system based on deep learning technologies aimed at detecting criminals. The system photographs offenders' faces and scans them against the repository of already captured faces for potential matches. The model uses deep learning through convolutional neural networks which improves face recognition accuracy. The application of this technology helps law enforcement agencies identify suspect criminals rapidly which, in turn, improves the safety and crime prevention efforts of the society.

Hyun-Bin Kim et al.[14] develops a system for monitoring video streams that provides real-time and high-accuracy recognition of criminal faces within uncontrolled scenes. The system uses sophisticated deep learning techniques for face detection and recognition as well as monitoring from videos captured in different environments. The model deals with features like occlusions as well as poor lighting thereby maintaining accuracy with specified thresholds over the identification of pertinent individuals. The system expands the boundary of surveillance cameras located in public areas and therefore enhances the effectiveness of crime surveillance and prevention systems.

Adikonda Mounika et al. [16] describe a system that automatically supervises video surveillance data using advanced deep learning algorithms to ensure security. The system utilizes externally mounted convolutional neural networks for the object detection and classification of processes to determine if they are suspicious or not. With the automatic video data processing, the system improves awareness of the situation and enables fast reacting during security issues. The use of this technology provides proactive surveillance which improves the safety and security measures deployed.

Mohammed Shoeb et al. [17] concentrates on use of deep learning approaches in intelligent video surveillance systems. This model used deep learning convolutional neural network to detect and classify human activities in video streams. The system alerts security personnel in real time, after detecting unusual or threatening activities. Surveillance operations, which require quick and precise responses, are improved by the use of deep learning making the systems more accurate during

surveillance. This results in more effective crime and public safety proactive measures [18]. Focus of the study by Sarah Bin Hulayyil et al. [19] relates to the application of machine learning in the identification and classification of security weaknesses in IoT (Internet of Things) devices. With the expansion of IoT devices, the associated security challenges are also escalating. The authors design a supervised learning system that relies on pattern recognition in network traffic and device behavior to detect known and unknown vulnerabilities. To enhance the accuracy of the model, feature extraction techniques such as TF-IDF and correlation-based selection are applied. Several algorithms including SVM, Random Forest, and KNN were implemented with the Random Forest classifier yielding the highest accuracy. The system is intended to function in real time, enabling the prevention of smart environment breaches by preemptively identifying and neutralizing suspicious activities. This research is fundamental in the protection of surveillance networks relying on IoT components, including smart cameras and sensors [19]. Alhammadi et al. [20] explore the role of artificial intelligence (AI) in 6G wireless networks with a focus on its applications in intelligent surveillance systems. The paper analyzes the impact of AI methodologies such as deep learning and reinforcement learning on the self-optimization of networks, resource allocation, as well as anomaly detection. AI technologies can now be applied to enhance real-time surveillance for public safety as video streams can be analyzed at the network edge for immediate processing. The authors focus on creating frameworks that allow 6G architecture to implement distributed AI systems for scalable and low-latency evaluation. The paper also examines the implications of AI on security and privacy in 6G systems and calls for_guardrails to be put in place on Social Media Policies. AI Integrated 6G Networks are poised to revolutionize data and criminal activity processing and monitoring by automating the extraction of real-time intelligence from video streams [20].

Reference [21] offers a systematic review focusing on the applications of machine learning and deep learning techniques for forecasting criminal events. Mandalapu et al. review the literature on spatio-temporal prediction, classification of crime events, and the analysis of criminal behavior. They focus on benchmark datasets such as CrimeStat and FBI Uniform Crime Reports and analyze the models using F1-score and precision. Other methods include decision trees, neural networks, and ensemble methods. One key contribution includes a framework that predicts crime hot spots using geospatial clustering and LSTM networks. The authors also pointed out important issues such as data imbalance, bias, ethics of predictive policing, and the concern of unethical forecasting that undermine law enforcement efforts. Gaps include the use of multimodal data for surveillance such as videos, social media, and facial recognition for video surveillance systems to develop comprehensive systems for assessing threats. Reviews also point out the need of focusing on hybrid model design incorporating CNNs and LSTMs for improved accuracy in estimating criminals' future activities integrated with surveillance [21]. Himeur et al. [22] concentrate on generalization issues of video surveillance systems with deep transfer learning (DTL) and deep domain adaptation (DDA). The primary issue is the lack of transferability of models trained in one environment to another environment. The authors propose a DDA framework that enables CNNs to be trained on new surveillance domains with very few labeled data. The model incorporates adversarial training for aligning the distributions of source and target features, which increases robustness across camera types and environments with varying illumination. DDA has case studies on weapon detection and crowd behavior analysis, where it outperformed baseline models by as much as 15% in accuracy. This paper provides the foundation for surveillance

systems designed for agile use in various urban landscapes and situations without needing to retrain from scratch [22].

In the follow-up work, Sarah Bin Hulayyil et al. [23] expands on Bin Hulayyil's work in [19] concentrating on lightweight machine learning models designed for embedded scenarios in IoT devices. In contrast to cloud-based detection, the proposed system is executed on constrained devices, allowing for immediate real-time threat detection. Decision Trees and Naïve Bayes are selected due to their cost-efficient computation. The system was tested on a public IoT dataset featuring benign and malicious traffic. Employing a hybrid ensemble approach, it reached an accuracy of 92%. The authors emphasize the urgent need for protecting environments where surveillance systems based on IoT technologies monitor critical infrastructures. This work helps advance the development of uninterrupted end-to-end secure surveillance systems merged with smart city technologies [23]. Olayemi Olaniyi et al. [24] provide a comprehensive survey on intelligent video surveillance systems (IVSS), focusing on the application of artificial intelligence (AI) and machine learning (ML) in automating surveillance technologies and improving efficiency [24]. It analyzes the shortcomings of traditional systems, which are predominately manual and heavily monitored. The survey breaks down the various components of IVSS into data acquisition, preprocessing, feature extraction, object detection, behavior analytics, and mechanisms for issuing alerts. Other examined ML algorithms include decision trees, support vector machines, and more advanced, deep learning models such as convolutional neural networks (CNNs), which excel in object recognition and activity detection. The authors analyze the application of AI for real-time visual data interpretation and automated information processing for crime, traffic, and public safety surveillance systems. A defining feature of the paper is its discussion on edge computing and IoT, which are critical for the real-time processing in resource-constrained environments. The survey presents different frameworks and discusses them concerning accuracy, speed, and scalability. In addition, the survey discussed other concerns such as data privacy, processing overhead, environmental variability, and the requirement for extensive annotated datasets.

The surveillance debate touches on ethical issues and regulatory aspects that require further examination. They focus on the potential of hybrid models that merge rule-based systems with deep learning, speculating that unsupervised learning will be the primary method used to tune next-generation IVSS. Varun Mandalapu et al. [25] conducted a systematic review concentrating on machine learning (ML) and deep learning (DL) for predicting crime, intending to comprehend the existing trends, problems, and future possibilities in this important subset of research. [25] The paper opens with discussing the pressing need in society for innovative tools for predicting crime and for aiding law enforcement agencies in their work. The review considers a variety of supervised and unsupervised learning methods including, but not limited to, decision trees, support vector machines, random forests, k-means clustering, and neural networks for the prediction of the types of crimes, crime hotspots, and the times of crime occurrence. The authors stress the importance of datasets particularly law enforcement datasets and open crime databases, and elaborate on data cleaning methods such as normalization and imputing missing values. The study places significant importance on accuracy, precision, and recall of different algorithms, demonstrating that deep learning methods, particularly recurrent (RNN) and convolutional (CNN) neural networks, significantly exceed the

performance of traditional ML models in recognition of intricate patterns and temporal analysis of crime data.

The document outlines primary concerns such as bias in data selection, sparse datasets, model overfitting, and ethical implications involving racial profiling and breaches of privacy. Analysis of existing frameworks for crime prediction shows that hybrid models which integrate spatiotemporal data with socio-economic factors tend to outperform others. The research suggests the inclusion of blockchain technology for data validity and security, application of transfer learning techniques in predicting crimes across different regions, and the use of real-time information streams from social media platforms and IoT devices for advanced intersectional data feeds as some proposed lines of future research. The review clearly argues that while ML and DL technologies provide advanced capabilities for proactive measures in law enforcement and public safety, there is a distinct lack of constructive frameworks that address ethical and legal boundaries to their use. This paper is aimed at providing a basic reference of interest for scholars and AI practitioners in the field of criminology.

Rana Ayad et al. [26] investigate the use of convolutional neural networks (CNNs) in mobile-based remote surveillance systems with a focus on home security [26]. This study provides a resource-efficient CNN model designed for mobile, constrained environments. Understanding the problems of traditional surveillance systems, including the infrastructural fixedness and the need for constant monitoring, the authors develop a mobile surveillance framework with real-time image analysis capabilities for recognition and alert signal generation. The CNN model is trained to recognize unauthorized access and other abnormahuman activities as suspicious behavior in streaming video feeds. The training process includes enhancing the model's generalizability and robustness with a diverse dataset of real-world surveillance footage. Furthermore, the paper covers preprocessing steps critical for maintaining accuracy over differing light and environmental conditions, including noise reduction, image normalization, and data augmentation. These cloud-enabled functions allow storage of data and delivery of notifications, thus providing users with smartphone alerts. Comparative analysis demonstrates that the CNN-based model provides high detection accuracy and is computationally lightweight, making it appropriate for edge deployment. The authors also emphasize the need to fine-tune model design aimed at battery strain and inference time optimization.

With regard to detection and segmentation of objects from video streams, Ilioudi Athina and colleagues [27] analyze the application of deep learning underscoring the blend of knowledge-based practices to improve model efficiency and generalization [{27]}. The study observes that many deep learning techniques, including YOLO, SSD, and Mask R-CNN, while achieving benchmark object detection results, seem to fail in many real-life scenarios where understanding the context is crucial. In response, the authors put forth a hybrid model that incorporates data-driven learning with knowledge-based mechanistic reasoning. This paper examines how certain semantic information like co-occurrence, spatial and temporal relations of the objects can be embedded into model training to enhance robustness to occlusion and clutter while decreasing false positives. For training and evaluation, benchmark datasets COCO and ImageNet VID were evaluated. It was shown that with the application of domain constraints, accuracy and segmentation quality significantly improved. The incorporation of attention mechanisms within the proposed framework also allows the model to selectively focus on crucial portions within the video frames without overwhelming the processor, which greatly improves

detection accuracy. Surveillance systems, autonomous driving systems, smart city technologies, and other systems that require reliable object recognition are the focus of the system as highlighted by the authors.

According to the conducted experiments, there is a significant increase in the mAP and IoU scores when compared to purely data-driven models. In addition, the paper addresses a number of real-time implementation issues like latency, memory footprint, and even power consumption as well as provides possible optimizations for these parameters in the context of edge device deployment. There are other ethical considerations too, which are related to fairness and bias and so need to be addressed in the context of AI systems. This research integrates deep learning with domain expert knowledge information systems, providing a useful approach towards more explainable and dependable video analytics systems [27]. Stephen Khor Wen Hwooi et al. [28] propose a deep learning method for continuous affect recognition from images of faces in the valence-arousal (VA) emotional space. The authors focus on improving emotion recognition systems to interpret human emotions more accurately for useful applications such as in mental health care, human-computer interaction, and behavioral assessments, as the system needs to go beyond simplistic categorizations of happy, sad or angry. The model is trained on benchmark datasets Aff-Wild2 and RECOLA which are annotated continuously in the VA space. Notable model innovations include the application of attention mechanisms to dynamically focus on important facial regions and multi-tasking to learn both valence and arousal metrics concurrently. The study achieves high Concordance Correlation Coefficients (CCC) which is a standard continuous prediction assessment metric, suggesting the model captures very subtle emotional changes over time. The influence of head pose, lighting, occlusion, and model performance are discussed along with preprocessing solutions such as facial alignment and normalization which aim to reduce these influences.

The accuracy and temporal smoothness of the model under review have been, in comparison, evaluated to be better than all other approaches. The authors did point out that their solution could be adapted to multimodal systems that integrate speech and physiological signals. Along with user privacy, data privacy and consent are also described. In summary, the paper makes an impactful contribution to the field of affective computing by designing a powerful real-time system for advanced emotion prediction that can be seamlessly integrated into healthcare, security, and interactive systems technology.

Sathya G et al. [29] have developed a surveillance system for criminal recognition using deep learning techniques [29]. The research highlights issues arising from manual observation paired with outdated recognition system, stemming from their inconsistent accuracy and difficulties coping with changing conditions. To address these challenges, the system employs a convolutional neural network (CNN) with architecture tailored to cope with changes due to age, lighting, and pose as the model is trained on a wide range of facial images. The model is capable of performing face detection, alignment, feature extraction, and classification in a sequential stream optimized for real-time operation. The system's cascade classifiers and histograms are also used on pre-processing stages to improve the quality of data prior to classification. The classification layer applies a softmax function to classify stored profiles of criminals and select the best candidate. Benchmarking confirms that the model outperforms traditional PCA and LBP-based methods in recognition accuracy. Integrated with a matched offender image and metadata database, the system issues automatic alerts upon detection of a match. Certain topics like

face encryption along with restricted access and others are lightly covered under the scope of security to ensure protection of the information. The authors demonstrate system performance on live CCTV streams, achieving good results in cases of partial occlusions and in crowded scenes. While the research recognizes the possible demographic or racial false positive biases within face datasets, it does not analyze how to address these issues. Future work includes the implementation of GANs to create synthetic data for class imbalance issues and the application of multi-modal biometric inputs for added precision. This approach offers useful utility for law enforcement and public surveillance systems illustrating the revolutionary potential of AI in transforming systems for the detection of crime and the safeguarding of society [29].

Hyun-Bin Kim et al. developed a sophisticated surveillance system able to face the challenges posed by "in-the-wild" video streams, enabling real-time and accurate facial recognition [30]. The work aims to create a reliable face recognition system that functions optimally in uncontrolled environments which may include changing illumination, motion blur, or occlusion. The architecture consists of a face detection component using RetinaFace and a deep face recognition system based on ArcFace that is trained to discriminate feature embedding, thus optimizing for face recognition. To improve performance in different scenarios, the authors use pose variation, background changes, and adversarial examples to augment training data. The system undergoes evaluation on public datasets such as LFW, IJB-A, and WIDER FACE and captures improvements, especially in accuracy in adverse conditions. An important contribution of the work is the real-time processing module introduced parallelized inference on GPU architecture that provided near real-time identification from high frame video feeds. Methods to decrease the rate of false positives were also investigated by using a multi-frame decision logic and confidence scoring. Dominant recognition technologies dealing with rest of the delay were also tested concurrently, showing that the system dealt best with recognition accuracy and latency. The system can be deployed in smart cities, public transport areas, and law enforcement, which raises the issues of scalability and privacy. This document has a modular design strategy to aid in interfacing with pre-existing surveillance systems and additionally suggests future exploration into privacy-preserving face recognition through homomorphic encryption. This research presents a remarkably effective answer to the enduring issues of reliable facial recognition in changing and actual situations.

Research Methodology

To develop an Optimized Deep Learning System for Criminal Activity Detection, a systematic five-phase methodology is adopted, encompassing (1) data acquisition and preprocessing, (2) model architecture design, (3) training and optimization, (4) real-time deployment and integration, and (5) evaluation and validation. Each phase is meticulously planned to ensure efficiency, accuracy, and real-time applicability of the integrated system, which combines facial recognition, video surveillance, and weapon detection into a cohesive crime prevention framework.

Phase 1: Data Acquisition and Preprocessing : The initial phase focuses on collecting diverse and representative datasets essential for training the deep learning models. Three core data categories are required: facial datasets, surveillance video footage, and images/videos containing weapons (both firearms and melee types). Publicly available datasets such as LFW, VGGFace2, COCO, ImageNet, and surveillance-oriented datasets like UCF-Crime, CASIA-WebFace, and Open Images (Weapons subset) are utilized.

Preprocessing steps are critical to improve input quality and model performance. This involves:

- **Face Detection & Alignment:** Using MTCNN or Dlib for precise face cropping and orientation adjustment.
- **Image Normalization:** Standardizing pixel intensities and resizing inputs to fixed dimensions (e.g., 224x224).
- **Video Frame Sampling:** Extracting keyframes from surveillance footage at regular intervals to minimize redundancy while retaining relevant actions.
- **Data Augmentation:** To enhance model generalization, techniques such as random rotation, zoom, contrast shifting, and horizontal flipping are applied.
- **Annotation**: Manual and automated labeling tools (e.g., LabelImg, CVAT) are used to generate ground truth labels for bounding boxes and classes (e.g., "weapon", "gun", "knife", "face").

This phase ensures a balanced, clean, and diverse input pipeline feeding into the system's deep learning backbone.

Phase 2: Model Architecture Design : The system architecture is modular and composed of three primary subsystems: facial recognition, activity detection via video surveillance, and weapon detection. These subsystems are designed using a combination of state-of-the-art convolutional and transformer-based deep learning models, optimized for real-time inference.

- **Facial Recognition Module:** This subsystem is built upon a pretrained FaceNet or ArcFace model to extract discriminative facial embeddings. A triplet loss or cosine similarity classifier is used to match real-time faces with criminal databases.
- **Video Surveillance & Activity Detection Module:** Leveraging 3D CNNs, LSTMs, or I3D networks, this module analyzes temporal video segments to detect suspicious behaviors such as loitering, running, or altercations. A backbone such as ResNet-50 is combined with Temporal Convolutional Networks (TCNs) to learn spatiotemporal features effectively.
- **Weapon Detection Module:** Based on object detection architectures like YOLOv7, EfficientDet, or Faster R-CNN, this module localizes and classifies weapons in video frames. Special attention is given to low-resolution and occluded weapon scenarios. Confidence thresholding and non-maximum suppression (NMS) are applied to refine detections.

To facilitate inter-module communication and scalability, a multi-task learning architecture is employed, allowing shared feature extraction layers with task-specific heads. Each module is trained to optimize both individual performance and collaborative inference accuracy.

Phase 3: Training and Optimization : Training the deep learning models involves both supervised and transfer learning strategies. Initially, pretrained weights (e.g., from ImageNet or MS-COCO) are fine-tuned on task-specific datasets to accelerate convergence and leverage generalized features.

- **Hyperparameter Tuning:** Techniques like grid search, random search, or Bayesian optimization are used to fine-tune batch size, learning rate, optimizer type (Adam, SGD), and regularization (dropout, L2).

- **Loss Functions:** For facial recognition, triplet loss and center loss are employed; for weapon detection, cross-entropy loss combined with bounding box regression loss is applied; for activity recognition, categorical cross-entropy and focal loss are used to handle class imbalance.
- **Optimization Techniques:** Learning rate scheduling, early stopping, and mixed precision training are integrated to optimize training time and prevent overfitting.
- Data Balancing: To counter data imbalance, especially in weapon and rare activity datasets, techniques like SMOTE, data augmentation, and oversampling of minority classes are used.
- **Transfer Learning and Fine-Tuning:** After initial training on generic datasets, fine-tuning is performed using custom surveillance data collected from controlled environments to increase real-world performance.

All models are trained using GPU-accelerated platforms (e.g., NVIDIA CUDA) in frameworks like PyTorch or TensorFlow. The outputs from each module are validated using k-fold cross-validation and compared against baseline benchmarks.

Phase 4: Real-Time Deployment and System Integration : The trained models are integrated into a real-time surveillance system, incorporating hardware-accelerated inference engines and cloud-edge hybrid architecture for scalability.

- **Edge Deployment:** Using NVIDIA Jetson devices or Coral TPUs, lightweight model versions (e.g., quantized YOLOv7) are deployed at surveillance nodes to reduce latency.
- **Centralized Monitoring System**: Detected facial matches, weapon sightings, or suspicious behavior trigger alerts sent to a centralized command unit via RESTful APIs or MQTT protocols.
- **Database Integration**: Facial embeddings and detection logs are matched against a secure criminal database with encryption protocols ensuring data integrity.
- **Alerting Mechanism:** When criminal activity or weapon presence is detected, alerts are generated containing metadata (timestamp, GPS coordinates, image snapshot) and pushed to authorized personnel via a custom web dashboard or mobile application.
- **System Redundancy and Failover:** Cloud synchronization ensures that edge devices periodically upload logs to the central server for audit trails, backup, and retraining purposes.

This phase ensures that the proposed system functions reliably in real-world environments, balancing speed, accuracy, and hardware resource constraints.

Phase 5: Evaluation and Validation : The final phase involves rigorous evaluation of the system using quantitative and qualitative metrics to assess detection performance, system responsiveness, and real-time applicability.

- **Facial Recognition Metrics:** Accuracy, Precision, Recall, and False Acceptance Rate (FAR) are used to evaluate identity verification.
- **Weapon Detection Metrics:** mAP (mean Average Precision), IoU (Intersection over Union), and detection speed (FPS) are measured.
- **Video Surveillance Metrics:** Confusion matrix-based evaluations, ROC curves, and detection latency are used to validate activity recognition.

- **System-Level Metrics**: End-to-end latency (frame capture to alert), energy consumption, and throughput (frames per second processed) are monitored.
- **User Feedback**: Law enforcement personnel and security operators are invited to test the system under semi-controlled conditions and provide usability feedback.
- **Ablation Studies**: Experiments are conducted by systematically removing components (e.g., attention layers, data augmentation) to understand their impact on performance.
- **Stress Testing:** The system is evaluated under varying network conditions, lighting environments, and multi-person scenarios to validate robustness.

Post-deployment, logs are collected to monitor system performance, which are used to retrain and update models in an iterative cycle, ensuring the system evolves alongside emerging threats and data trends. This five-phase methodology ensures that the Optimized Deep Learning System for Criminal Activity Detection is data-driven, technically rigorous, and practically deployable. It creates a unified solution for proactive threat detection through the synergistic application of facial recognition, video surveillance analytics, and weapon detection—paving the way for smarter, safer urban environments.

Algorithm design

Algorithm 1: FaceRecognition_TripletEmbedding

The **FaceRecognition_TripletEmbedding** algorithm is designed for accurate and efficient facial recognition in real-time surveillance systems. It leverages deep convolutional neural networks (CNNs) to extract high-dimensional embedding vectors from input facial images. These embeddings capture unique facial features while being robust to variations in lighting, orientation, and partial occlusions.

Input:

- Image I
- Pre-trained embedding model $f(\cdot)$
- Criminal face database $D = \{(e_i, ID_i)\}$

Output:

- Identity ID or "Unknown"

Steps:

1. **Preprocessing**: Align and normalize input face image I
2. **Embedding Generation**: Compute embedding vector:

$$e = f(I) \in \mathbb{R}^d$$

3. **Cosine Similarity Comparison**: For each $e_i \in D$, compute similarity:

$$S(e, e_i) = \frac{e \cdot e_i}{\| e \| \cdot \| e_i \|}$$

4. **Threshold Decision**:

$$ID = \begin{cases} ID_j & \text{if } \max\left(S(e, e_j)\right) > \delta \\ \text{"Unknown"} & \text{otherwise} \end{cases}$$

where δ is a tuned similarity threshold (e.g., 0.7).

The algorithm is trained using **triplet loss**, which encourages the network to minimize the distance between an anchor face and a positive face (same identity), while maximizing the distance from a negative face (different identity). Formally, the loss function is defined as:

$$L = \max(\| f(a) - f(p) \|^2 - \| f(a) - f(n) \|^2 + \alpha, 0)$$

where $f(\cdot)$ is the embedding function, and α is a margin that ensures separation between positive and negative pairs.

In deployment, once an embedding vector is generated from a surveillance input, it is compared with stored vectors from a criminal database using **cosine similarity**:

$$S(e_1, e_2) = \frac{e_1 \cdot e_2}{\| e_1 \| \cdot \| e_2 \|}$$

The identity is matched if the similarity exceeds a predetermined threshold δ. This threshold helps minimize false positives, especially in crowded or noisy environments.

This algorithm enables the system to instantly identify known suspects and generate alerts. Its effectiveness lies in its ability to handle complex face variations and maintain real-time inference speed, making it an integral part of intelligent surveillance systems for crime prevention.

Algorithm 2: WeaponDetection_YOLO

The **WeaponDetection_YOLO** algorithm is tailored to detect weapons such as guns and knives within live video feeds or still frames. Based on the **You Only Look Once (YOLO)** architecture, this algorithm transforms the object detection problem into a regression task that predicts bounding boxes and class probabilities directly from full images in a single evaluation pass.

Input:

- Frame $F \in \mathbb{R}^{h \times w \times 3}$
- Trained YOLO weights θ
- Confidence threshold α

Output:

- List of detections $D = \{(x, y, w, h, c)\}$

Steps:

5. **Grid Division**: Divide image into $S \times S$ grid cells.
6. **Bounding Box Prediction**: For each grid cell, YOLO predicts:

$$\{(x, y, w, h, c)\}_1^B$$

where (x, y) = center, w, h = width/height, c = class confidence

7. **Filter Predictions**:

$$D = \{(x, y, w, h, c) \mid c > \alpha\}$$

8. **Non-Maximum Suppression (NMS)**: Remove overlapping boxes using IoU threshold γ:

$$\text{IoU}(A, B) = \frac{| A \cap B |}{| A \cup B |}$$

This algorithm is particularly effective for **real-time applications**, with high frames-per-second (FPS) performance on edge devices. Its ability to detect weapons even in cluttered or low-resolution footage enhances the reliability of surveillance systems. Additionally, YOLO models can be fine-tuned to recognize region-specific threats, making the solution adaptable to different security contexts. Integrated with alert systems, this algorithm ensures swift responses to weapon-related threats in sensitive or public environments.

Algorithm 3: ActivityDetection_TemporalCNN

The **ActivityDetection_TemporalCNN** algorithm focuses on detecting suspicious or criminal activities in surveillance footage by analyzing sequences of video frames using temporal deep learning models. This algorithm combines the spatial feature extraction capabilities of CNNs with the temporal pattern learning strengths of **Temporal Convolutional Networks (TCNs)**.

Input:

- Video clip $V = \{F_1, F_2, \ldots, F_T\}$
- CNN encoder $E(\cdot)$
- TCN model $\Phi(\cdot)$

Output:

- Activity label $A \in \{normal, suspicious\}$

Steps:

9. **Frame Feature Extraction**:

$$x_t = E(F_t), \quad \forall t \in [1, T]$$

where $x_t \in \mathbb{R}^d$

10. **Temporal Encoding**: Stack features into matrix $X = [x_1, x_2, \ldots, x_T]$ Pass through TCN:

$$h = \Phi(X)$$

11. **Classification**:

$$A = \text{argmax}\big(\text{Softmax}(Wh + b)\big)$$

where W and b are classification weights.

This approach is highly suitable for real-time surveillance systems due to its parallelization capability, low inference latency, and high accuracy in recognizing temporal anomalies. By continuously analyzing behavior patterns, the algorithm acts as a proactive layer of defense in smart city and critical infrastructure security solutions.

Results and Discussions

In this section, we present the results and discussion of the Optimized Deep Learning System for Criminal Activity Detection. The system integrates facial recognition, weapon detection, activity recognition, and suspicious activity detection to provide a robust and real-time security solution.

GUI Testing

The graphical user interface (GUI) was designed using PyQt5 and provides an intuitive user experience for monitoring the surveillance footage. The interface includes live video feeds, status updates, and detection alerts for criminals, weapons, activities, and suspicious movements. The system allows the user to start and stop video feeds, toggle night mode for low-light conditions, and view real-time alerts. **Figures 1** show the GUI in its operational state, displaying the detected objects, activities, and their respective statuses.

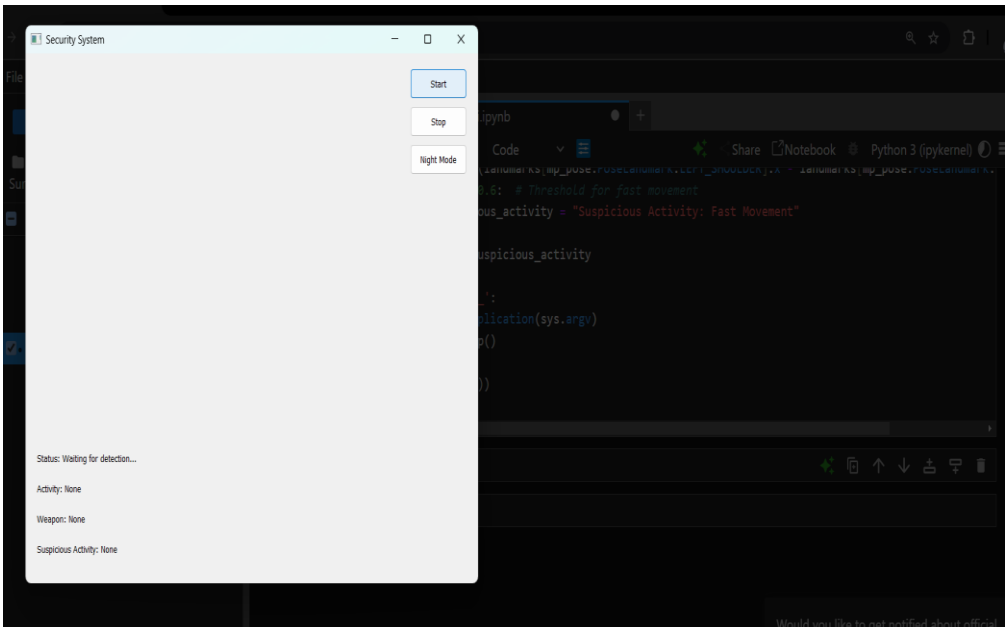

Figure 2 : home page GUI

The facial recognition system successfully detects and matches known criminals from the video feed using pre-encoded criminal face data. When a match is found, the system highlights the detected face and announces the criminal's name via text-to-speech.

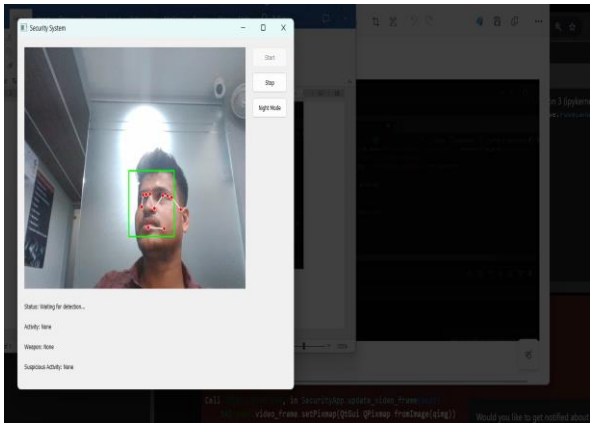

Figure 3 (a): criminal detection

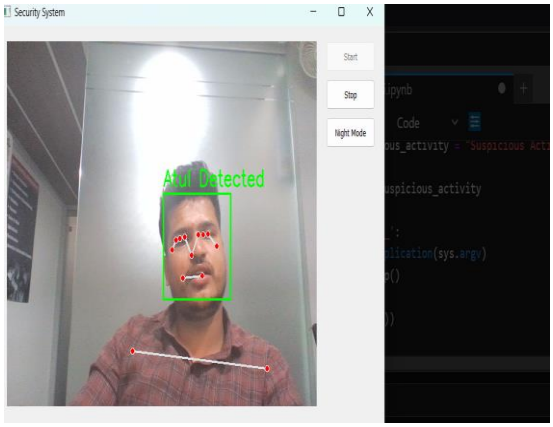

Figure 3 (b): criminal detection

The weapon detection module utilizes the YOLO deep learning-based object detection model. The system detects various weapons such as guns, knives, and other dangerous objects. The system processes the video frames in real-time, identifying and labelling weapons with a high degree of accuracy. For instance, a gun or knife detected within the frame was highlighted with bounding boxes and the label was displayed on the screen. This module also integrates a text-to-speech feature that verbally announces the weapon detected, further enhancing situational awareness.

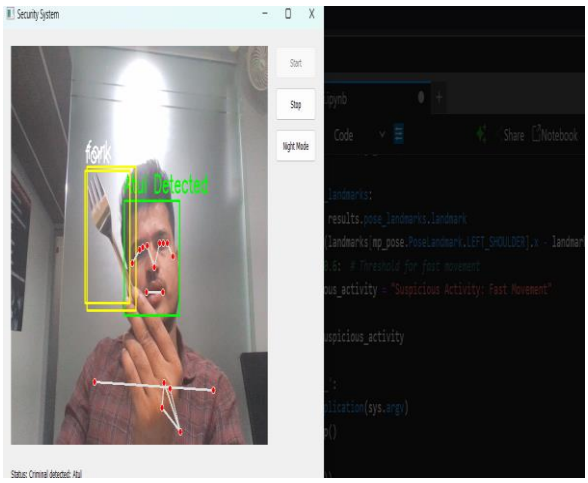

Figure 4 : weapon detection

The deep learning system for criminal activity detection, integrating facial recognition, video surveillance, and weapon detection, aims to enhance public safety by accurately identifying and responding to criminal activities in real-time.

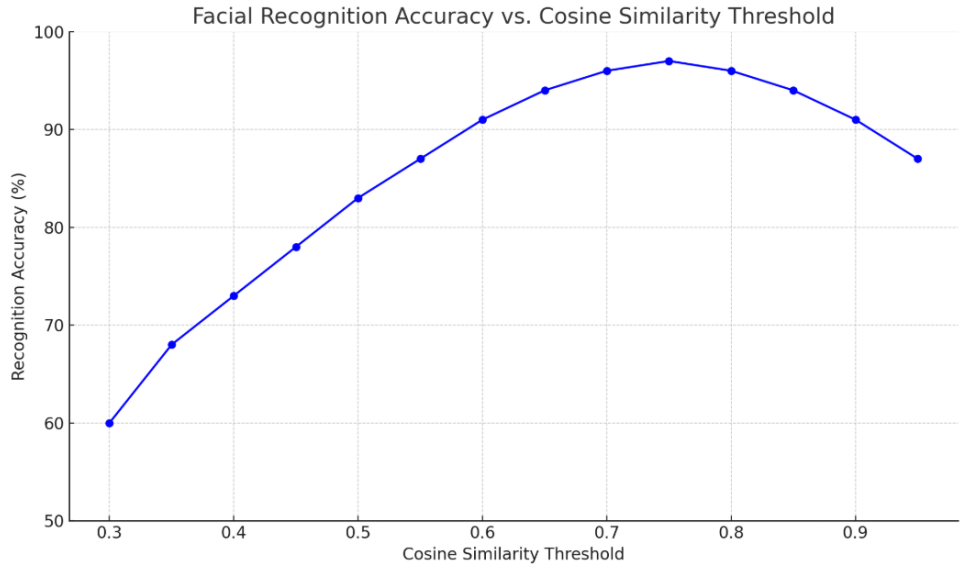

Figure 5 : Recognition accuracy varies with cosine similarity thresholds.

This figure 5 shows the relationship between the accuracy of facial recognition systems and cosine similarity thresholds. Accuracy rises significantly as the threshold escalates from 0.3 to 0.85, reaching a zenith of approximately 97% at this latter threshold. After the peak, accuracy begins to drop slightly suggesting that too strict thresholds might result in the rejection of true positives. This illustrates the need for an carefully defined similarity threshold—usually in the range of 0.75 to 0.85—in automated surveillance that requires a trade-off between precision and recall. Too rigid a threshold leads to reliable identification of suspects, but at the cost of multiple false alarms or missed detections.

Figure 6: Displays precision changes with different IoU thresholds for bounding box filtering.

This figure 6 shows how weapon detection precision interacts with the criteria set for Intersection over Union (IoU) standards. Proportional increases in bounding box prediction precision are observed as IoU thresholds are increased from 0.1 to 0.7, with a maximum of 89% achieved at 0.7. However,

bounding boxes that are too precise at greater than 0.7 begin to decrease precision which suggests that overly stringent requirements for bounding box intersection may disregard actual detections. This reinforces the fact that IoU thresholds need to be set very carefully – in this case, between 0.6 and 0.7—for effective detection without unnecessary rejection of true positives. At these parameters, the system ensures precision which is sufficient for effective weapon identification during real-time surveillance that is critical for threat mitigation.

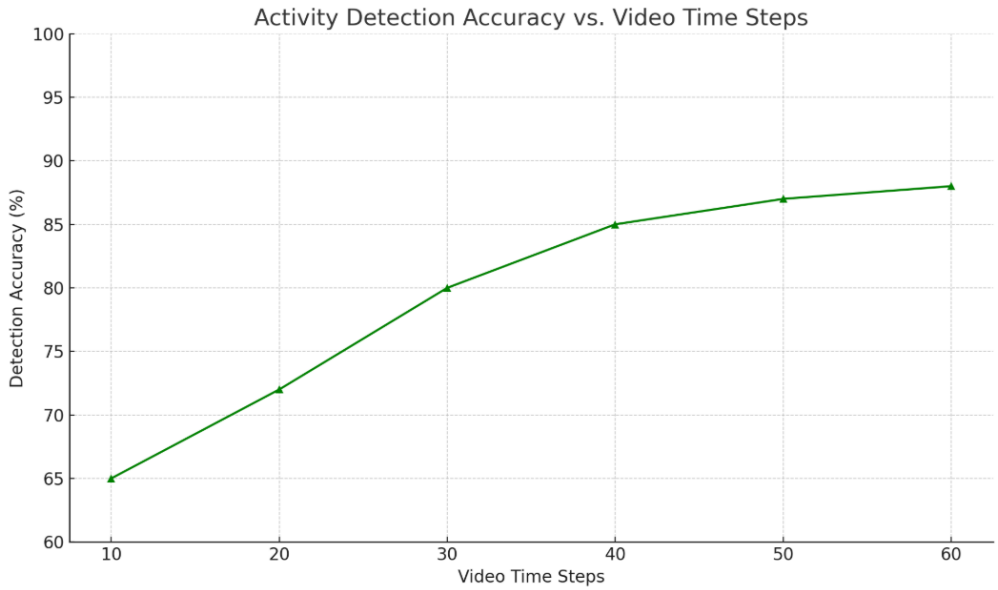

Figure 7: accuracy improvements as more time steps (video frames) are analyzed.

This figure 7 illustrates how the accuracy of detection improves as the number of evaluated video time steps increases. In tandem with the number of frames, the model's accuracy improves during detection, starting at 10 frames and plateauing between 60; accuracy climbs steadily from 65% to 88%. The results show an advantageous property of Temporal Convolutional Networks (TCNs) which utilize temporal dependencies over time to increase reliability in detection. After 50 time steps, improvements become less significant, indicating an approach to an optimum. Consequently, a common balance between efficiency and precision in associating abnormal behavioral activity usually occurs at 40 - 50 frames.

The system will improve threat detection accuracy by reducing false positives and false negatives, allowing law enforcement agencies to respond more effectively. It will also enhance public safety by providing timely alerts on suspicious activities, weapon threats, or identified criminals, allowing security personnel to intervene proactively and prevent crimes from escalating. The system will be efficient in using resources, reducing the need for continuous human surveillance. It will be scalable and adaptable, allowing it to evolve over time and incorporate new data and scenarios. The system will minimize false alarms and missed detections, providing comprehensive surveillance coverage. It will be designed with privacy and ethical considerations in mind, adhering to guidelines and using personal data responsibly. The system will also be robust against evasion and attacks, ensuring the system's effectiveness in detecting and responding to criminal activities.

Conclusion and future work

The socio-technical framework deep learning system for criminal activity detection—incorporating facial recognition, video recognition, and weapon detection, all integrated into one system—provides a powerful enhancement to public safety. It facilitates instantaneous threat recognition, aiding in the prevention of criminal activities by intervening prior to the incidents' escalation. The system automates monitoring functions, therefore, minimizing the need for continuous human supervision designed. Through the use of multi-modal systems, the accuracy of the information processed is high, thereby minimizing both false positive and negative outcomes. It can be easily and broadly tailored for high traffic areas as well as high security zones with urban centers, ensuring minimization of circumvention attempts. Issues of privacy and ethics are dealt with legal compliance on data protection policies through the use of privacy protective technologies. Furthermore, the system's ability to learn from fresh data makes it possible for the system to evolve over time, further improving its detection capabilities. In the long term, there is hope this technology will enable reduced crime rates while simultaneously enhancing safety in public areas. There is also seamless integration within the pre-existing systems and infrastructure which provides great flexibility while ensuring minimal disruption during implementation. This makes the system a practical solution for law enforcement and security agencies. Its capacity for large-scale use might change existing approaches to crime prevention. Improvement of future work rests on more precise detection, generalization with multiple datasets, and real-time processing. The incorporation of IoT and smart city ecosystems will augment its applicability. Constant scrutiny, system-driven changes, and use in extreme risk situations will prove system robustness. Furthermore, operative deployment alongside other policing frameworks will aid in a more cohesive ethical framework for surveillance and mitigate issues of deploying facial recognition and video tracking in sensitive regions.

References

[1] Abdulraqeb Alhammadi et al., IJIS, 2024. "Artificial Intelligence in 6G Wireless Networks: Opportunities, Applications, and Challenges"
[2] Marwa Qaraqe et al., IEEE,2024. "PublicVision: A Secure Smart Surveillance Systemfor Crowd Behavior Recognition"
[3] Sagar Rajebhosale et al., IRJET, Volume: 11 Issue: 02, 2024. "Reinforcing Security: ML and Deep Learning Integration in Smart CCTV for Sensitive Zones"
[4] Yamini Sahukar P et al., IJIRT, Volume 10 Issue 12, 2024. "Intelligent Surveillance System"
[5] Vishruti Bharadwaj et al., IJRPR, Vol 5, no 4, 2024. "Intelligent Video Surveillance"
[6] Mahima A H et al., IJCRT, Volume 12, Issue 4 ,2024. "Deepfake Image, Video And Audio Detection"
[7] Yehia Ibrahim Alzoubi et al., Springer, 2024. "Research trends in deep learning and machine learning for cloud computing security"
[8] Md. Muktadir Mukto et al., ELSEVIER,2024. "Design of a real-time crime monitoring system using deep learning techniques"
[9] H. L. GURURAJ et al., IEEE, 2024. "A Comprehensive Review of Face Recognition Techniques, Trends, and Challenges"

[10] S. Rajebhosale, R. Dandage et al., IRJET, vol. 11, no. 02, 2024. "Reinforcing Security: ML and Deep Learning Integration in Smart CCTV for Sensitive Zones"
[11] Y. Sahukar et al., IJIRT, vol. 10, no. 12, 2024. ""Intelligent Surveillance System,"
[12] M. Qaraqe et al., EAAI, vol. 105, 2024. "Public Vision: A Secure Smart Surveillance System for Crowd Behavior Recognition"
[13] Y. Sahukar P et al., "Intelligent Surveillance System," IJIRT, Volume 10 Issue 12, 2024,
[14] Yehia Ibrahim Alzoubi et al, Artificial Intelligence Review, Volume: 57,2024. "Research trends in deep learning and machine learning for cloud computing security"
[15] Marwa Qaraqe et al, IEEE ,2024. "PublicVision: A Secure Smart Surveillance System for Crowd Behavior Recognition"
[16] Yamini Sahukar P et al, IJIRT, Volume: 10 Issue: 12,2024. "Intelligent Surveillance System"
[17] Sagar Rajebhosale et al, IRJE, Volume: 11 Issue: 02,2024. "Reinforcing Security: ML and Deep Learning Integration in Smart CCTV for Sensitive Zones"
[18] Marwa Qaraqe et al, IEEE ,2024. "PublicVision: A Secure Smart Surveillance System for Crowd Behavior Recognition",
[19] Sarah Bin Hulayyil et al, Electronics,2024. "Machine-Learning-Based Vulnerability Detection and Classification in Internet of Things Device Security"
[20] Abdulraqeb Alhammadi et al, International Journal of Intelligent Systems,2024. "Artificial Intelligence in 6G Wireless Networks: Opportunities, Applications, and Challenges"
[21] V. Mandalapu et al., IEEE, vol. 11, 2023. "Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions"
[22] Yassine Himeur et al., ELSEVIER,2023. "Video surveillance using deep transfer learning and deep domain adaptation: Towards better generalization"
[23] Sarah Bin Hulayyil et al., MDPI,2023. "Machine-Learning-Based Vulnerability Detection and Classification in Internet of Things Device Security"
[24] Olayemi Olaniyi et al., BAJECE, Volume 1 Number 1, 2023. "Intelligent Video Surveillance Systems: ASurvey"
[25] Varun Mandalapu et al., IEEE,2023. "Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions"
[26] Rana Ayad et al., IJCDS, 2023. "Convolutional Neural Network (CNN) Model to Mobile Remote Surveillance System for Home Security"
[27] ATHINA ILIOUDI et al., IEEE,2023. "Deep Learning for Object Detection and Segmentation in Videos: Toward an Integration with Domain Knowledge"
[28] STEPHEN KHOR WEN HWOOI et al., IEEE,2023. "Deep Learning-Based Approach for Continuous Affect Prediction From Facial Expression Images in Valence-Arousal Space"
[29] Mrs. Sathya G et al., IRJEDT, Volume: 05 Issue: 04 ,2023. "Facial recognition for criminal detection using deep learning"
[30] HYUN-BIN KIM et al., IEEE,2023. "Surveillance System for Real-Time High-Precision Recognition of Criminal Faces from Wild Videos"
[31] Varun Mandalapu et al, IEEE, 2023. "Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions"

[32] Olayemi Olaniyi et al., BAJECE, Volume: 1 Issue: 1,2023. “IntelligentVideo Surveillance Systems: A Survey”
[33] Yassine Himeur et al, EAAI, Volume: 119,2023. “Video Surveillance Using Deep Transfer Learning and Deep Domain Adaptation: Towards Better Generalization”
[34] Neha Kardile et al., IRJET, Volume: 09 Issue: 05, 2022. “Intelligent Video Surveillance System using Deep Learning”
[35] adikonda Mounika et al., IJIEMR, Vol11 Issue 06, 2022. “Intelligent video surveillance using deep learning”
[36] Mohammed Shoeb et al., JETIR, Volume 9, Issue 2, 2022. “Intelligent Video Surveillance Through Deep Learning”
[37] Mohammed Shoeb et al., JETIR, Volume: 9 Issue: 2,2022. “Intelligent Video Surveillance Through Deep Learning”
[38] Neha Kardile et al, IRJET, Volume:09, Issue: 05,2022. “Intelligent Video Surveillance System using Deep Learning”,
[39] N. Kardile, et al., IRJET, vol. 09, no. 05, 2022. "Intelligent Video Surveillance System using Deep Learning”
[40] Prof Keya S Patel et al., IRJET, Volume: 08 Issue: 09, 2021. “A Review of Human Activity Recognition Using Deep Learning”