Technical Report

Anomaly Detection in CCTV Footage for Snatching Incident Identification

By: Muhammad Ali Khan

Supervisor: Mr. Salahuddin

Islamia College Peshawar

Abstract

This technical report presents an AI-driven framework for anomaly detection in CCTV footage, specifically focused on snatching incident identification involving weapons. Leveraging YOLO models (v5–v12) and a custom dataset of more than 5,000 annotated CCTV images, the system achieves high accuracy in detecting firearms under challenging conditions such as low lighting, occlusion, and motion blur. The developed solution enhances public safety by enabling real-time monitoring and rapid threat response in dynamic urban surveillance environments.

Introduction & Background

Urban surveillance systems employ millions of CCTV cameras worldwide, generating vast amounts of video data. Traditional manual monitoring methods are inefficient and prone to error, especially in high-crime urban areas. The need for scalable, automated solutions drives the adoption of AI-based surveillance systems. This project utilizes advanced YOLO models to enhance real-time weapon detection and contribute to the development of smart city security frameworks.

Problem Statement & Objectives

The rising number of firearm-related snatching incidents in urban environments highlights the urgent requirement for automated real-time detection systems. Manual CCTV monitoring is slow, unreliable, and often fails under poor visibility, motion blur, or occlusion. This project addresses these issues by proposing an AI-driven anomaly detection system. The objectives are: - To develop a robust YOLO-based model for weapon detection. - To train the model on a diverse dataset of CCTV footage. - To achieve high accuracy in real-time detection for public safety.

Dataset & Methodology

A custom dataset of over 5,000 annotated CCTV images was used, capturing real-world snatching scenarios involving firearms. The dataset includes challenges such as low lighting, motion blur, and occlusion. Frames were preprocessed (resized to 420x420 pixels) and augmented for robustness. YOLO models (v8–v12) were fine-tuned using stochastic gradient descent and CIoU loss for 100 epochs on an NVIDIA Tesla T4 GPU. Real-time inference was achieved at 60 FPS with non-maximum suppression (IoU 0.5).

Results & Contributions

The system achieved a peak accuracy of 89.8% (mAP@0.5) using YOLOv10, outperforming baseline detectors. Contributions include: - Introduction of a novel dataset with diverse conditions for weapon detection. - Integration of attention-based mechanisms for spatiotemporal feature learning. - Optimization of loss functions, improving detection performance by 15%. - Development of an ethical, privacy-preserving framework for scalable urban security.

Conclusion & Future Work

This work demonstrates the effectiveness of YOLO-based models in detecting weapons for anomaly detection in CCTV footage. By overcoming challenges of occlusion, low lighting, and motion blur, the framework contributes to proactive crime prevention in urban surveillance. Future work may involve extending the dataset, incorporating human action recognition, and deploying the system in large-scale smart city networks for real-time security enhancement.

References

1. Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767. 2. Jocher, G. (2023). YOLOv8: State-of-the-Art Object Detection. Ultralytics. 3. Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934.