Abnormal Trajectory Type Loitering Detection Using YOLOv9 and CNN

1st Abdul Wasiul Khair
School of Computing
Telkom University
Bandung, Indonesia
abdulwasiulkhair@student.telkomuniversity.ac.id

2nd Gamma Kosala
School of Computing
Telkom University
Bandung, Indonesia
gammakosala@telkomuniversity.ac.id

***Abstract*— Closed-circuit television (CCTV), also referred to as security cameras, plays a crucial role for both preventing and detecting crimes. However, manually monitoring CCTV footage presents significant challenges, especially when managing multiple screens. Missed incidents often occur due to human limitations such as loss of focus, boredom, distractions, and inexperience. Research has been conducted to automate the analysis of CCTV footage to detect unusual events with minimal human intervention. One example of such an unusual event is loitering, where individuals remain in a place without an apparent reason, potentially posing a public security threat. This research proposes an automated loitering detection method to address these challenges. The method focuses on abnormal trajectory patterns, which are characterized by frequent direction changes, zigzag movements, and circling objects without clear intent. The proposed system involves three main steps. First, YOLOv9 is employed to detect individuals in video frames and generate bounding boxes. Second, trajectories are reconstructed by sequentially connecting centroids, and these trajectories are saved as images with black lines on a white background. Finally, the EfficientNet model is used to classify these trajectory images and identify loitering behavior. The proposed method achieves an accuracy of 92.65% in classifying abnormal trajectory type loitering.**

***Keywords*— Loitering Detection, Abnormal Trajectories, YOLOv9, EfficientNet, CCTV.**

I. INTRODUCTION

Security cameras, also known as Closed Circuit Television (CCTV), are an important tool in preventing and investigating crimes [1]. Manual monitoring of CCTV footage faces various challenges if done manually by humans, especially when monitoring multiple CCTV footage [1], [2], [3]. Crimes can occur on monitoring screens that could be unattended because the CCTV supervisor is focused on another monitoring screen. Other factors that can make crime poorly monitored are difficulty maintaining focus, boredom, distraction, and lack of experience [4].

Therefore, an intelligent surveillance system is needed. This system utilizes CCTV installed in various areas and analyzes the video footage captured by the CCTV to identify unnatural events that occur with little help from humans [1], [5], [6]. One of the unusual events is loitering, which refers to the behavior of being in an area with an unclear purpose, staying in public areas for a long time is often considered suspicious and can be a sign of a potential threat to public safety. In addition to loitering, there is also abnormal trajectory loitering, where an individual walks in a zigzag or circular pattern [7]. Intelligent surveillance systems can reduce the number of crimes by detecting possible thieves before committing the act of theft [5].

Research conducted by Wahyono et al. [5] focuses on loitering detection using spatial-temporal data. The method first detects if a human appears in the video frame using features extracted via Histogram of Oriented Gradients (HOG) then classifies the data using Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest. If a human is detected, their movement is tracked to obtain the path traveled. Spatial-temporal data extraction is then performed by calculating the angle at each step. This feature is used for classification to determine whether the video shows normal movement or not.

T Huang et al. [8] propose a loitering detection technique based on classification in pedestrian areas. To identify and track individuals in the video footage, the methods employed are Gaussian mixture models (GMMs) and MeanShift. The coordinate points of their trajectories are then transformed into a defined area. The pedestrian loitering behaviors are categorized into three types: rectangle loitering, ellipse loitering, and sector loitering. The time an individual spends within the area is used to determine whether loitering occurs.

J Núñez et al. [7] define and detect different types of loitering behaviors, including seated loitering, random walk loitering, abnormal trajectory, and no motion. Their method combines motion, stationary, and geometric analyses. Abnormal trajectory loitering is characterized by zigzag walking patterns, frequent turns, and direction changes. This study provides foundational methods for identifying loitering behaviors through prolonged inactivity or irregular movement patterns.

This study approaches loitering detection as a classification problem based on trajectory images that represent an individual's movement pattern. The trajectory images are obtained by using YOLOv9 to extract the centroid positions of the detected individual in each frame. These centroid positions are then connected according to the order of occurrence to form the movement pattern, which is used to determine loitering. The proposed method combines YOLOv9 for real-time person detection and trajectory reconstruction with EfficientNet for trajectory image classification. It utilizes deep learning models to improve performance, in contrast to other methods that rely on hand-crafted features.