

BAB II

TINJAUAN PUSTAKA

Bab ini membahas berbagai teori, konsep, dan hasil penelitian terdahulu yang menjadi dasar dalam pengembangan sistem keamanan berbasis kamera yang menggunakan teknologi *Artificial Intelligence* (AI). Pembahasan difokuskan pada beberapa topik utama yaitu *Face Recognition*, *Object Detection* serta penelitian terdahulu yang relevan. Tinjauan Pustaka ini bertujuan untuk memberikan pemahaman yang komprehensif mengenai perkembangan teknologi yang mendukung penelitian ini, sekaligus menegaskan posisi dan kontribusi penelitian yang dilakukan dibandingkan dengan penelitian sebelumnya.

1. Sistem Keamanan Berbasis Kamera

Perkembangan teknologi Computer Vision dan Deep Learning telah mendorong sistem keamanan berbasis kamera menuju arah yang lebih cerdas. Sistem CCTV konvensional hanya berfungsi sebagai dokumentasi pasif, sehingga kurang efektif dalam memberikan peringatan dini. Penelitian terbaru mengenai deteksi anomali dalam video pengawasan menunjukkan bahwa sistem berbasis AI mampu mengidentifikasi perilaku abnormal, seperti loitering atau kerumunan yang tidak wajar, dengan akurasi yang semakin baik [2], [3]. Pendekatan berbasis deep learning juga menjadi fokus utama dalam deteksi aktivitas, baik melalui metode supervised maupun unsupervised, karena kemampuannya dalam memodelkan pola kompleks pada data video [4].

2. Face Recognition dalam Sistem Keamanan

Teknologi pengenalan wajah merupakan salah satu komponen utama dalam sistem keamanan modern, terdapat survei awal yang komprehensif mengenai face recognition yang menekankan pentingnya tahap deteksi, ekstraksi fitur, dan

pencocokan wajah [5]. Perkembangan signifikan terjadi dengan diperkenalkannya FaceNet, yang menggunakan pendekatan embedding untuk menghasilkan representasi wajah yang lebih diskriminatif dan akurat [6].

Dalam aplikasi nyata, banyak studi lainnya yang telah mengimplementasikan sistem *Face Recognition* pada sistem keamanan rumah, kontrol akses pintu (*smart door*), dan integrasi IoT (*Internet of Things*) di mana CNN digunakan pada Raspberry Pi dengan tingkat akurasi ~97,5% [7]. Studi-studi ini menunjukkan bahwa *Face Recognition* bukanlah hanya tentang teori, akan tetapi sudah ada uji praktik yang juga sudah diterapkan dalam skala perangkat keras nyata di sistem keamanan rumah [8].

Beberapa penelitian lainnya memadukan *Face Recognition* dengan pendekatan lain seperti *anomaly detection*, *hybrid learning*, atau sistem multi-modal untuk memperkuat keamanan. Contoh terkini adalah studi yang menerapkan kombinasi model CNN dengan deteksi anomali (model boosting) untuk mendeteksi aktivitas tak wajar sekaligus mengenali wajah [9]. Pendekatan tersebut memungkinkan sistem tidak hanya mengenali identitas, tetapi juga memantau aktivitas abnormal (misalnya penyusupan atau gerakan tak terduga).

3. Activity Detection dalam Sistem Keamanan

Selain identifikasi wajah, deteksi aktivitas manusia juga memainkan peran penting dalam sistem pengawasan. Aktivitas yang tidak wajar seperti orang yang berdiam terlalu lama di depan pintu (*loitering*) atau munculnya kerumunan mendadak yang bersifat mengkhawatirkan, dapat menjadi indikator potensi ancaman. Dengan perkembangan teknologi seperti menggunakan metode *deep learning* terbukti efektif untuk mendeteksi anomali di dalam rekaman video [4]. Selain itu, penggunaan metode OpenPose dan MediaPipe Pose mampu mengidentifikasi titik-titik sendi manusia secara real-time untuk mengenali pola gerakan seperti berjalan, duduk atau berlari [10], [11]. Di sisi lain, pengembangan model pengenalan aksi berbasis video

juga berkembang pesat di mana pada arsitektur Inflated 3D ConvNet (I3D) mampu menangkap informasi spasial dan temporal secara bersamaan. Dengan pendekatan tersebut, sistem keamanan dapat lebih akurat dalam mengenali perilaku kompleks pada lingkungan nyata [12].

Penelitian terkait *Activity Detection* juga banyak dilakukan untuk mendukung sistem keamanan dan pemantauan berbasis AI. Perkembangan metode *Activity Detection* berbasis *Ensembles of Deep Long Short-Term Memory* (LSTM) yang mampu meningkatkan ketepatan pengenalan aktivitas manusia menggunakan data dari sensor *wearable*. Model ini diuji pada 3 dataset standar dan mencapai rata-rata skor F1 hingga 0.85 yang menunjukkan peningkatan signifikan dibandingkan model LSTM tunggal [13]. Model *Transformer Network* untuk pengenalan aktivitas manusia berbasis data sensor inersia memanfaatkan mekanisme *self-attention* yang mampu mengekstraksi hubungan temporal jangka panjang antar-sekuens data dan menghasilkan akurasi hingga 96.1% pada dataset UCI HAR. Hasil penelitian ini menunjukkan bahwa arsitektur *transformer* memiliki potensi besar dalam mendeteksi aktivitas manusia secara efisien tanpa ketergantungan pada struktur rekuren seperti LSTM [14].

4. Perbandingan Penelitian Terdahulu

Beberapa penelitian terdahulu telah membahas komponen sistem keamanan berbasis *computer vision* seperti FaceNet untuk pengenalan wajah, ArcFace untuk peningkatan akurasi embedding, serta YOLO untuk deteksi objek secara real-time. Namun penelitian tersebut umumnya berfokus pada satu aspek tertentu, misalnya hanya *Face Recognition* atau hanya *Activity Detection* [6], [15], [16]. Berbeda dengan penelitian ini yang mengintegrasikan beberapa komponen sekaligus.

Tabel 2.1 di bawah menunjukkan hasil dan perbandingan penelitian ini dengan penelitian yang serupa.

Tabel 2. 1 Tabel Perbandingan Dengan Penelitian Serupa Secara Ilmiah

Peneliti	Metode	Model	Dataset	Input	Akurasi	Kelas
Schroff et al [6].	Face Recognition	FaceNet (CNN + Embedding)	Dataset wajah (CASIA-Webface)	Citra wajah (image dataset)	~99% (LFW Benchmark)	Wajah individu (multiclass)
Deng et al [15].	Face Recognition	ArcFace (Angular margin loss)	MS-Celeb-1M, LFW	Citra wajah (image dataset)	~99,8% (LFW Benchmark)	Wajah individu (multiclass)
Kiran et al [4].	Activity detection	Autoencoder + 3D-CNN	UCSD Pedestrian	Video Surveillance	~92% anomaly detection	Aktivitas abnormal
Bochkovsky et al [16]	Object detection	YOLOv4 (CSPDarkNet + FPN)	COCO dataset	Video / image stream	mAP 43.5% (COCO)	80 kelas objek umum
Penelitian Ini	Face recognition + object & activity detection + bot notification	ArcFace (InsightFace) + YOLOv8 + EfficientNet-B0 (head classifier)	Dataset wajah + dataset helm, masker, hoodie	Video streaming (IP Camera 720p, Webcam 640x480p)	Face recognition (80%), object + activity detection (90%)	Wajah (registered/unregistered), head classifier (4 kelas), activity (loitering, crowd)

Tabel 2. 2 Perbandingan Dengan Penelitian Serupa Berdasarkan Fitur

Peneliti	Face Recognition	Activity Detection	Bot Notification	Dashboard/Interface	Real-time Processing	Catatan
Schroff et al [6].	FaceNet	Tidak	Tidak	Tidak	Tidak	Fokus ke Face Recognition embedding
Kiran et al[4]	Tidak	Autoencoder + 3D-CNN	Tidak	Tidak	~ 15 FPS (offline)	Fokus ke video anomaly detection
Bochkovski et al [16]	Tidak	Person/Object detection	Tidak	Tidak	~ 30 FPS (GPU)	Real-time deteksi objek umum
Osman et al [17]	Tidak	Motion detection (IoT sensor)	Telegram Bot	Mobile app (basic)	Realtime Notification	Fokus integrasi IoT + Bot
Penelitian ini	ArcFace + InsightFace	Loitering + crowd detection	Telegram Bot	Streamlit	15-20 FPS	Integrasi multi-fitur: FR + AD + Bot + Dashboard

Berdasarkan tabel 2. 1 Perbandingan Penelitian Serupa Secara Ilmiah dan tabel 2. 2 Perbandingan Dengan Penelitian Serupa Berdasarkan Fitur, dapat disimpulkan bahwa penelitian terdahulu umumnya berfokus pada salah satu aspek tertentu dari sistem keamanan berbasis computer vision seperti penelitian face recognition lebih berfokus pada akurasi, activity detection yang lebih berfokus pada kinerja sistem untuk pengawasan real-time, serta penggunaan Telegram Bot untuk sistem notifikasi rumah pintar.

Berbeda dengan penelitian terdahulu, penelitian ini menggabungkan beberapa komponen sekaligus dalam satu sistem terpadu. Dari sisi model, penelitian ini memanfaatkan ArcFace melalui framework InsightFace untuk face recognition, YOLOv8 untuk activity detection (loitering and crowd), serta EfficientNet-B0 untuk

klasifikasi kondisi kepala (helm, masker, hoodie, bare face). Dari sisi evaluasi, sistem diuji tidak hanya pada akurasi face recognition (80%) dan activity detection (90%), tetapi juga dari segi performa real-time seperti frame rate (15-20 FPS) dan latensi pemrosesan. Dari sisi sistem, penelitian ini menambahkan mekanisme data logging berupa snapshot, video, dan log aktivitas dalam format CSV, serta integrasi Telegram Bot untuk notifikasi real-time dan Streamlit dashboard sebagai antarmuka monitoring.

Dengan integrasi multi-fitur ini, penelitian ini memberikan kontribusi yang masif dan adaptif. Sistem dibangun tidak hanya mampu mendeteksi wajah dan aktivitas, tetapi juga memberikan notifikasi instan, menyimpan rekaman sebagai bukti, dan menyediakan antarmuka yang mudah diakses pengguna. Hal ini menunjukkan perbedaan signifikan baik dari sisi pemilihan model, evaluasi sistem, dan implementasi integrasi fitur jika dibandingkan dengan penelitian lain yang umumnya berfokus pada satu aspek saja.