

Digital Twin

Literature Review
presented to the academic faculty
by

Nehal Naeem Haji	nh07884
Manal Hasan	mh08438
Eeshal Khalidnadeem Qureshi	eq08433
Muhammad Shawaiz Khan	mk07899



In partial fulfillment of the requirements for
Bachelor of Science
Computer Science

Dhanani School of Science and Engineering

Habib University
Fall 2025

1 Introduction

University campuses are dynamic environments where students and staff continuously move through classrooms, laboratories, and shared facilities. However, campus management lacks a unified, real-time understanding of how these spaces are being utilized, relying instead on manual counts, static schedules, or fragmented CCTV monitoring. Digital twin technology has recently emerged as a promising solution for improving operational awareness by creating virtual representations of spaces that update in real time based on sensor and vision data [1].

Recent research in multi-camera multi-people tracking has demonstrated that computer vision can accurately extract per-person positions, handle occlusions, and provide spatially aligned occupancy insights suitable for real-world deployments. Multi-camera fusion and coordinate mapping techniques have enabled reliable and privacy-aware tracking in complex indoor environments, supporting robust real time analytics even under challenging visual conditions [2]. In parallel, advances in digital twin based building automation demonstrate that integrating real time occupancy, environmental sensing, and predictive control can improve energy efficiency and maintain occupant comfort [3]. Together, these developments highlight the growing importance of unified data-driven facility management tools.

Therefore, this project presents an anonymity preserving digital twin for two adjacent university laboratories, integrating capabilities that existing systems traditionally address separately. The proposed system combines anonymous multi-camera crowd monitoring, automated emergency evacuation alerts to identify uncleared spaces, and energy-occupancy correlation using simulated IoT parameters. Through the use of non-identifiable avatars and interactive 3D visualization, the system aims to provide campus operators with a centralized, actionable platform for improving safety, reducing energy waste, and enabling proactive space management.

2 Current State-of-the-Art

2.1 Computer Vision for Anonymous Crowd Monitoring

2.1.1 Object Detection Models

Rane (2023) highlights that YOLO and Faster R-CNN are central to the evolution of real-time object detection technologies [4]. The author discusses key differences between these two widely used algorithms.

YOLO:

- YOLO is recognized for its high speed and strong accuracy. It divides the image

into a grid and simultaneously predicts bounding boxes and class probabilities for each grid cell [4].

- The recently introduced YOLOv8 by Ultralytics continues the YOLO lineage through an anchor-free design, a decoupled detection head, and a lightweight CSPDarknet-style backbone that supports detection, segmentation, and classification within a unified architecture [5].
- YOLOv8 maintains competitive accuracy while providing high throughput suitable for live CCTV feeds and real-time digital-twin applications.

Faster R-CNN:

- Faster R-CNN is another widely adopted object detection algorithm known for its accuracy and robustness. Unlike YOLO, it employs a two-stage architecture: a Region Proposal Network (RPN) to generate candidate bounding boxes, followed by a detection network to classify and refine the proposals.
- Ren et al. introduced the Region Proposal Network (RPN), which produces object proposals directly from convolutional feature maps, significantly improving efficiency over traditional selective search methods [6].

For digital-twin components of crowd monitoring, YOLO is preferable because of its speed and ability to track people and objects in real time, enhancing overall security monitoring [4].

2.1.2 Multi-Object Tracking (MOT) Techniques

Multi-Object Tracking (MOT) is essential for digital-twin indoor monitoring, ensuring consistent identity and trajectory estimation of individuals across video frames. Most MOT systems follow the 'tracking by detection' paradigm, pairing detectors such as YOLO with association algorithms to maintain persistent identities. This subsection summarizes three widely adopted MOT techniques.

SORT (Simple Online and Realtime Tracking): SORT is a fast, motion-based tracker that uses a Kalman Filter for prediction and the Hungarian Algorithm for frame-to-frame association. Its simplicity enables high speed, but the lack of appearance features makes it vulnerable to identity switches during occlusions or reentries [7].

DeepSORT: DeepSORT enhances SORT by incorporating appearance embeddings

through a deep association metric, enabling more reliable identity tracking in crowded or occluded scenes [8]. This addition significantly reduces ID switches while maintaining real-time performance.

ByteTrack: ByteTrack advances MOT by associating both high- and low-confidence detections, allowing it to recover missed objects and maintain identities more effectively [9]. Its robust data-association strategy yields strong benchmark performance and makes it suitable for complex, real-time digital-twin environments.

Overall, SORT provides speed, while DeepSORT and ByteTrack offer stronger robustness and identity consistency, key requirements for accurate multi-camera digital-twin systems.

2.1.3 Multi-Camera People Re-Identification (ReID)

Person Re-Identification (ReID) focuses on matching the same person across different camera views [10]. ReID is essential in digital-twin or multi-camera systems to avoid counting the same person multiple times or to stitch person trajectories across non-overlapping camera views.

Feature Learning: Modern ReID methods learn embeddings that cluster images of the same person close together while pushing different identities apart. Approaches commonly use triplet loss, cross-entropy ID loss, and hard-mining strategies. Benchmarks such as Market-1501 and DukeMTMC-ReID significantly advanced this domain by providing large multi-camera datasets [11].

Integration with Tracking and Trajectory Models: Practical multi-camera pipelines combine:

- detection (e.g., YOLOv8),
- single-camera tracking (creating tracklets),
- inter-camera ReID or trajectory association,

to resolve identity duplication. Systems integrating spatio-temporal priors, trajectory continuity, locality-aware metrics, and motion cues significantly reduce false duplicates [10].

Overlapping and Non-Overlapping Camera Views: In multi-camera networks

with mixed overlapping and non-overlapping fields of view, modeling spatial layout and appearance similarity helps enhance identity accuracy and reduce duplication errors [12].

In our project multiple cameras may detect the same individual simultaneously—especially in overlapping views, ReID is necessary to avoid duplicate counting. ReID enables accurate movement tracking across camera zones, improving the fidelity of the digital twin.

2.2 Translating CCTV Co-ordinates using Homography

Homography is a foundational technique in multi-camera computer vision for projecting image coordinates onto a shared planar reference frame [13].

2.2.1 Ground-Plane Homography for CCTV Monitoring

Ground-plane homography is widely used in surveillance and pedestrian analytics to project detections onto a 2D top-down map.

This paper [14] demonstrated application of this technique for multi-camera pedestrian localization, where detections from multiple cameras were projected onto a unified ground-plane mosaic to compute trajectories and occupancy maps [14].

By using known camera parameters or calibration patterns, each camera estimates its homography matrix, enabling all detections to be fused into a shared reference coordinate system. This process is essential for digital-twin systems, which require consistent spatial reasoning across multiple non-overlapping views [13].

2.2.2 Beyond Planar Homography: 3D Mapping and Bird’s-Eye-View (BEV) Transformations

BEV (bird’s-eye-view) transformations and 3D lifting methods estimate a pseudo-top-down representation of the scene from monocular or multi-camera setups.

Philion and Fidler introduced the Lift-Splat-Shoot framework, where image features are lifted into 3D space and then projected into a BEV representation for autonomous driving applications [15]. Such 3D-aware approaches offer richer spatial context compared to simple planar warping and are increasingly adopted in advanced monitoring and digital-twin pipelines.

2.3 Digital Twin Visualization

Digital twin systems depend on real-time 3D visualization to represent physical spaces, user interactions, and sensor-driven updates. Following are the engines widely used for 3D modeling.

2.3.1 Unity

Unity is one of the most widely used real-time engines for digital twin development due to its flexibility, fast rendering pipeline, and support for geospatial data integration. Unity enables interactive simulation, terrain modeling, and real-time visual updates, making it ideal for built-environment twins.

Gong and Ding (2025) demonstrate Unity’s ability to model complex urban environments through a real-time traffic digital twin, showcasing its capacity for scalable visualization and dynamic system behavior [16]. Similarly, Woo et al. (2025) propose an automated Unity-based pipeline that generates terrain and building geometries directly from open geospatial datasets, significantly reducing manual modeling effort for large-scale twins [17].

Lee et al. (2020) extend Unity’s capabilities to planetary-scale environments using VWorld geospatial data, illustrating how Unity can handle large-capacity datasets while supporting interactive 3D geospatial applications [18]. Their work emphasizes Unity’s strengths in rendering extensive terrain and structural information efficiently.

2.3.2 Unreal Engine

Unreal Engine is increasingly used in digital twin research due to its high-fidelity rendering, strong support for procedural content, and advanced simulation capabilities. Somanath et al.(2024) develop a procedural pipeline in Unreal Engine that transforms GIS and LiDAR datasets into detailed urban digital twins, capturing building footprints and volumetric spatial context [19].

In the construction domain, Ellul et al. (2024) build an Unreal Engine-based twin integrating BIM, IoT, and site monitoring data to support safety analysis and real-time progress visualization [20]. Lyu et al. (2024) describe the engineering workflow of a UE5-based twin platform, emphasizing Unreal’s advantages in large-scene rendering, physics simulation, material editing, and cinematic visualization [21]. Their comparison highlights Unreal’s suitability for graphically intensive and large-scale

twins.

Kilijanek and Miłosz (2025) provide a direct performance comparison between Unity and Unreal Engine, concluding that Unreal performs better in graphically complex scenes with many dynamic objects due to its advanced rendering tools, while Unity remains more resource-efficient for simpler projects with lower complexity [22].

2.3.3 Three.js

WebGL frameworks like **Three.js** enable browser-based 3D visualization for digital twins, offering real-time rendering without client-side installation. La Guardia et al. (2024) demonstrate a web-accessible urban digital twin integrating geospatial datasets and live sensor data. Their system, built using an HTML5/WebGL stack, demonstrates how libraries like Three.js can serve as an efficient front-end engine for visualizing heterogeneous 3D assets, including buildings, terrain models, and environmental datasets, directly in the browser [23].

A detailed example of Three.js integration in BIM-oriented digital twins is provided by the Carleton University **CDC Digital Twin** project [32]. In the project, Three.js was used to render campus BIM models in real time, integrating geometric data (pre-processed as glTF) and attribute data (exported as JSON) linked via element IDs.

Compared to game-engine solutions like Unreal Engine Pixel Streaming, Three.js offers key advantages:

- **Lightweight Deployment & Scalability:** Runs entirely in-browser with no installation or heavy server rendering, allowing potentially unlimited global users.
- **Flexible Data Integration:** Geometry and attributes are separated, supporting updates, expansions, and real-time sensor data integration.
- **Broad 3D Asset Support:** Handles glTF, OBJ, FBX, dense point clouds (LAS, E57), textured meshes, and IFC via IFC.js.
- **Open-Source and Customizable:** Fully open-source with extensive community support for customizing rendering pipelines, shaders, and interactions.

Limitations include lower visual fidelity compared to AAA game engines and the need for careful data subdivision for large models. Despite this, Three.js provides a robust foundation for web-based Digital Twins.

3 Similar Work

3.1 Multi-Camera Fusion and Bird-Eye View Location Mapping for Deep Learning-Based Cattle Behavior Monitoring

Nasir et al. propose a multi-viewpoint surveillance and behavior monitoring system for precision livestock farming, integrating deep learning and multi-camera fusion for real-time cattle behavior understanding [24].

Domain: Precision Livestock Farming (PLF), deep learning-based surveillance, and multi-camera behavior analytics.

Key Contribution: The authors present an end-to-end methodology for accurate behavior detection and location mapping of individual cattle using synchronized multi-camera footage. The system is deployed in real cattle barns and enables continuous monitoring for welfare analysis.

Problem Addressed: Traditional bovine surveillance systems struggle due to occlusions, pose changes, limited viewpoints, and difficulty distinguishing visually similar cattle. The proposed system mitigates these challenges using a fusion of detection, pose estimation, and BEV mapping.

System Components: The architecture uses four synchronized cameras, action detection, and pose estimation:

- **Action Detector:** Identifies 14 distinct cattle behaviors with a mean Average Precision (mAP) of 99.2%.
- **Ground Point Estimator:** Built on YOLOv8 Pose Estimation, it extracts the ground point for each cow, enabling precise 2D-to-3D location mapping.

BEV Mapping and Fusion: An Enhanced Bird’s-Eye-View (BEV) algorithm projects detections onto a unified top-down map using homography. Fusion strategies resolve duplicate projections and identity ambiguities by combining views from all cameras.

3.2 Crowded Event Management in Smart Cities Using a Digital Twin Approach

Villanueva et al. introduce a digital twin-driven platform for event management in smart cities, emphasizing situational awareness and resource planning during crowded events [25].

Domain: Smart-city systems, crowd event management, and augmented virtuality interfaces.

Key Contribution: The paper proposes a modular architecture and GUI design leveraging augmented virtuality to merge real-world sensor inputs with a 3D virtual city model. The system aids decision-makers during large events such as concerts or marathons.

Digital Twin Concept: The digital twin acts as a dynamically updated virtual replica of the physical urban environment. It supports simulation and forecasting using live data streams (IoT sensors, citizen apps, social media).

System Architecture: Data flows from the physical city layer (sensors, citizens, IoT devices) into the digital city layer containing 3D models, simulation engines, and virtual objects. Augmented virtuality overlays real-time observations into this environment for improved situational awareness.

3.3 Intelligent Real-Time Crowd Density Estimation for Proactive Event Safety

Maharajpet and Hegde propose a hybrid dual-model system to estimate crowd density in real time, combining YOLOv8 with CSRNet for high accuracy and responsiveness [26].

Domain: Real-time crowd monitoring, event safety, density estimation.

Key Contribution: The system strategically fuses the fast inference speed of YOLOv8 with the density map precision of CSRNet, addressing the trade-off between speed and accuracy faced by earlier systems.

System Architecture:

- **YOLOv8:** Provides real-time person detection and handles occlusion effectively.
- **CSRNet:** Generates high-resolution density maps for dense crowd conditions.

Density Classification and Interface: The monitored area is divided into zones and classified into four density levels: Low, Medium, High, and Critical. A web dashboard visualizes heatmaps (e.g., red for high density) and issues alerts when thresholds are exceeded.

3.4 Crowd Density Estimation and Real-Time Monitoring Using Digital Twins in Transportation Hubs

Shah et al. develop a digital twin-based platform for real-time monitoring in transportation hubs, emphasizing low-latency inference and 3D visualization [27].

Domain: Transportation hubs, digital twins for crowd monitoring, edge analytics.

Key Contribution: The system employs YOLOv8 for detection and uses Three.js for 3D visualization. Designed for edge computing, it minimizes latency and supports rapid response in real-time operations.

System Pipeline: Video frames are processed through YOLOv8 to identify individuals, compute bounding boxes, and calculate centroids.

3D Visualization: Detected coordinates are transformed into 3D space using inferred depth and camera intrinsics. The results are rendered as spheres on a virtual 3D plane using React Three Fiber, forming an interactive digital-twin display.

3.5 Interactive 3D and VR-Based Digital Twin for an Industrial Machine - FYP 2024

The project *Design and Development of an Interactive 3D Web-Based and Virtual-Reality Digital Twin for an Industrial Machine* [28] applies digital-twin concepts to an Industry 4.0 setting. The system features both a web-based 3D visualization and a virtual reality (VR) environment for a spice packaging machine.

Key features include:

- ingestion of real-time IoT sensor data (voltage, current, packet count) via bidirectional WebSocket communication,
- remote machine actuation (e.g., ON/OFF control),
- a 2D dashboard integrated with a 3D machine model,
- immersive VR for enhanced monitoring and operator training.

This demonstrates the utility of digital twins for remote monitoring, predictive maintenance, and industrial supervision.

3.6 GitHub Repositories

3.6.1 Crowd Management using a Digital Twin Approach

The repository *Crowd Management using a Digital Twin Approach* [29] proposes an end-to-end framework integrating CSRNet-based density estimation to compute crowd counts in highly congested environments such as the Mahakumbh Mela.

The system processes continuous video streams to generate density heatmaps, detect congestion spikes, and trigger automated emergency alerts via SMS and WhatsApp using Twilio. The digital-twin visualization primarily overlays 2D density maps onto the physical scene, providing an abstract yet informative view of spatial crowd pressure.

3.6.2 CCTV-Person-ReID-MultiCam

The repository *CCTV-Person-ReID-MultiCam* [30] presents a multi-camera person re-identification (ReID) pipeline built on the WildTrack dataset. The system integrates:

- YOLOv8m for person detection,
- DeepSORT for intra-camera tracking,
- custom feature extractors for ReID embeddings,
- and OpenCV-based camera calibration for geometric alignment.

This enables consistent identity association across non-overlapping CCTV views.

3.6.3 Multi-Camera Person Tracking and Re-Identification

The repository *Multi-Camera-Person-Tracking-and-Re-Identification* [31] proposes a practical implementation that uses YOLOv3/YOLOv4 for person detection and Torchreid models for feature embedding extraction. The pipeline outputs per-camera tracking trajectories and cross-camera identity matches using distance-based metric learning.

4 Difference between existing work and our own

While existing literature demonstrates various approaches to crowd monitoring, digital twin visualization, and multi-camera tracking systems, our Digital Twin project for Habib University’s laboratories introduces several distinctive features and design choices that differentiate it from prior work.

4.1 Identity Handling and Privacy Preservation

Nasir et al. [24] and ReID-based GitHub repositories such as CCTV-Person-ReID-MultiCam [30] and Multi-Camera-Person-Tracking-and-Re-Identification [31] perform persistent identity tracking, storing appearance embeddings to maintain identity continuity across cameras. Similarly, Maharajpet & Hegde [26] and Pattanaik [29] do not focus on privacy because their systems handle aggregated density data rather than individuals.

Our system instead prioritizes anonymity by:

- Using ReID only for duplicate removal,
- Representing each person as an anonymous avatar,

4.2 Domain & Functional Scope

Most of the reviewed digital-twin and monitoring systems are designed for single, domain-specific purposes. Nasir et al. focus on cattle behaviour monitoring in agriculture [24], Villanueva et al. address city-scale event management [25], while Maharajpet and Hegde target crowd density estimation [26]. Shah et al. develop solutions for transportation hub monitoring [27], industrial twins such as Mehmood and Razzaq emphasize machine-level diagnostics [28], and Pattanaik proposes a congestion-based emergency alerting system [29]. ReID repositories similarly concentrate on identity tracking within multi-camera networks. In contrast, our system

is a multi-functional digital twin built for indoor university spaces, integrating capabilities that prior works treat separately. Specifically, it combines (i) real-time occupancy monitoring with per-person mapping as an avatar, (ii) IoT-based energy-occupancy correlation for facility optimization, and (iii) emergency evacuation tracking with automated notifications to the Campus Security Office (CSO). This unified, multi-capability design makes our digital twin significantly broader and more operationally integrated than existing single-purpose solutions.

4.3 Visualization and 3D Digital Twin Technology

Visual outputs in existing systems are largely 2D or low-detail: Maharajpet and Hegde produce density heatmaps [26], Pattanaik overlays CSRNet maps [29], and ReID repositories show 2D bounding boxes. Nasir et al.’s BEV mapping [24] is also 2D in nature. While Shah et al. [27] also employ Three.js with React Three Fiber for 3D visualization in transportation hubs, our system extends this approach by rendering a spatially accurate 3D digital twin of laboratory environments with live avatar movements synchronized to detected person coordinates from multiple CCTV streams. Unlike the VR-based industrial digital twin [28], which requires dedicated VR hardware and focuses on machine monitoring with bidirectional IoT control, our system prioritizes accessibility through standard web browsers without requiring specialized equipment.

4.4 Detection & Tracking Methodology

Our technical pipeline combines YOLO 12 for real-time person detection with DeepSORT/ByteTrack for multi-frame tracking and duplicate mitigation across overlapping camera views. While Maharajpet and Hegde [26] propose a hybrid YOLOv8-CSRNet system for density estimation and CSRNet-based systems such as Pattanaik’s digital-twin implementation [29] perform well in highly congested scenes, they do not support precise individual-level tracking. In contrast, our system focuses on per-person coordinate tracking and accurate occupancy counting, enabling precise avatar placement within the 3D digital twin rather than generating only density heatmaps.

The ReID systems [30,31] rely on embedding-based identity maintenance, and Shah et al. [27] employ simple centroid tracking without multi-camera fusion. Nasir et al. [24] use multi-view pose estimation for identity consistency.

Our system integrates YOLO 12 for detection, DeepSORT/ByteTrack for short-term ID continuity, homography-based spatial alignment for multi-camera fusion, and tem-

porarily non-persistent IDs to enforce privacy. This combination is unique among the surveyed systems.

4.5 Emergency Evacuation and Safety Response

None of the existing works whether the smart-city framework of Villanueva et al. [25], the transportation twin of Shah et al. [27], the livestock system of Nasir et al. [24], or Pattanaik’s crowd-twin [29] provide room-level evacuation verification. Only Pattanaik includes SMS alerts, but these relate to density rather than evacuation.

Our system introduces the ability to declare emergencies, track un-evacuated rooms in real time, and automatically notify the CSOs via email. This emergency intelligence module is absent across the entire body of compared work.

4.6 Energy Occupancy Correlation

None of the reviewed systems analyzes building energy usage or correlates occupancy with AC or lighting inefficiencies. The VR Industrial Digital Twin [28] employs IoT data, but only for machine diagnostics rather than facility energy management. Our system fills this gap by detecting energy waste in unoccupied rooms, generating temperature threshold alerts, and offering an Energy Dashboard dedicated to the Facilities Manager, positioning energy analysis as a core component of the digital twin.

References

- [1] A. Sayed, M. Hassan, and M. Elhoseny, “Real-Time Digital Twin for Indoor Monitoring Using IoT and 3D Visualization,” *Applied Sciences*, vol. 15, no. 9, p. 4939, 2023. [Online]. Available: <https://www.mdpi.com/2076-3417/15/9/4939>
- [2] J. Kwon, S. Lee, and H. Kim, “Real-time Multi-person Tracking Using Sparse Camera Networks and Edge Computing,” 2023. [Online]. Available: <https://arxiv.org/abs/2305.05062>
- [3] A. Clausen, K. Arendt, A. Johansen, F. C. Sangogboye, M. B. Kjærgaard, C. T. Veje, and B. N. Jørgensen, “A digital twin framework for improving energy efficiency and occupant comfort in public and commercial buildings,” *Energy Informatics*, vol. 4, Suppl. 2, p. 40, 2021. [Online]. Avail-

able: <https://energyinformatics.springeropen.com/articles/10.1186/s42162-021-00153-9>

- [4] N. Rane, “YOLO and Faster R-CNN object detection for smart Industry 4.0 and Industry 5.0: Applications, challenges, and opportunities,” 2023. [Online]. Available: <http://dx.doi.org/10.2139/ssrn.4624206>
- [5] Ultralytics, “What is YOLOv8? An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector,” 2023. [Online]. Available: https://www.researchgate.net/publication/383494295_What_is_YOLOv8_An_In-Depth_Exploration_of_the_Internal_Features_of_the_Next-Generation_Object_Detector
- [6] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” 2015. [Online]. Available: https://www.researchgate.net/publication/277722488_Faster_R-CNN_Towards_Real-Time_Object_Detection_with_Region_Proposal_Networks
- [7] Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple Online and Realtime Tracking. *IEEE International Conference on Image Processing*, pp. 3464–3468. doi:10.1109/ICIP.2016.7533003. Available at: <https://arxiv.org/abs/1602.00763>
- [8] Wojke, N., Bewley, A., & Paulus, D. (2017). Simple Online and Realtime Tracking with a Deep Association Metric. *IEEE International Conference on Image Processing*, pp. 3645–3649. doi:10.1109/ICIP.2017.8296962. Available at: <https://arxiv.org/abs/1703.07402>
- [9] Zhang, Y., Sun, P., Jiang, Y., Yu, Q., Weng, X., Yuan, Z., Wang, L., Luo, P., Liu, W., & Wang, X. (2022). ByteTrack: Multi-Object Tracking by Associating Every Detection Box. *European Conference on Computer Vision (ECCV)*, LNCS 13682. doi:10.1007/978-3-031-20047-2_1. Available at: <https://arxiv.org/abs/2110.06864>
- [10] E. Ristani and C. Tomasi, “Features for Multi-Target Multi-Camera Tracking and Re-Identification,” in *CVPR*, 2018. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/papers/Ristani_Features_for_Multi-Target_CVPR_2018_paper.pdf
- [11] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian, “Scalable Person Re-Identification: A Benchmark,” in *ICCV*, 2015. [Online]. Available: <https://arxiv.org/abs/1506.02651>

//www.cv-foundation.org/openaccess/content_iccv_2015/papers/
Zheng_Scalable_Person_Re-Identification_ICCV_2015_paper.pdf

- [12] F. Author, “Multi-camera crowd analysis in overlapping and non-overlapping views,” 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0925231223006811>
- [13] S. Baker and I. Matthews, “A survey of planar homography estimation,” 2005. [Online]. Available: https://cdn.iiit.ac.in/cdn/cvit.iiit.ac.in/images/JournalPublications/2005/A_survey_of_planar_homography_estimation.pdf
- [14] R. Eshel and Y. Moses, “Homography-based multiple camera detection and tracking of people in a dense crowd,” in *Proc. CVPR*, 2008. [Online]. Available: <https://ieeexplore.ieee.org/document/4587539>
- [15] J. Philion and S. Fidler, “Lift, Splat, Shoot: Encoding Images from Arbitrary Camera Rigs by Implicitly Unprojecting to 3D,” 2020. [Online]. Available: https://www.researchgate.net/publication/343649048_Lift_Splat_Shoot_Encoding_Images_From_Arbitrary_Camera_Rigs_by_Implicitly_Unprojecting_to_3D
- [16] Gong, L., & Ding, M. (2025). Urban traffic digital twin system development in Unity. *Scientific Reports*, 15, 40085. <https://doi.org/10.1038/s41598-025-23943-7>
- [17] Woo, D., Choi, H., Espejo Jr., R. D., Kim, J., & Yu, S. (2025). Open-Data-Driven Unity Digital Twin Pipeline: Automatic Terrain and Building Generation with Unity-Native Evaluation. *Applied Sciences*, 15(21), 11801. <https://doi.org/10.3390/app152111801>
- [18] Lee, A., Chang, Y.-S., & Jang, I. (2020). Planetary-Scale Geospatial Open Platform Based on the Unity3D Environment. *Sensors*, 20(5967). <https://doi.org/10.3390/s20205967>
- [19] Somanath, S., Naserentin, V., Eleftheriou, O., Sjölie, D., Wästberg, B. S., & Logg, A. (2024). Towards Urban Digital Twins: A Workflow for Procedural Visualization Using Geospatial Data. *Remote Sensing*, 16(11), 1939. <https://doi.org/10.3390/rs16111939>

- [20] Ellul, C., Hamilton, N., Pieri, A., & Floros, G. (2024). Exploring Data for Construction Digital Twins: Building Health and Safety and Progress Monitoring Twins Using the Unreal Gaming Engine. *Buildings*, 14(2216). <https://doi.org/10.3390/buildings14072216>
- [21] Lyu, S., Wang, M., Zhang, S., & Wang, S. (2024). Design of a UE5-based digital twin platform. arXiv:2407.03107.
- [22] Kilijanek, R., & Milosz, M. (2025). Comparative analysis of the performance of Unity and Unreal Engine. *Journal of Computer Sciences Institute*, 35, 197–201. <https://doi.org/10.35784/jcsi.7298>
- [23] La Guardia, M., Koeva, M., Díaz-Vilariño, L., & Nourian, P. (2024). *Open-Source Solutions for Real-Time 3D Geospatial Web Integration*. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-4-2024, 289–295. <https://doi.org/10.5194/isprs-archives-XLVIII-4-2024-289-2024>
- [24] M. F. Nasir *et al.*, “Multi-camera fusion and bird-eye view location mapping for deep learning-based cattle behavior monitoring,” *Artificial Intelligence in Agriculture*, vol. 15, no. 4, pp. 724–743, 2025.
- [25] F. J. Villanueva *et al.*, “Crowded event management in smart cities using a digital twin approach,” in *Proc. IEEE ISC2*, 2022, pp. 1–7.
- [26] S. S. Maharajpet and A. Hegde, “Intelligent real-time crowd density estimation for proactive event safety: A machine learning approach,” *Recent Research Reviews Journal*, vol. 4, no. 2, pp. 264–280, 2025.
- [27] M. Shah, S. Goyal, and S. Salvi, “Crowd density estimation and real-time monitoring using digital twins in transportation hubs,” in *Proc. IEEE PuneCon*, 2024, pp. 1–6.
- [28] Habib University, “Design and Development of an Interactive 3D Web-Based and Virtual Reality Digital Twin for an Industrial Machine,” FYP Project, 2021. [Online]. Available: <https://hira.habib.edu.pk/jspui/handle/123456789/741>
- [29] A. Pattanaik and S. Pati, “Crowd management using a digital twin approach,” GitHub Repository, 2023. [Online]. Available: <https://github.com/Anjaneya-Pattanaik/Crowd-Management-using-a-Digital-Twin-Approach>

- [30] U. Purwar, “CCTV-Person-ReID-MultiCam,” GitHub Repository, 2024. [Online]. Available: <https://github.com/umangpurwar03/CCTV-Person-ReID-MultiCam>
- [31] S. Hormi, “Multi-Camera Person Tracking and Re-Identification,” GitHub Repository, 2022. [Online]. Available: <https://github.com/samihormi/Multi-Camera-Person-Tracking-and-Re-Identification>
- [32] Carleton Immersive Media Studio, *CIMS’ Canada’s Digital Twin project recognized for innovation, research*, Azrieli School of Architecture & Urbanism, Carleton University, December 10, 2024. Available at: <https://architecture.carleton.ca/2024/cims-canadas-digital-twin-project-recognized-for-innovation-research/>