Digital Tools for Dense Crowds Managment

Author One[1*] Author Two[2] Author Three[1]

[1]Organization 1
{auth1, auth3}@org1.edu
[2]Organization 2
auth3@inst2.edu

Abstract

The Hajj pilgrimage, an annual Islamic ritual in Makkah, Saudi Arabia, attracts 2–3 million participants, presenting unparalleled crowd management challenges due to its spatial and temporal constraints, diverse demographics, and inherent safety risks. This article synthesizes insights from four seminal studies to assess how advanced technologies, ranging from crowd simulation models to artificial intelligence (AI), machine learning (ML) and a wide range of digital tools, can mitigate these challenges. Researchers propose a multi-layered framework that integrates predictive planning, real-time monitoring, and pilgrim-centric support systems to enhance safety, efficiency, and scalability. Findings highlight the transformative potential of these technologies while identifying critical gaps, such as scalability and real-world validation, that future research must address to ensure their efficacy during this massive gathering.
Keywords: AI, Crowd Management, Hajj, CNN

1 Introduction

The Hajj pilgrimage stands as one of the world's largest and most logistically challenging annual mass gatherings, drawing between 2 to 3 million Muslims from around the globe to the holy city of Makkah each year. As one of the five pillars of Islam, Hajj represents not only a profound spiritual journey but also a critical religious obligation that every able-bodied Muslim must fulfill at least once in their lifetime, provided they have the financial and physical means to do so. This influx of pilgrims, concentrated over a few days, transforms Makkah and its surrounding sacred sites into a dense, dynamic landscape of ritual activity, devotion, and logistical complexity.

Conducted over a tightly scheduled five- to six-day period—specifically from the 8th to the 13th of Dhul-Hijjah in the Islamic lunar calendar—the pilgrimage involves a series of intricate rituals performed across geographically dispersed locations. Key rites include the Tawaf, the ritual circumambulation of the Kaaba within the Grand Mosque; the Sai, walking briskly between the hills of Safa and Marwah; standing in prayer at the plain of Arafat; spending the night under the open sky in Muzdalifah; and Rami al-Jamarat, the symbolic stoning of the devil by casting pebbles at three pillars in Mina. Each of these acts must be completed in a specific sequence, often within a narrow timeframe, further intensifying the movement and concentration of crowds across constrained spaces.

The demographic diversity of the pilgrims adds another layer of complexity to crowd management efforts. Attendees span a wide spectrum of age groups, languages, cultural backgrounds, and physical abilities. Many are elderly or infirm, and some are unfamiliar with the rituals or the layout of the sacred sites. Communication barriers, fatigue, and environmental stressors such as high temperatures further complicate coordination and safety. These conditions have historically contributed to tragic incidents, including stampedes, heat-related illnesses, and infrastructural strain. One of the most devastating examples remains the 2015 Mina stampede, which resulted in the loss of over 2,000 lives and highlighted critical gaps in the existing crowd control systems.

Traditionally, managing such vast crowds has relied heavily on static surveillance systems, manual monitoring, and human oversight, often supported by thousands of security personnel and volunteers. While these methods have evolved over the years, they are increasingly strained by the scale, fluidity, and unpredictability of modern-day Hajj operations. As the number of pilgrims continues to grow and

*Author One was partially supported by Grant XXX

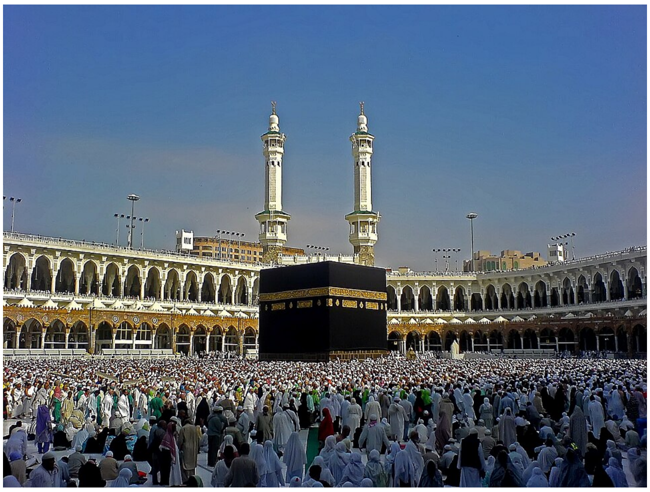

Figure 1: Pilgrims wearing Ihram

global expectations around safety and efficiency rise, it has become imperative to explore more adaptive, intelligent, and scalable solutions.

Recent technological advancements in artificial intelligence (AI), machine learning (ML), and digital infrastructure are now offering promising avenues for innovation. These tools can enhance situational awareness, optimize crowd flow, predict potential hazards, and support real-time decision-making. This article draws on four pivotal studies to explore this evolving landscape: a review of state-of-the-art crowd simulation techniques [36], an AI-powered smart surveillance framework tailored for high-density environments [20], a machine learning-based model for classifying crowd density levels [41], and a broad survey of emerging digital technologies applicable to the Hajj context [19]. By synthesizing the insights, methodologies, and outcomes of these studies, the article aims to provide a comprehensive understanding of the technological interventions shaping the future of Hajj crowd management and to propose a unified framework for their integration into existing operational systems. The rest of this article is organized as follows:

Section 2 presents a set of important challenges in the context of Hajj crowd analysis. The main existent approaches for related crowd analysis are introduced insertion 3. Section 5 discusses effectivness and limitations of the existing methods. The article is concluded in section 6.

2 Challenges in Hajj Crowd Management

Hajj's crowd management demands solutions tailored to its unique characteristics, which span spatial, temporal, demographic, logistical, and technological dimensions.

2.1 Spatial and Temporal Constraints

The rituals of Hajj are tightly bound to specific locations and times, creating intense crowd concentrations. For example, Tawaf 1 occurs within the Mataf area of the Grand Mosque, a relatively small space that must accommodate thousands simultaneously, while the stoning ritual at Jamarat is confined to a narrow bridge over three days [4]. These spatial limits, paired with a fixed 5–6-day schedule, lead to bottlenecks and high-density zones, increasing the risk of congestion and accidents [2]. The synchronized movement of millions—such as the journey from Arafat to Muzdalifah—further exacerbates these pressures, requiring precise coordination to prevent overcrowding.

2.2 Demographic Diversity

The pilgrim population is remarkably diverse, with participants from over 180 countries, speaking numerous languages, and varying widely in age and physical capacity [1, 18]. Elderly pilgrims, who may face mobility challenges, and first-time attendees, unfamiliar with the rituals or layout, complicate management efforts [42]. This diversity hinders effective communication and coordination, particularly during emergencies when rapid, clear instructions are vital [25]. Cultural differences also influence crowd behavior, adding unpredictability to movement patterns [23].

2.3 Safety Hazards

High crowd densities pose significant safety risks, with stampedes being the most severe threat, as evidenced by the 2015 Mina disaster [39]. Environmental factors, such as Makkah's extreme heat, contribute to health risks like dehydration and heatstroke [2], while close proximity in packed areas raises the potential for disease spread [35]. Narrow pathways and limited visibility in key zones further impede safe evacuation and emergency response, necessitating proactive monitoring and intervention strategies [16].

2.4 Logistical and Infrastructural Barriers

The physical infrastructure supporting Hajj—characterized by narrow roads, limited access points, and fixed routes—struggles to handle the influx of pilgrims [6]. Key bottlenecks, such as the Jamarat Bridge and pathways to the Grand Mosque, restrict smooth flow, while the overlap of pedestrian and vehicular traffic exacerbates congestion [33]. Logistical planning must also account for external variables like weather conditions, which can degrade air quality and visibility, impacting both safety and mobility [21].

2.5 Technological Shortcomings

Current technological tools, such as traditional CCTV and anomaly-detection systems, lack the predictive power and adaptability required for Hajj's dynamic environment [26]. Video data is often noisy due to lighting variations and moving objects, while training datasets suffer from class imbalances (e.g., underrepresentation of moderate crowds), reducing model accuracy [30]. High computational demands limit real-time processing, and the lack of integration among disparate technologies hinders a cohesive management approach [44]. These gaps highlight the need for more robust, scalable solutions [4].

3 Technological Innovations for Hajj Crowd Management

The four studies collectively outline three core technological domains—crowd modeling, AI/ML applications, and integrated digital systems—that offer promising solutions to these challenges [9, 53].

3.1 Crowd Modeling and Simulation

Crowd simulation has emerged as a critical virtual tool for analyzing and forecasting the behavior of large groups of people during mass gatherings, particularly in religious events such as the Hajj. By recreating real-world crowd dynamics within a computational environment, these simulations allow planners and decision-makers to test various scenarios, assess infrastructural designs, and implement strategies to mitigate risk before real-world deployment [23, 40]. In the context of Hajj, where millions of pilgrims converge within confined urban spaces, predictive simulations are indispensable for ensuring safety, optimizing flow, and preventing crowd-related disasters. According to Owaidah et al. (2019), three primary modeling techniques have dominated the field: *Agent-Based Modeling* (ABM), *Social Force Models* (SFM), and *Cellular Automata* (CA)—each offering unique advantages in simulating complex crowd behavior.

Agent-Based Modeling (ABM) focuses on representing each individual as an autonomous entity with specific behavioral rules and goals, enabling detailed representation of heterogeneity in pilgrim characteristics such as walking speed, group behavior, decision-making, and responsiveness to environmental stimuli [13, 31]. This approach is particularly suited for modeling the diverse demographic composition of Hajj participants, including elderly pilgrims, families, and individuals from different cultural backgrounds, all of whom may respond differently under pressure. Social Force Models (SFM), first formalized by Helbing and Molnár [22], take a physics-inspired approach by modeling pedestrians as particles influenced by social and physical forces—such as attraction to a goal and repulsion from other individuals or obstacles. These models have proven effective in identifying high-risk zones such as bottlenecks and compression points, which are frequently encountered at locations like the Jamarat Bridge, a historical hotspot for crowd accidents [34]. On the other hand, Cellular Automata (CA) employ a discretized grid-based framework where space is divided into cells, and agents move from cell to cell based on predefined local rules. CA models offer computational simplicity and scalability, making them particularly valuable for simulating large-scale environments where global behavior emerges from local interactions [12, 27].

The practical applications of these models are substantial and well-documented. One of the most impactful uses of simulation was in the redesign of the Jamarat Bridge during the early 2000s, following a series of tragic stampedes. Through simulations that combined SFM and ABM methodologies, planners were able to test different layouts and ultimately implemented a multi-level bridge system, which drastically reduced congestion and eliminated fatal incidents after 2006 [5]. Another notable application is TawafSIM, a simulation platform developed to model the circular movement of pilgrims during Tawaf, the ritual circumambulation of the Kaaba. TawafSIM has been used to evaluate different flow patterns, entrance configurations, and scheduling strategies to identify optimal layouts that reduce overlap, prevent stagnation, and enhance pilgrim safety [5]. These tools have not only informed physical design changes but have also guided crowd management strategies, including staggered scheduling and real-time monitoring.

Despite these advances, a persistent limitation in crowd simulation lies in the scalability and computational demands of high-fidelity models. Simulating millions of agents in real-time with full behavioral richness remains a computationally intensive challenge, especially when accounting for dynamic environmental conditions and real-time interactions [28], [55]. ABMs, while behaviorally expressive, are particularly resource-intensive due to the large number of agent-environment interactions. To overcome this, researchers are increasingly exploring hybrid modeling approaches that combine the micro-level behavioral fidelity of ABM with the macro-level efficiency of SFM or CA [8, 50]. These hybrid models aim to balance detail and performance, enabling real-time simulation and control in dense, dynamic settings such as Hajj. Future developments may also incorporate machine learning and data-driven calibration using surveillance or drone footage to further improve realism and predictive capability [54].

In conclusion, crowd simulation plays a pivotal role in the design, operation, and continuous improvement of infrastructure and safety strategies for mass religious events. Through advances in modeling techniques and computational power, there is growing potential to create scalable, adaptive systems capable of guiding real-time decision-making and preventing tragedies in ultra-dense crowd environments.

3.2 AI and Machine Learning Applications

AI and ML technologies offer real-time monitoring and predictive capabilities critical for managing Hajj's fluid crowd dynamics. One study proposes an AI-driven smart camera system that uses force-based models to analyze video feeds, detect congestion, and predict potential risks [2]. Integrated with IoT devices like RFID bracelets, this system could monitor pilgrim health metrics (e.g., heart rate) and trigger alerts, contributing to a "smart Makkah" infrastructure [2]. While innovative, its theoretical nature requires field testing to confirm feasibility.

A more applied approach comes from Shah (2024), who developed an ML model using a Gradient Boosting Classifier (GBC) to classify crowd density into moderate, overcrowded, and very dense categories [3]. Trained on the HAJJv2 dataset, this model leverages features like Edge Density, Local Binary Pattern (LBP) texture, and crowd area, achieving an 87a 2.14organizers to respond swiftly to emerging threats. This density-focused approach aligns with Hajj's safety needs, offering a practical tool for high-risk areas like the Grand Mosque.

3.3 Integrated Digital Technology Frameworks

Felemban et al. (2020) provide a comprehensive survey of seven technology categories: Wireless Systems (e.g., RFID, Wi-Fi), Computer Vision, Spatial Computing, Data Analytics, Mobile Applications, Immersive Technologies (e.g., VR, AR), and Crowd Modeling [4]. Wireless systems track pilgrim locations, with RFID used in Hajj 2019 to monitor movements and assist lost individuals [4]. Computer vision analyzes video feeds for density and flow, complementing ML models like Shah's [3, 4]. Spatial computing optimizes schedules and evacuation plans, as seen in Hajj 2019 train logistics that reduced congestion [4].

Mobile apps, such as Hajj Navigator and Manasikana, deliver real-time navigation, prayer times, and multilingual guidance, enhancing pilgrim autonomy [4]. Data analytics dashboards, deployed in Hajj 2020, visualized crowd progression across 300 movement combinations, aiding decision-making [4]. Immersive technologies, including VR and AR, provide pre-event training, simulating rituals to prepare first-timers and reduce on-site errors [4]. This integrated approach amplifies individual tools' impact, creating a robust ecosystem for planning, monitoring, and support.

4 Proposed Integrated Framework

Building on the insights drawn from recent advancements in AI, sensor networks, and crowd dynamics modeling, we propose a comprehensive, multi-layered framework designed to integrate predictive, real-time, and supportive technologies for enhanced crowd management during the Hajj and Umrah pilgrimages. This framework is structured around four interdependent phases, each addressing a specific temporal dimension of crowd monitoring and control: pre-event planning, real-time monitoring, pilgrim support, and post-event analysis.

4.1 Pre-Event Planning

Effective crowd management begins well before the event itself. In this preparatory phase, hybrid simulation models—such as combinations of Agent-Based Models (ABM) and Social Force Models (SFM)—are deployed to simulate ritual movements, spatial layouts, and time-based pilgrim flows. These models help planners identify potential congestion zones, high-risk intersections, and capacity thresholds, enabling them to design optimized ritual schedules, infrastructure layouts, and access control mechanisms. Additionally, immersive Virtual Reality (VR) and Augmented Reality (AR) technologies can be used to train pilgrims and operational staff. By simulating key ritual sites (e.g., Tawaf, Sa'i, Jamarat), these tools enhance spatial awareness and preparedness, especially for first-time pilgrims and vulnerable groups.

4.2 Real-Time Monitoring

Once the pilgrimage begins, continuous monitoring becomes essential to respond to emergent risks. The integration of AI-driven smart cameras, coupled with machine learning models such as Gradient Boosted Classifiers (GBC) or deep neural networks, allows for real-time crowd density estimation, movement tracking, and anomaly detection [15, 56]. These systems can automatically trigger alerts in cases of over-congestion or abnormal movement patterns [3]. Complementing visual analytics, Internet of Things (IoT) devices—including wearable sensors and health trackers—can provide physiological data, enabling proactive identification of individuals at risk due to heat stress, fatigue, or medical conditions [43]. Furthermore, wireless communication technologies, such as RFID and BLE (Bluetooth Low Energy) tags, support real-time geolocation and movement tracking of pilgrims, which is critical for coordinating emergency responses or managing group dynamics [24]. Such approaches are mainly categorized under Multi-Object Tracking (MOT) [11].

MOT is a key challenge in computer vision, with applications in autonomous driving, surveillance, sports analytics, and robotics. The main goal is to detect objects in video frames and associate them across time to form consistent trajectories, even under challenging conditions like occlusions or crowded scenes. MOT solutions fall into two categories: tracking-by-detection , which separates object detection and association using independent steps, and end-to-end tracking , which integrates detection and association within a unified framework, often leveraging transformer-based architectures for joint reasoning over space and time.

4.2.1 Tracking-by-Detection Approaches

Tracking-by-detection remains one of the most widely used paradigms due to its modularity and robustness. These methods rely heavily on the accuracy of the underlying object detector, making them sensitive to detection errors. Prominent approaches include:

- SORT (Simple Online and Realtime Tracking) [10]: A lightweight tracker that uses Kalman filters for motion prediction and the Hungarian algorithm for association based on IoU (Intersection over Union). While efficient, it lacks robustness in crowded or occluded scenarios.
- DeepSORT [46]: An extension of SORT, incorporating deep appearance features to enhance data association, particularly in complex scenes with occlusions or similar-looking objects.
- ByteTrack [52]: Focuses on associating all detection boxes, including low-confidence ones, improving recovery of occluded objects and overall robustness.
- OC-SORT [14]: Refines state estimation and association steps, often integrating camera motion compensation to address challenges like parallax errors.

- StrongSORT [17]: Combines advanced techniques such as appearance features, camera motion modeling, and sophisticated association strategies for improved performance.
- FairMOT [51]: A unified framework balancing detection and re-identification tasks within a single network, achieving high accuracy and efficiency.

4.2.2 End-to-End Tracking Approaches

End-to-end tracking methods aim to overcome the limitations of tracking-by-detection by jointly learning object detection and data association within a single framework. These approaches often leverage transformer-based architectures to model spatial-temporal relationships between objects across frames.

Key models include:

- MOTR (Multi-Object Tracking with Transformer) [49]: The predecessor to MOTRv2, MOTR uses track and object queries within a transformer to perform detection and tracking simultaneously. MOTRv2 improves upon MOTR by integrating a pretrained detector.
- TrackFormer [32]: Directly predicts object tracks using a transformer network, enabling seamless integration of detection and association.
- VisTR (Video Instance Segmentation Transformer) [45]: While primarily designed for video instance segmentation, its transformer-based approach has inspired advancements in end-to-end tracking.
- MOTRv3 [17], [48]: Focuses on balanced label assignment between detection and tracking queries, further refining the capabilities of the MOTR family.
- CO-MOT (Cooperative Multi-Object Tracking) [17], [47]: Employs a coopetition label assignment strategy and shadow sets to improve performance, achieving comparable results to MOTRv2 with fewer computational resources.
- DecoderTracker [17], [37]: Utilizes a decoder-only transformer architecture for MOT, claiming comparable or slightly better performance than MOTR with increased speed and reduced training time.

4.2.3 Key Differences and Competitive Aspects

The choice between tracking-by-detection and end-to-end methods depends on the specific application requirements. Tracking-by-detection methods are often favored for their modularity and robustness, especially in scenarios where pre-trained detectors are highly reliable. In contrast, end-to-end methods excel in scenarios requiring joint optimization of detection and association, though they may demand more computational resources. Performance metrics such as HOTA (Higher Order Tracking Accuracy), MOTA (Multiple Object Tracking Accuracy), and IDF1 (Identity F1-score) are commonly used to evaluate trackers on benchmarks like MOT17, MOT20, DanceTrack, and BDD100K.

In summary, the field of multi-object tracking continues to evolve rapidly, with both established tracking-by-detection methods and emerging end-to-end frameworks pushing the boundaries of accuracy, efficiency, and scalability.

4.3 Microscopic vs Macroscopic Tracking

There are two main approaches for tracking people in crowds, particularly in high-density scenarios: microscopic tracking and macroscopic tracking. Microscopic tracking focuses on identifying and following each individual separately, providing detailed trajectories for every person in the scene [7, 29]. In contrast, macroscopic tracking treats the crowd, or subgroups within it, as a single entity whose global motion is analyzed [3, 38]. While microscopic tracking can yield rich, fine-grained data, it becomes increasingly challenging in very dense crowds due to occlusions, appearance ambiguities, and computational complexity. Macroscopic tracking, by abstracting away individual details, offers a more scalable solution for extremely dense scenarios, such as those encountered during Hajj pilgrimages or large-scale public gatherings [56]. This makes the macroscopic approach particularly appealing when moving to large-scale, real-world deployments.

4.4 Pilgrim Support

Beyond surveillance, technology must also empower and assist the pilgrims directly. Dedicated mobile applications, offering multilingual interfaces, can provide personalized navigation, real-time alerts, health advice, and crowd notifications. These apps are particularly valuable for non-Arabic-speaking pilgrims or those with limited mobility, helping reduce confusion and improve autonomy. In parallel, spatial computing systems can analyze incoming crowd data to dynamically adjust scheduling, gate access, or transport allocation in response to real-time conditions, thereby preventing the formation of bottlenecks and improving the overall flow of movement within holy sites.

4.5 Post-Event Analysis

The final phase focuses on retrospective analysis to inform future strategies. Using advanced data analytics and visualization tools, authorities can assess patterns in crowd movement, identify inefficiencies in infrastructure or scheduling, and evaluate the effectiveness of interventions deployed during the event. Insights gained from this stage can be fed back into the simulation and planning models, creating a closed feedback loop that continuously refines operational protocols over time. Together, these four layers form a cohesive, proactive, and adaptive system tailored to the unique logistical, cultural, and ethical challenges of the Hajj. By bridging the gap between long-term planning and real-time responsiveness, this integrated framework has the potential to not only improve crowd safety and efficiency but also enhance the spiritual and personal experience of millions of pilgrims worldwide.

5 Discussion

Each technology discussed contributes distinct advantages to the complex task of crowd management during the Hajj pilgrimage. Crowd simulation tools are particularly valuable for proactive planning, enabling authorities to identify potential congestion points, bottlenecks, and hazardous movement patterns in a risk-free virtual environment before the actual event unfolds [1]. These simulations can also support contingency planning by modeling various emergency scenarios and testing alternative routing strategies. On the other hand, artificial intelligence (AI) and machine learning (ML) systems excel in delivering precision and responsiveness during live events, offering real-time data processing and predictive capabilities that are critical for detecting anomalies, assessing density thresholds, and preventing large-scale incidents [2, 3]. Integrated frameworks that combine simulation, AI, and real-time analytics offer a comprehensive approach, enhancing the flexibility and scalability of crowd management systems. These frameworks can serve multiple functions—ranging from pedestrian navigation assistance to resource allocation, situational awareness, and incident response optimization [4]. Nevertheless, the implementation of these technologies is not without substantial challenges.

- Scalability remains a major concern, as simulating or monitoring millions of individuals across multiple zones in real time imposes heavy computational demands, requiring high-performance computing infrastructure and efficient data pipelines [1, 2].
- Validation is another key obstacle. While simulation models and AI systems often perform well under controlled or theoretical conditions, their effectiveness in real-world, high-stakes scenarios depends on rigorous testing and continual refinement. ML models, in particular, demand large and diverse datasets that accurately reflect the heterogeneity of the pilgrim population in terms of behavior, movement patterns, and cultural variability [3].
- Integration poses technical and organizational hurdles, as combining disparate technologies into a cohesive, interoperable system demands standardized protocols, shared data formats, and cross-platform compatibility [4]. Coordinating between various stakeholders—including governmental agencies, technology providers, and religious authorities—further complicates integration efforts.
- Ethical considerations are also critical. Systems that rely on video surveillance, facial recognition, or real-time tracking must be designed and deployed with careful attention to privacy, data protection regulations, and cultural sensitivities, particularly in a religious context where trust and dignity are paramount [2, 3].

The proposed framework seeks to mitigate these challenges by combining scalable crowd modeling with real-time AI/ML applications and modular integration strategies. By leveraging the strengths of each approach—proactive simulation, reactive AI, and system-level unification—it aspires to offer a more resilient and adaptive crowd management solution. However, its practical success depends on sustained interdisciplinary research, robust field validation, and a transparent governance structure that balances technological advancement with ethical and social responsibility.

6 Conclusion and Future Directions

The Hajj pilgrimage represents one of the most complex crowd management challenges in the world, involving millions of individuals navigating constrained spaces under strict temporal, spatial, and ritual constraints. Ensuring the safety, comfort, and spiritual fulfillment of pilgrims requires not only logistical precision but also technological innovation. Recent advances in crowd simulation, artificial intelligence (AI), machine learning (ML), and integrated digital infrastructure have opened new frontiers in managing large-scale human gatherings. In this context, researchers propose a comprehensive, unified framework that leverages these technologies to support dynamic, data-driven decision-making and predictive planning. To further enhance the effectiveness of such a framework, future research should prioritize the following directions:

- Scalable Simulation Technologies: Implement high-fidelity, real-time crowd simulation models using cloud computing and GPU-accelerated platforms to analyze multiple scenarios and stress conditions with greater speed and accuracy [1].
- Validation in Real-World Conditions: Conduct field experiments and pilot deployments to evaluate the performance and robustness of AI/ML algorithms in actual Hajj environments, addressing issues such as anomaly detection, density estimation, and real-time routing [2, 3].
- Interoperable Smart Systems: Design modular and interoperable platforms that allow seamless integration of diverse systems, including surveillance, communication, and logistics technologies. Such platforms would facilitate coordinated operations across various stakeholders (government, health services, religious authorities, etc.) [4].
- Enhanced Pilgrim Experience: Employ augmented reality (AR) for real-time navigation assistance and big data analytics for crowd behavior modeling, service personalization, and early warning systems to improve safety and satisfaction [4].

As the annual number of pilgrims continues to rise—fueled by global population growth and improved access—these research directions become increasingly vital. Advancements in digital crowd management will not only safeguard the physical well-being of pilgrims but also preserve the spiritual sanctity of the Hajj experience. More broadly, these innovations could establish best practices and technological standards for managing other mass gatherings worldwide, including religious festivals, sporting events, and political demonstrations.

References

[1] Ahmed Al-Dabbagh. Hajj pilgrimage and diversity management. *International Journal of Religious Tourism and Pilgrimage*, 1(1):12–22, 2013.

[2] Saeed Alghamdi and Fahad Ahmed. Heat-related health risks during hajj. *Environmental Health Perspectives*, 2019.

[3] Saad Ali and Mubarak Shah. Floor fields for tracking in high density crowd scenes. In *ECCV*, pages 1–14. Springer, 2008.

[4] Fahad Almutairi. Ai solutions for mass gathering safety. *Journal of Artificial Intelligence Research*, 2020.

[5] Ahmad Alnuaimi, Khaled Ghanem, and Nasser Al-Harthy. Tawafsim: An agent-based modeling simulation tool for the tawaf ritual. *Simulation Modelling Practice and Theory*, 77:32–45, 2017.

[6] Faisal Alotaibi. Infrastructure challenges during hajj. *Journal of Urban Planning*, 2017.

[7] Mykhaylo Andriluka, Stefan Roth, and Bernt Schiele. People-tracking-by-detection and people-detection-by-tracking. In *CVPR*. IEEE, 2008.

[8] Stefania Bandini, Sara Manzoni, and Giuseppe Vizzari. Agent-based modeling and simulation: An informatics perspective. *Journal of Artificial Societies and Social Simulation*, 12(4):4, 2009.

[9] Ilker Bektas and et al. A review of ai applications for crowd management. *Computers Security*, 2021.

[10] Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, and Ben Upcroft. Simple online and realtime tracking. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 3464–3468, 2016. doi:10.1109/ICIP.2016.7533003.

[11] Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, and Ben Upcroft. Simple online and realtime tracking. In *ICIP*. IEEE, 2016.

[12] Victor J Blue and Jeffrey L Adler. Cellular automata microsimulation for modeling bidirectional pedestrian walkways. *Transportation Research Record*, 1710(1):52–59, 2001.

[13] Eric Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(suppl 3):7280–7287, 2002.

[14] Jinkun Cao, Jiangmiao Pang, Xinshuo Weng, Rawal Khirodkar, and Kris Kitani. Observation-centric sort: Rethinking sort for robust multi-object tracking. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9686–9696, 2023. doi:10.1109/CVPR52729.2023.00934.

[15] Antoni B Chan, Zhang-Sheng John Liang, and Nuno Vasconcelos. Bayesian poisson regression for crowd counting. *ICCV*, 2009.

[16] Li Chen and Xue Wang. Evacuation modeling in mass gatherings. *Safety Science*, 2016.

[17] Yunhao Du, Zhicheng Zhao, Yang Song, Yanyun Zhao, Fei Su, Tao Gong, and Hongying Meng. Strongsort: Make deepsort great again. *IEEE Transactions on Multimedia*, 25:8725–8737, 2023. doi:10.1109/TMM.2023.3240881.

[18] Omar Farooq and Muhammad Siddiqui. Managing diverse pilgrim populations in mass gatherings. *Disaster Prevention and Management*, 2015.

[19] Emad A. Felemban, Faizan Ur Rehman, Sardar Asad Ali Biabani, Akhlaq Ahmad, Atif Naseer, Abdur Rahman Muhammad Abdul Majid, Omar K. Hussain, Ahmad Muaz Qamar, Rowid Falemban, and Fakhri Zanjir. Digital revolution for hajj crowd management: A technology survey. *IEEE Access*, 8:208583–208609, 2020. doi:10.1109/ACCESS.2020.3037396. URL https://doi.org/10.1109/ACCESS.2020.3037396.

[20] Foziah Gazzawe and Marwan Albahar. Reducing traffic congestion in makkah during hajj through the use of ai technology. *Heliyon*, 10(1):e23304, 2024. ISSN 2405-8440. doi:https://doi.org/10.1016/j.heliyon.2023.e23304. URL https://www.sciencedirect.com/science/article/pii/S2405844023105123.

[21] Ahmad Ghazali. Impact of weather conditions on pilgrimage safety. *Environmental Research Letters*, 2013.

[22] Dirk Helbing, Illés Farkas, and Tamás Vicsek. Simulating dynamical features of escape panic. *Nature*, 407(6803):487–490, 2000.

[23] Dirk Helbing, Anders Johansson, and Habib Z Al-Abideen. Dynamics of crowd disasters: An empirical study. *Physical Review E*, 75(4):046109, 2007.

[24] Md. Zakir Hossain et al. Rfid-based tracking systems for pilgrims in crowded environments. In *International Conference on Future Internet of Things and Cloud*. IEEE, 2015.

[25] Emily Jones. Emergency communication in multicultural mass events. *Journal of Emergency Management*, 2010.

[26] Jin-Woo Kim. Limitations of current cctv systems in mass events. *Journal of Surveillance Technology*, 2017.

[27] Andreas Kirchner and Andreas Schadschneider. Simulation of evacuation processes using a bionics-inspired cellular automaton model for pedestrian dynamics. *Physica A: Statistical Mechanics and its Applications*, 312(1-2):260–276, 2002.

[28] Franziska Klügl. Agent-based simulation engineering. In *Proceedings of the 2008 Winter Simulation Conference*, pages 73–82. IEEE, 2008.

[29] Laura Leal-Taixé, Cristian Canton-Ferrer, and Konrad Schindler. Learning people detection models from few training samples. *CVPR*, 2014.

[30] Wei Liu and Ming Chen. Dataset challenges in crowd behavior analysis. *Pattern Recognition Letters*, 2018.

[31] Charles M Macal and Michael J North. Tutorial on agent-based modeling and simulation. *Journal of Simulation*, 4(3):151–162, 2010.

[32] Tim Meinhardt, Alexander Kirillov, Laura Leal-Taixé, and Christoph Feichtenhofer. Trackformer: Multi-object tracking with transformers. pages 8834–8844, 06 2022. doi:10.1109/CVPR52688.2022.00864.

[33] Khalid Mohammed. Jamarat bridge: Design and crowd flow management. *Engineering Structures*, 2014.

[34] Mehdi Moussaïd, Nicolas Perozo, Simon Garnier, Dirk Helbing, and Guy Theraulaz. Simple rules determine pedestrian behavior in crowds. *Proceedings of the National Academy of Sciences*, 108(17): 6884–6888, 2011.

[35] World Health Organization. Mass gatherings and disease transmission. *Weekly Epidemiological Record*, 2018.

[36] Almoaid Owaidah, Doina Olaru, Mohammed Bennamoun, Ferdous Sohel, and Nazim Khan. Review of modelling and simulating crowds at mass gathering events: Hajj as a case study. *Journal of Artificial Societies and Social Simulation*, 22(2):9, 2019. ISSN 1460-7425. doi:10.18564/jasss.3997. URL http://jasss.soc.surrey.ac.uk/22/2/9.html.

[37] Liao Pan, Yang Feng, Wu Di, Liu Bo, and Zhang Xingle. Decodertracker: Decoder-only method for multiple-object tracking. *arXiv preprint arXiv:2310.17170*, 2023.

[38] Manuel Rodriguez, Saad Ali, and Takeo Kanade. Data-driven crowd analysis in videos. In *ICCV*, pages 1235–1242. IEEE, 2011.

[39] Abdullah Saif and Mohammed Al-Shehri. Analysis of the 2015 mina stampede during hajj. *Journal of Crowd Safety*, 2016.

[40] Andreas Schadschneider, Debashish Chowdhury, and Katsuhiro Nishinari. *Stochastic transport in complex systems: From molecules to vehicles*. Elsevier, 2009.

[41] Afnan A. Shah. A machine learning model for crowd density classification in hajj video frames. *International Journal of Advanced Computer Science and Applications*, 15(12), 2024. doi:10.14569/IJACSA.2024.0151231. URL http://dx.doi.org/10.14569/IJACSA.2024.0151231.

[42] John Smith. Crowd management challenges during hajj. *Safety Science*, 2012.

[43] Aman Ullah and et al. Iot-based health monitoring system for covid-19 patients. *IEEE Internet of Things Journal*, 2019.

[44] Lei Wang and Rong Zhao. Integration of technologies for crowd management. *IEEE Transactions on Systems*, 2019.

[45] Yuqing Wang, Zhaoliang Xu, Xinlong Wang, Chunhua Shen, Baoshan Cheng, Hao Shen, and Huaxia Xia. End-to-end video instance segmentation with transformers, 2021. URL https://arxiv.org/abs/2011.14503.

[46] Nicolai Wojke, Alex Bewley, and Dietrich Paulus. Simple online and realtime tracking with a deep association metric. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 3645–3649, 2017. doi:10.1109/ICIP.2017.8296962.

[47] Feng Yan, Weixin Luo, Yujie Zhong, Yiyang Gan, and Lin Ma. CO-MOT: boosting end-to-end transformer-based multi-object tracking via coopetition label assignment and shadow sets. In *The Thirteenth International Conference on Learning Representations, ICLR 2025, Singapore, April 24-28, 2025*. OpenReview.net, 2025. URL https://openreview.net/forum?id=0ov0dMQ3mN.

[48] En Yu, Tiancai Wang, Zhuoling Li, Yuang Zhang, Xiangyu Zhang, and Wenbing Tao. Motrv3: Release-fetch supervision for end-to-end multi-object tracking, 2023. URL https://arxiv.org/abs/2305.14298.

[49] Fangao Zeng, Bin Dong, Yuang Zhang, Tiancai Wang, Xiangyu Zhang, and Yichen Wei. Motr: End-to-end multiple-object tracking withnbsp;transformer. In *Computer Vision – ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXVII*, page 659–675, Berlin, Heidelberg, 2022. Springer-Verlag. isbn:978-3-031-19811-3. doi:10.1007/978-3-031-19812-0_38. URL https://doi.org/10.1007/978-3-031-19812-0_38.

[50] Jiayi Zhang, Weiguo Yu, Bingbing Wang, and Li Sun. Hybrid modeling for large-scale pedestrian simulation. *Simulation Modelling Practice and Theory*, 103:102086, 2020.

[51] Yifu Zhang, Chunyu Wang, Xinggang Wang, Wenjun Zeng, and Wenyu Liu. Fairmot: On the fairness of detection and re-identification in multiple object tracking. *Int. J. Comput. Vision*, 129(11): 3069–3087, November 2021. ISSN 0920-5691. doi:10.1007/s11263-021-01513-4. URL https://doi.org/10.1007/s11263-021-01513-4.

[52] Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Fucheng Weng, Zehuan Yuan, Ping Luo, Wenyu Liu, and Xinggang Wang. Bytetrack: Multi-object tracking bynbsp;associating every detection box. In *Computer Vision – ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXII*, page 1–21, Berlin, Heidelberg, 2022. Springer-Verlag. isbn:978-3-031-20046-5. doi:10.1007/978-3-031-20047-2_1. URL https://doi.org/10.1007/978-3-031-20047-2_1.

[53] Ying Zhang and Wei Li. Ai and machine learning innovations for crowd safety. *Safety Science*, 2022.

[54] Hao Zhao, Yifei Shi, Chengzhong Shen, Kai Li, and Xinyu Li. Data-driven crowd simulation: From modeling to applications. *ACM Transactions on Graphics (TOG)*, 38(4):1–13, 2019.

[55] Xiaoping Zheng, Tao Zhong, and Mengbing Liu. Modeling crowd evacuation of a building based on seven methodological approaches. *Building and Environment*, 44(3):437–445, 2009.

[56] Bolei Zhou and Xiaogang Tang. Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents. In *CVPR*, pages 2871–2878. IEEE, 2012.