IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 10 March 2025, accepted 23 March 2025, date of publication 26 March 2025, date of current version 8 April 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3555154

APPLIED RESEARCH

Toward an Integrated Intelligent Framework for Crowd Control and Management (IICCM)

TARIK ALAFIF [1], MOHAMMAD JASSAS [1], ALAA E. ABDEL-HAKIM [1,2], GHADA ALFATTNI [1], HASSAN ALTHOBAITI [1], MOHAMMED IKRAM [1], AMIRAH ALHARBI [3], HUSSAM ALSHARIF [1], MAZIN ALSHAMRANI [4], EBTISAM ALHARBI [3], TAHANI ALSUBAIT [5], (Senior Member, IEEE), ABDULLAH ALHAWSAWI [4], BADR ALSOLAMI [6], AND KHALID KHAYYAT [7]

[1]Department of Computer Science, Jamoum University College, Umm Al-Qura University, Al Jumum 25371, Saudi Arabia
[2]Electrical Engineering Department, Assiut University, Asyut 71516, Egypt
[3]Department of Data Science, Umm Al-Qura University, Makkah 24382, Saudi Arabia
[4]Hajj and Umrah Research Institute, Umm Al-Qura University, Makkah 24382, Saudi Arabia
[5]Department of Computer Science and Artificial Intelligence, Umm Al-Qura University, Makkah 24382, Saudi Arabia
[6]Department of Islamic Architecture, Umm Al-Qura University, Makkah 24382, Saudi Arabia
[7]Department of Computer and Network Engineering, Umm Al-Qura University, Makkah 24382, Saudi Arabia

Corresponding author: Tarik Alafif (tkafif@uqu.edu.sa)

ABSTRACT Managing large-scale gatherings, such as global festivals, sporting events, and religious congregations, presents substantial challenges in ensuring crowd safety and control. Innovative frameworks are essential to address these complexities effectively. The Integrated Intelligent Crowd Control and Management (IICCM) framework combines cutting-edge technologies, including Computer Vision (CV), Artificial Intelligence (AI), and the Internet of Things (IoT), to enhance participant safety and optimize crowd management. CV enables precise real time identification and tracking, AI analyzes crowd behavior to anticipate risks, and IoT gathers environmental data to improve crowd flow, alleviate congestion, and provide timely assistance. Additionally, the framework facilitates emergency evacuation planning by modeling crowd dynamics and identifying safe, efficient escape routes. Although suitable for diverse events, the Hajj pilgrimage—a uniquely large and dynamic annual gathering—provides a rigorous test case for the IICCM framework. Managing millions of participants from varied cultural and linguistic backgrounds highlights the system's adaptability and robustness. By effectively addressing Hajj specific challenges, the IICCM framework demonstrates its scalability and applicability to other large-scale events. This research offers valuable insights for decision-makers seeking to implement advanced crowd management technologies.

INDEX TERMS Crowd management, large-scale gatherings, artificial intelligence, computer vision, Internet of Things, emergency evacuation, crowd health, Hajj.

I. INTRODUCTION

Hajj, the annual Islamic pilgrimage to Makkah, is one of the largest religious gatherings in the world, attracting Muslims from all countries and continents. It is a mandatory religious duty for every Muslim to perform at least once in their lifetime. Each year, nearly three million pilgrims simultaneously converge in Makkah to perform the Hajj, creating massive, diverse crowds with significant cultural variation. Figure 1 illustrates a frame from a large-scale Hajj crowd video stream captured by a closed circuit television camera (CCTV). Managing such large crowds requires an integrated intelligent framework to ensure pilgrim safety, security, and assistance while mitigating potential risks, including injuries or fatalities caused by overcrowding and uncontrolled crowd flows.

The size and dynamics of crowds vary by context, with small-scale crowds involving tens of people and large-scale crowds encompassing hundreds, thousands, or even hundreds of thousands. Large-scale crowds pose significant challenges

The associate editor coordinating the review of this manuscript and approving it for publication was Sotirios Goudos.

FIGURE 1. A sample frame from a CCTV video stream capturing a large-scale crowd scene during Hajj.

in identifying individuals and monitoring their health during crowd movements. In addition, ensuring the sustainability of crowd data, video streams, and their associated applications is crucial. These challenges make analyzing and managing large-scale crowds significantly more complex than similar tasks in smaller crowds. This research project tackles these challenges by developing key components and advanced methodologies for analyzing large-scale crowds in the context of Hajj.

Artificial Intelligence (AI), a branch of computer science, focuses on simulating human intelligence in machines that are designed to think, learn, and act autonomously [1]. The primary goal of AI is to develop intelligent agents capable of reasoning, learning, and solving real-world problems. As an overarching field, AI encompasses optimization, search algorithms, machine learning, and deep learning techniques. In this research project, we suggest cutting-edge machine learning and deep learning algorithms to address control challenges in areas such as crowd vision, crowd health monitoring using the Internet of Things (IoT), and crowd evacuation during large-scale Hajj events.

Crowd vision, a specialized branch of computer vision, aims to understand and analyze crowd dynamics [2]. This field faces significant challenges due to the dense, dynamic, and occluded nature of crowded environments. Despite these obstacles, recent advancements in crowd vision have facilitated the development of systems suitable for real-world applications [3], [4], [5]. Among the primary challenges are effectively monitoring individuals within a crowd and accurately tracking their movements, both of which are complicated by occlusions caused by people or objects. Deep learning and multi-camera systems have emerged as transformative solutions, enhancing performance in these tasks [6], [7], [8]. As this technology continues to advance, crowd vision is expected to play a pivotal role in diverse applications, delivering increasingly accurate and efficient solutions.

The IoT refers to a network of physical devices embedded with sensors, software, and technologies that enable them to connect and exchange data over the internet [9]. These devices range from everyday household items to advanced industrial tools. Although IoT is still in its developmental stages, it holds the potential to transform numerous aspects of daily life. By integrating IoT devices with cloud computing, significant benefits are realized, including scalability and enhanced performance for IoT-enabled applications [10], [11], [12]. This synergy has driven the creation of innovative solutions across various domains, such as healthcare [13], transportation [14], and energy [15]. However, deploying efficient IoT platforms remains a challenge, particularly for modern cloud-based applications such as smart cities, home automation, and crowd e-health, which demand scalable and reliable framework capable of storage, computing, and data analysis. This research project harnesses state-of-the-art IoT technologies and advanced AI algorithms to enhance crowd vision, health monitoring, and evacuation strategies, ensuring public safety and security in large-scale gatherings.

Cloud computing faces challenges in meeting the Quality of Service (QoS) requirements for IoT applications, such as low latency and high availability, due to its remote infrastructure and the vast number of connected IoT devices [16]. To overcome these limitations, we are designing a cloud-IoT framework that incorporates proactive fault-tolerance techniques to improve the reliability and availability of IoT applications, particularly for crowd-related Software as a service (SaaS) solutions that are used in the cloud over the internet [17]. This framework also aims to optimize critical performance metrics, including scalability, throughput, and response time, ensuring more efficient and dependable application performance.

Crowd health, or "crowd e-Health," is an emerging field that harnesses collective knowledge, experience, and resources to enhance healthcare delivery and outcomes [18]. By tapping into the power of large groups, crowd e-Health aims to create more effective and efficient strategies for managing health and wellness. One major application of crowd e-Health is in patient decision-making [19]. Online forums and social media platforms connect patients, caregivers, and healthcare professionals, providing a rich source of information and support. This empowers patients to make informed decisions about their care and take greater control of their health.

Additionally, crowd e-Health supports clinical research by aggregating data from large populations, enabling researchers to uncover patterns and trends not visible in smaller studies as shown in [20] and [21]. These insights can drive the development of new treatments, diagnostic tools, and a deeper understanding of diseases. Moreover, crowd e-Health fosters innovation in healthcare tools and technologies, such as crowd-sourced medical transcription services that can reduce workloads and enhance record accuracy, and AI systems that assist with diagnosis, treatment planning, and risk assessment. Through these tasks, the crowd e-Health can transform healthcare into a more collaborative and data-driven field.

While crowd e-Health holds great promise for improving healthcare delivery and outcomes, it also faces several challenges and limitations. Ensuring the quality and accuracy of information shared online is a significant concern, as misinformation can lead to poor decision-making [22], [23]. Protecting patient privacy and securing sensitive health data is another critical issue that must be addressed [24], [25]. Furthermore, it is essential to ensure that crowd e-Health tools and technologies are accessible to all individuals, regardless of income level or technical expertise. Despite these hurdles, crowd e-Health has the potential to revolutionize healthcare. By leveraging the collective knowledge, experiences, and resources of large groups, it offers innovative and efficient approaches to managing health and wellness on a large scale.

Developing an intelligent e-Health framework requires effective analysis of health-related data to support real-time decision-making. This involves creating a framework that integrates data from diverse sources, such as Electronic Health Records (EHRs) and wearable devices, to detect patterns and trends. These insights can be used to predict patient outcomes and recommend timely interventions for individuals in crowded environments. The framework is designed to provide real-time alerts to clinicians, individuals, and decision-makers in such settings, enabling them to take proactive measures to prevent adverse events and enhance patient care, particularly for pilgrims in mass gatherings.

Crowd evacuation, a critical component of crowd management, involves guiding large groups of people from a hazardous area to a safe location [26]. This process is inherently complex, requiring consideration of various factors such as crowd size and density, the nature of the hazard, available evacuation routes, and the psychological and physical condition of evacuees. Several challenges complicate evacuations, including the difficulty of controlling large crowds, the potential for panic in emergencies, congestion on evacuation routes, and the vulnerabilities of individuals with specific needs, such as the elderly or those with disabilities. Effective planning and management are essential to ensure the safety of evacuees and minimize casualties. Key principles for successful crowd evacuation include early detection of hazards, timely emergency warnings, and real-time path recommendations to facilitate fast and secure movement to safety.

The primary contribution of this article is the design of an Integrated Intelligent Crowd Control and Management (IICCM) framework, which leverages advanced technologies such as Computer Vision (CV), AI, and IoT to improve participant safety and optimize crowd management, as illustrated in Figure 2. Furthermore, a comprehensive analysis of related studies is conducted based on this framework.

The proposed framework is inspired by advancements in cutting-edge technologies within CV, AI, and the IoT. Key AI technologies integrated into the framework include:

- Generative Adversarial Networks (GANs) [27]: GANs are machine learning models capable of generating realistic images and videos [28], [29], [30], [31], [32]. These can be leveraged to create synthetic training data for crowd management tasks or to simulate crowd behavior, enabling testing and refinement of the framework.
- Vision Transformers (ViTs) [33]: ViTs, known for their superior performance in image classification compared to traditional Convolutional Neural Networks (CNNs) [34], [35], [36], [37], [38], can enhance the framework's accuracy in identifying and tracking individuals within crowds based on attention mechanism.
- Zero-shot Learning [39]: This technique enables models to generalize and perform tasks without explicit training, allowing the framework to adapt to unforeseen scenarios dynamically.
- Few-shot Learning [40]: Few-shot learning minimizes the amount of training data required for new tasks, streamlining the development process for the intelligent crowd control and management framework.
- Contrastive Learning [41]: By learning to differentiate between positive and negative pairs of examples, contrastive learning can help the framework identify anomalous or suspicious crowd behaviors effectively.
- Continual Learning [42]: This approach allows the framework to learn and adapt to new tasks over time without forgetting previously acquired knowledge, ensuring adaptability to real-world changes while maintaining performance.

The remainder of this paper is structured as follows: Section II provides a review of the background and related work, focusing on crowd vision, crowd health and IoT, crowd evacuation, and high availability in IoT-based crowd applications. Section III details the preliminary proposed framework, including its tasks and methodologies. Section IV outlines the future implementation aspects. Section V offers an in-depth discussion of the framework. Finally, Section VI concludes the paper and highlights directions for future work.

II. BACKGROUND AND RELATED WORK

Several crowd research studies have been presented over the past two decades. In this work, we focus on recent advancements in the domains of crowd vision, crowd health, and crowd evacuation. These advancements are categorized into four themes aligned with the core components of our framework, as detailed below.

A. THEME 1. CROWD VISION

Recent studies have significantly advanced people counting and crowd density estimation using vision and deep learning methods [43]. Ahuja and Charniya [44] provide a comprehensive survey on modern image-processing-based crowd density estimation techniques, summarizing key developments in the field. Ding et al. [45] introduced an encoder-decoder Convolutional Neural Network (CNN) that fuses feature maps across encoding and decoding layers, enhancing density map accuracy. Their Patch Absolute Error (PAE) metric set a new standard for evaluating density estimation performance.

Chen et al. [46] developed a Multi-Scale and Multi-Column Convolutional Neural Network (MSMC), addressing varying object sizes to improve robustness and accuracy. Saleem et al. [47] presented an ensemble regression-based model focusing on memory efficiency and computational speed, achieving high performance in crowd estimation. Zhu et al. [48] tackled challenges such as occlusion and varying density levels by proposing a Patch Scale Discriminant Regression Network (PSDR) and a person Classification Activation Map (CAM), delivering robust estimation results. Additionally, the study in [49] aimed to improve facial detection performance in the wild, particularly under occlusions and varying poses.

Face recognition has also played a critical role in crowd management. For instance, locating missing persons was explored in [50], utilizing a four-step process including spatio-temporal search boundary estimation and face detection. Another security-focused application [51] employed Local Binary Patterns Histogram (LBPH) for surveillance, enhancing public safety via drone-captured imagery. Other works, such as [52], [53], and [54], utilize robust principal component analysis for face recognition. Other work in [55] attempted to classify races from faces using the CNN. Additionally, Facial Expression Recognition (FER) has been used to analyze crowd mood [56], providing insights into emotional states like happiness or anger, enabling tailored services and recommendations.

People counting remains a cornerstone of crowd estimation. Frameworks like the one proposed in [43] utilize multi-sensor video data and advanced feature extraction techniques to track and count individuals dynamically. Similarly, [57] employed multi-sensor detectors and robust feature extraction to estimate foot traffic, demonstrating high precision. In addition, Alotaibi et al. [58] made a performance comparison for large-scale crowd counting using CNNs.

Monitoring and managing crowds for safety and efficiency have also been extensively studied. Authors in [59] and [59] integrated edge computing with Unmanned Aerial Vehicles (UAVs) and ground sensors to develop a multi-modal crowd monitoring system. Their hybrid model leverages behavioral sensing and classification for enhanced crowd safety. Wang et al. [60] introduced a lightweight CNN-based model optimized for edge intelligence, providing real-time density estimation to prevent stampedes.

Abnormal behavior detection in large crowds has been addressed using hybrid models. Alafif et al. [61] combined CNNs with Random Forests (RFs) for behavior detection, achieving superior Area Under the Curve (AUC) metrics. Additionally, [62] leveraged optical flow and Generative GANs for abnormal behavior detection in both small and large-scale crowd scenarios, offering valuable solutions for managing high-density gatherings such as Hajj.

B. THEME 2. CROWD HEALTH AND IOT

Health-related information holds immense value and can play a critical role in real-time decision-making processes. Analyzing this vast data is essential for enhancing healthcare services and aligning with Saudi Arabia's digital transformation initiatives in the health sector. Big data analytics enables the discovery of hidden patterns and trends in health-related data that are challenging to identify manually. These insights can be leveraged to predict patient outcomes and recommend targeted interventions [63].

For instance, big data analytics can help identify patients at risk of developing specific diseases or those who may experience adverse reactions to certain medications under particular conditions, including crowded environments. This proactive approach can significantly improve patient care and resource allocation, ensuring better health outcomes.

Real-time decision-making is crucial for delivering high-quality healthcare, enabling clinicians to act swiftly and accurately, even in complex cases. Data analysis plays a vital role in this process by equipping clinicians with the necessary insights to assess a patient's condition and determine the most effective course of treatment [64].

This framework, specifically the crowd health and IoT component, leverages big data analytics to process health-related information and deliver real-time decision support to clinicians and patients. It enables real-time alerts for clinicians, individuals, and decision-makers, facilitating timely interventions to prevent adverse events and enhance patient care. The component integrates data from various sources, including:

- EHRs: EHRs provide a comprehensive repository of patients' medical history, encompassing diagnoses, prescribed medications, laboratory results, and other critical health information.
- Wearable Devices: Devices like smartwatches and fitness trackers gather valuable data on patients' vital signs, activity levels, and sleep patterns, offering real-time insights into their health status. The component leverages various machine learning algorithms to analyze the collected data, uncover patterns and trends, and make actionable predictions. It can forecast patient outcomes and recommend appropriate interventions. For instance, the component might predict that a patient is at risk of heatstroke and suggest consulting a doctor for further evaluation. Consequently, this component provides real-time decision support to clinicians and patients through multiple channels, such as:
    1) Alerts: The component generates and delivers notifications to clinicians and patients when it detects potential issues, such as a patient at risk of developing a specific condition or experiencing an adverse reaction to a particular medication.
    2) Recommendations: The component offers personalized suggestions to clinicians and patients based on its analysis of health data. These recommendations may include lifestyle modifications or initiating specific medications to improve health outcomes.
    3) Dashboards: The component delivers interactive dashboards to clinicians and patients, offering

visualizations of health data along with actionable insights into their overall health status.

The proposed framework offers several benefits to both clinicians and patients, including the following:

- Enhanced decision-making: The framework supports clinicians in making more informed and timely decisions by providing essential information to assess the patient's condition and recommend the most effective treatment options.
- Cost reduction: The framework helps lower healthcare costs by preventing adverse events and identifying patients at risk of developing chronic diseases.
- Enhanced patient outcomes: The framework improves patient outcomes by equipping patients with the information needed to make informed health decisions and by recommending interventions that help prevent diseases and manage chronic conditions.

The proposed framework can be implemented in four key phases:

- Phase 1: Data Collection.
  In this phase, data are gathered from various sources, such as EHRs, wearable devices, and social media. Then, they are stored in appropriate storage systems.
- Phase 2: Data Integration.
  This phase focuses on ensuring the consistency and compatibility of the data collected from multiple sources. It involves transforming and loading the data into a centralized data warehouse, preparing it for analysis and enabling efficient utilization in the following phases.
  In both Phase 1 and Phase 2, the Extract, Transform, Load (ETL) process is applied to aggregate the data and create a unified data warehouse. This ensures that the collected health-related data is standardized and prepared consistently, facilitating seamless analysis.
- Phase 3: Model Development and Evaluation.
  In this phase, various machine learning models are developed and assessed to identify patterns and trends within the data. The most effective models are then selected for deployment in production.
- Phase 4: Component Deployment and Evaluation.
  In this phase, the component is deployed in a production environment, and its performance is evaluated. Feedback are gathered from clinicians and patients, and the component is assessed using various metrics, including accuracy, precision, recall, and F1 score. These metrics measure the component's ability to identify relevant information and deliver personalized recommendations. Additionally, a team of medical experts can evaluate the component to ensure it provides accurate, reliable information and maintains consistency in decision-making.

Smartwatches can monitor and measure vital signs of individuals in a crowd, potentially saving lives through active alerts for critical health symptoms such as elevated heart rates. These alerts can be automatically linked to emergency services to ensure timely intervention. Additionally, smartwatches can detect critical events like falls, further enhancing individual safety. Advanced features available in premium models include Electrocardiogram (ECG) monitoring and Blood Oxygen Level (SpO2) measurement. The ECG functionality helps detect signs of atrial fibrillation, while SpO2 monitoring provides valuable insights into respiratory health, making these devices highly beneficial for both personal health management and large-scale crowd health monitoring.

The ECG is a medical test that evaluates heart performance by measuring the electrical activity generated during the heart's contraction and relaxation phases. Modern devices and smartwatches equipped with medical alert systems often include fall detection algorithms. These systems can automatically detect falls by analyzing collected signals and trigger an alert without requiring any user action. This feature can be enabled automatically or disabled manually, offering flexibility to users.

Smartwatches are essential tools for health monitoring and stress detection, especially in scenarios requiring individual or group behavioural studies to understand health conditions. They enable continuous monitoring and provide insights into factors affecting the well-being of both individuals and groups. Various smartwatch models, with features tailored to specific monitoring needs, can be selected for this purpose based on their functionalities and capabilities, as outlined in Table 1.

C. THEME 3. CROWD EVACUATION

Crowd evacuation is a complex process involving multiple human actions, such as evacuation movement and behavioral responses [65]. Tragically, incidents involving large crowds result in an average annual death toll exceeding 2,000 people. This underscores the critical importance of developing not only effective but also safe evacuation plans to handle emergencies and ensure crowd safety. Over the years, numerous research studies have focused on designing efficient and secure evacuation strategies [66].

According to Chertkoff et al. [67], the primary cause of deaths in such events is not the initial catastrophe itself, but the collective human reaction to it. Actions such as stampeding, pushing, knocking others down, and trampling are often the leading causes of death [65]. Real-time indoor crowd evacuation has emerged as a significant research area aimed at preventing accidents and injuries. Recently, several crowd evacuation methods have been thoroughly reviewed in [26].

To refine the scope of the literature review, only studies utilizing smartphone devices for guided crowd evacuation were considered. For instance, Ikeda and Inoue [68] proposed an evacuation route planning method leveraging a multi-objective genetic algorithm. This approach employs Global Positioning System (GPS) and a cloud server to collect data and display a safety map of evacuation routes on users' smartphones. Similarly, Chittaro and Nadalutti [69] developed a mobile application featuring a 3D visualized location model to facilitate visual crowd evacuation.

Inoue et al. [70] introduced a system designed for indoor crowd evacuation, which connects to smartphones to determine evacuees' positions and provide real-time indoor navigation. Chen et al. [71] presented a crowd evacuation framework aimed at minimizing evacuation time. In studies [72] and [73], radio frequency techniques were applied via smartphones to compute evacuation routes and locate evacuees during fire-related emergencies. Lastly, Iizuka and Iizuka [74] utilized a combination of multi-agent systems and smartphones to evaluate evacuees' locations and estimate evacuation times.

D. THEME 4. HIGH AVAILABILITY IN IOT-BASED CROWD APPLICATIONS

Cloud computing and the IoT are among the most transformative technologies of our era, reshaping how businesses function. Enterprises increasingly migrate workloads to public clouds and adopt multi-cloud strategies to enhance cost-efficiency, agility, and flexibility. The rising demand for cloud computing has driven providers to establish massive data centers globally, designed to store, manage, and analyze user data via the Internet, eliminating the reliance on local machines. These cloud providers ensure the maintenance and management of data centers, delivering high Quality of Services (QoS) to users who access resources like virtual machines and cloud storage through APIs that facilitate communication between applications and data sources.

Designing and implementing highly available crowd applications requires minimizing task failures through effective failure analysis and prediction. Unsupervised machine learning techniques have been employed to classify cloud applications based on job and task events. For example, Alam et al. [75] utilized K-means clustering to analyze Google workload patterns and categorize jobs. Gao et al. [76] improved the accuracy of failure prediction with a multilayer Bidirectional Long Short Term Memory (Bi-LSTM) algorithm, capable of identifying task and job failures based on task completion. Although the Bi-LSTM algorithm delivers superior accuracy, its training process is slow and computationally intensive. Jassas and Mahmoud [77] developed a model to predict failed tasks in cloud applications by evaluating multiple algorithms and selecting the most precise approach.

The IoT is revolutionizing sectors such as healthcare, transportation, and beyond. By 2025, the number of connected devices and machines is projected to reach 38.6 billion, with an anticipated economic impact ranging from $3.9 trillion to $11.1 trillion [78]. Advances in edge and cloud computing are unlocking new research possibilities by merging local IoT resources with modern computational capabilities. For example, smart cities are pivotal in driving innovation across industries, leveraging the integration of multiple urban systems such as transportation, healthcare, and operations research to enhance crowds' quality of life [79].

The convergence of IoT and cloud computing, driven by advancements in data connectivity, enables more devices to seamlessly connect to the cloud, forming systems with enhanced capabilities to handle crowd flows. This progress highlights the growing need for frameworks that effectively combine cloud and IoT technologies, particularly for crowd applications that demand scalability, high availability, and superior performance.

A review of related works reveals that tasks such as crowd vision, crowd health, crowd evacuation, and high availability in IoT-based crowd applications are typically addressed in isolation. To the best of our knowledge, this is the first framework to integrate these tasks into a cohesive, intelligent control and management crowd system. The proposed framework achieves this integration, offering robust crowd monitoring and management services to support security authorities in their operations.

III. PROPOSED FRAMEWORK, TASKS, AND METHODOLOGIES

Figure 2 illustrates the proposed framework, which is composed of four components: (1) Crowd Vision, (2) Crowd e-Health and IoT, (3) Crowd Evacuation, and (4) Availability and Reliability of Cloud-IoT Large-Scale Crowd Applications. Each component includes several tasks, all of which are integrated into a single user interface as the final product.

The first component is Crowd Vision, which encompasses several tasks designed to assist security authorities in monitoring crowds intelligently using video streams from surveillance cameras. This component includes tasks such as large-scale crowd counting and area measurement, detection and tracking of crowd direction and speed, crowd density estimation, location identification of individuals, identification of missing or suspicious persons, detection, tracking, and recognition of abnormal crowd and individual behaviors, early crowd prediction, and detection of crowd demographics such as race, sex, and age. These tasks are further detailed in Section III-A.

The second component is Crowd e-Health and IoT, which includes several tasks designed to intelligently manage pilgrim health information. Health records are obtained from the EHRs. Additionally, this component features capabilities that assist with tracking directions and detecting accidents involving pilgrims. Geolocation or GPS sensors connected to pilgrims can identify accidents and monitor their movement using intelligent algorithms. The Crowd e-Health and IoT component includes tasks such as the e-Health smart system utilizing IoT devices, crowd accident detection, crowd flow direction monitoring, and real-time analysis of health records to support decision-makers in ensuring a safe Hajj. These tasks are detailed in Section III-B.

The third component, Crowd Evacuation, encompasses tasks such as real-time evacuation path recommendations, early fire detection, and emergency alert identification. These tasks are discussed in Section III-C.

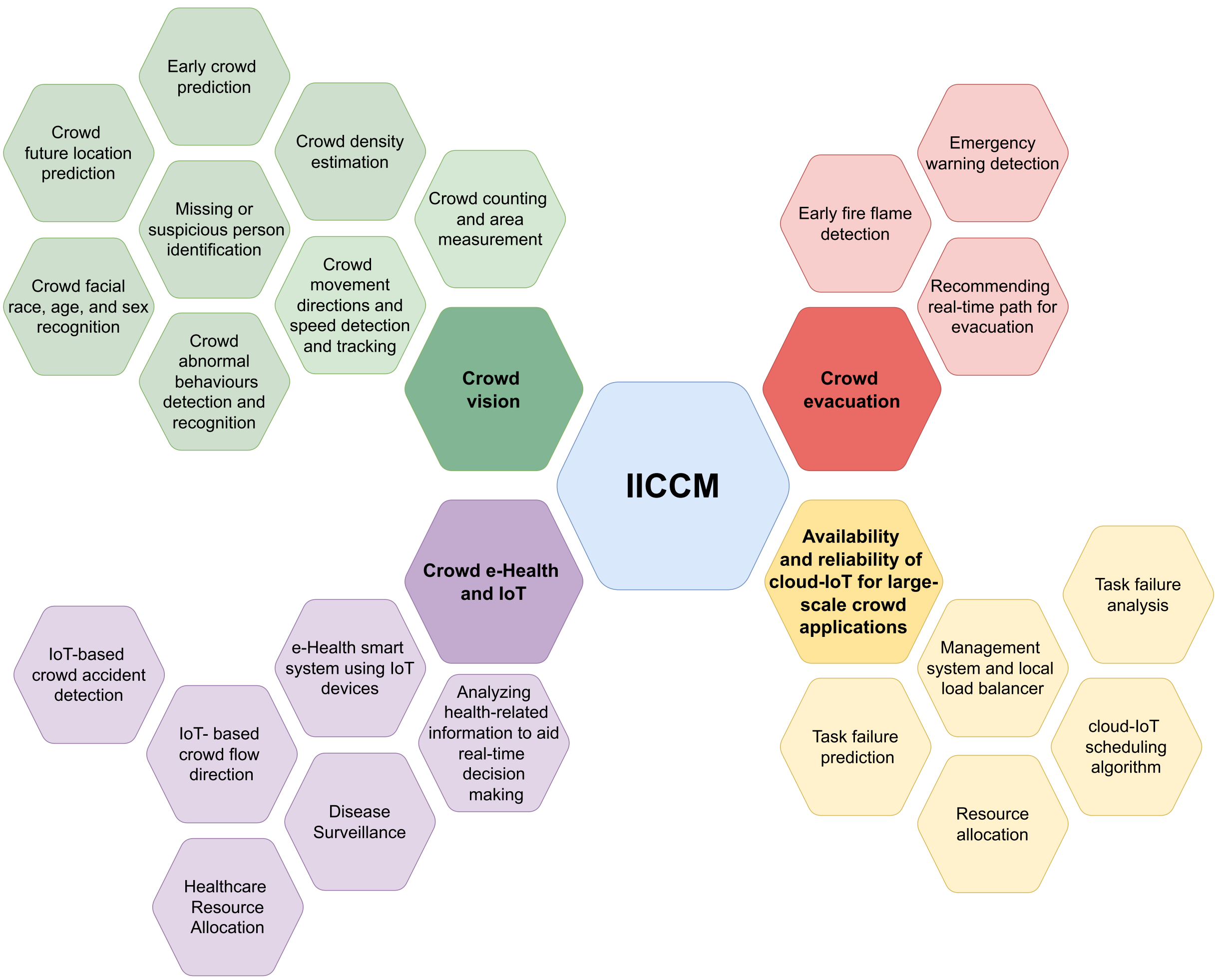

FIGURE 2. **The proposed Intelligent Integrated framework for Crowd Control and Management (IICCM).**

The final component addresses the availability and reliability of cloud-IoT applications, focusing on real-time solutions for large-scale crowds. The tasks include system management, local load balancing, task failure analysis and prediction, resource allocation, and cloud-IoT scheduling algorithms. These tasks are discussed in Section III-D.

A. CROWD VISION COMPONENT

In this section, we present the tasks related to Crowd Vision. The tasks are as follows:

1) CROWD COUNTING AND AREA MEASUREMENT

The task of crowd counting and area measurement primarily focuses on estimating the number of people in a crowd and the area they occupy. This is achieved by counting the heads of individuals and measuring the surrounding area. A CNN method and advanced learning techniques, such as few-shot learning, are used to perform this task without the need for extensive labelling. The CNN extracts and learns the features of people's heads, and sliding windows or image pyramids are applied to locate these heads with bounding boxes. The area is measured by calculating the number of pixels surrounding the crowd. Input data is sourced from video streams captured by CCTV cameras, and the CNN classifier requires only a few annotated examples of people's heads to perform crowd counting. The area measurement is based on the pixel count around the crowd.

2) CROWD MOVEMENT DIRECTIONS AND SPEED DETECTION AND TRACKING

The task of crowd movement direction and speed detection focuses on analyzing crowd movement behaviors. Directions are classified into north, south, east, and west, while speed is categorized into low, medium, and high, based on predefined threshold values. Direction detection is achieved using optical flow features and ViT with attention maps. Optical flows capture the magnitudes and orientations of moving objects, and direction trajectories are tracked using algorithms such as Kalman filters [85], particle filters [86], or other robust tracking methods. Crowd movement and tracklets are predicted by forecasting the crowd's position in subsequent frames. Video streams from CCTV cameras

TABLE 1. A comparison of smartwatch features.

Features	HUAWEI GT2 Pro	HUAWEI FIT	Amazfit Bip S	Galaxy Active 2	FitBit Sense	Apple Watch Series 8	Apple Watch Series 9	Apple Watch Ultra 2
GPS tracking	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Heart rate surveillance	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Oxygen ratio in blood	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Skin temperature	No	No	No	Yes	Yes	Yes	Yes	Yes
Heart Electrocardiogram (ECG)	No	No	No	Yes	Yes	Yes	Yes	Yes
Fall detection	No	No	No	Yes	No	Yes	Yes	Yes
Average battery life	14 Days	10 Days	15 Days	61 hours	6 Days	18 hours	36 hours	36 hours
Accuracy level	Medium	Medium	Medium	High	Medium	Very high	Very high	Very high

provide the input data. For speed detection, the change in distance is calculated based on the target pixels representing the crowd across consecutive frames in the video streams.

3) CROWD DENSITY ESTIMATION

The goal of crowd density estimation is to classify the density level within a crowd scene. This task is performed using two main functions. The first function involves counting the number of people's heads using the AI-based counter described in previously. Next, the frames are divided into grids, and density levels are estimated based on predefined thresholds. To count heads, the CNN network is fine-tuned to extract and learn high-level features of heads. Sliding windows or image pyramids are then used to locate and count the heads. Afterward, various thresholds are applied to estimate the density of the crowd and the density level (high, medium, or low) is classified.

4) MISSING OR SUSPICIOUS PERSON IDENTIFICATION

The goal of this task is to identify and locate missing or suspicious individuals in large-scale crowds in real time, which is crucial to ensuring security and public safety during events such as Hajj. The identification process should be linked to the facial recognition database of government or security authorities to effectively perform this task. Data acquisition involves video streams from CCTV cameras.

The first step is detecting all faces in real-time video surveillance. A pre-trained ViT with attention maps is used to localize faces within bounding boxes. Next, a pre-trained CNN based YOLO is employed to extract high-level facial features. These features are then compared with the facial database using an SNN to identify the individual. It's essential that the identification process is fast and real-time, as any delay could result in losing track of the missing or suspicious person. However, challenges such as partial or full occlusions and individuals being out of camera range may arise in large-scale crowds.

5) CROWD ABNORMAL BEHAVIORS DETECTION AND RECOGNITION

Based on video streams from CCTV cameras, it is essential for the system to detect and recognize abnormal behaviors in the crowd during the Hajj event, as these behaviors can pose potential risks to crowd flow. Abnormal behaviors include actions such as sudden standing, fighting, sleeping, running, jumping, pickpocketing, stealing, and crowds moving in the opposite or wrong direction. It is important to note that the definition of abnormal behaviors in the context of Hajj may not apply to other events, and vice versa.

To accomplish this task, we generate normal and abnormal behavior samples using GANs. An attention map based ViT is then used to localize abnormal behaviors. Video streams from CCTV cameras, along with datasets such as HAJJv1 [62], HAJJv2 [61], UCF Crime [82], UCSD [83], and UMN [84], are utilized as inputs for training and evaluating this task. These datasets are selected due to their availability and their representation of abnormal human behaviors.

6) EARLY CROWD PREDICTION

The goal of this task is to predict the formation of a crowd in a given area, based on the accumulation of people in video frames over time. This task utilizes GANs, optical flows, and ViT. Optical flows provide the magnitudes and orientations of crowd movement between consecutive frames. The GAN generates various optical flow patterns, which are then trained using the ViT classifier to predict the future location of the crowd. For data acquisition, video streams from CCTV cameras should be directed at areas expected to become crowded, allowing for the generation of optical flows, followed by training and testing the classifier.

7) CROWD FUTURE LOCATION PREDICTION

In contrast to the task of early crowd prediction, this task aims to forecast the future location of the crowd. The prediction is based on prior information about the crowd's past locations

TABLE 2. A list of the framework tasks, descriptions, and expected outcomes.

Task	Method	Algorithm Description	Expected Input	Expected Output
Crowd Future Location Prediction	Optical flows, GAN, and ViT classifiers	The GAN generates different optical flows. Optical flows are trained by the ViT classifier to predict the future location of the crowd.	Video streams from CCTV cameras; generated different crowd directions and optical flow samples.	Predicted crowd future location.
Early Crowd Prediction	GAN and ViT	Generated crowd density using GAN; a ViT classifier predicts the density of the crowd over time using the same scene.	Video streams from CCTV cameras; generated different crowd density samples.	Predicted crowd occurrences.
Crowd Facial Race, Age, and Sex Recognition	GAN and DNN	Generated different facial race, age, and sex images using GAN. A DNN classifier is used to learn facial features and predict their categories.	Video streams from CCTV cameras; generated different race, age, and sex facial images.	Predicted crowd facial race, age, and sex.
e-Health Smart System using IoT Devices	Based on health sensors such as ECG	Collecting information from the perception layer, which is then sent through a reliable network system to the cloud for analyzing and monitoring the health status of pilgrims/visitors.	Blood pressure, heart rate, temperature, and oxygen saturation data.	Real-time health status presented on the dashboard system.
IoT-Based Crowd Flow Direction	Based on collecting geolocation information of pilgrims/visitors	Real-time collecting geolocation information classified into four colors: (1) red for north, (2) green for south, (3) yellow for east, and (4) orange for west.	Geolocation of pilgrims/visitors.	Monitoring and displaying the direction of pilgrims on a big screen (dashboard presentation).
IoT-Based Crowd Accident Detection	Based on collecting information related to geolocation of pilgrims	Collecting real-time active geolocation data of pilgrims and providing estimated travel times between locations. If there's a significant delay compared to the expected duration, the system alerts the Hajj authority. Alerts are classified into low, medium, and high danger using fuzzy logic.	Geolocation of pilgrims/visitors.	Alert notification to Hajj authority.
Task Failure Analysis	Statistical Analysis	Identify patterns and characteristics of failed crowd application jobs.	Crowd application traces.	Abnormal behavior that causes incoming tasks to fail.
Task Failure Prediction	RF and DNN	Predicting task termination status.	Crowd application traces.	Failed or completed task.
Resource Allocation	Scheduling Algorithms	Select the desired level of fault tolerance based on application requirements; transfer incoming tasks based on the requirements.	Required level of computation and availability.	Efficient resource allocation.
Analyzing Health-Related Information for Real-Time Decision-Making	LSTM	The LSTM is effective in capturing sequential dependencies and handling time-series data when analyzing health-related data. Input includes symptoms, medical reports, demographics, treatment history, etc.	Health-related information from wearable devices.	Disease diagnosis, treatment recommendations, and personalized health advice.
Disease Surveillance	SVM and RF	Hybrid model uses relevant features to facilitate early interventions against potential disease outbreaks.	Disease symptoms reported, weather conditions, and people's activity patterns.	Alarm detection for early interventions against disease outbreaks.
Healthcare Resource Allocation	Different Machine Learning Methods	Certain models are built using effective aspects to predict the requirements for diverse healthcare services.	Demographic data of pilgrims with health conditions, geographic data, historical data on aid usage, and disease prevalence.	An optimal distribution plan for healthcare resources.

and the directions detected in the earlier task. A fused data framework is introduced to predict the next crowd location. Expert input from both geographical sources and Hajj domain specialists is integrated to provide guidance on identifying directions, location names, and relevant information during Hajj events. For data acquisition, video streams from CCTV cameras should focus on areas anticipated to be crowded.

TABLE 3. Additional framework tasks, descriptions, and expected outcomes.

Task	Method	Algorithm Description	Expected Input	Expected Output
Large-scale Crowd Counting and Place Measurement	CNN with novel unlabeled learning methods	The CNN extracts features from people's heads using recent techniques such as few-shot learning, employing sliding windows or image pyramids for head localization. The crowd area is quantified by the pixel count surrounding it.	Video streams from CCTV cameras; few annotated head examples for counting.	Large-scale crowd counts and place measurements.
Crowd Directions/Speed Detection and Tracking	Crowd direction detected via optical flows and ViT; tracking with Kalman filter; speed estimation through distance changes	Detects crowd direction using optical flows and attention maps, predicts crowd movements for tracking in subsequent frames, and calculates distance changes for speed estimation.	Video streams from CCTV cameras.	Detection and tracking of crowd direction and speed.
Crowd Density Estimation	CNN for crowd counting; threshold-based density estimation	Counts people in video by extracting head features with CNN methods; classifies density levels based on head counts and thresholds.	Video streams from CCTV cameras.	Total head counts and categorized crowd density (high, medium, low).
Missing/Suspicious Person Identification	Face detection via pre-trained CNN based YOLO person detector; feature extraction with another CNN; identification using an SNN	Implements YOLO for real-time face detection, extracts features using a pre-trained CNN, and identifies individuals by comparing feature similarities through an SNN.	Video streams from CCTV cameras; facial datasets (CelebA [83], LFW [84], HAJJv1 [85], HAJJv2 [86]).	Identification of missing or suspicious persons in crowds.
Crowd Abnormal Behavior Detection and Recognition	GAN and ViT	Generates anomalous behavior patterns with GAN, then detects and recognizes these abnormalities using attention maps based ViT classifier.	Video streams from CCTV cameras; HAJJv1 [85], HAJJv2 [86], UCF Crime [87], UCSD [88], and UMN [89] datasets.	Detection and recognition of abnormal crowd behaviors.
Early Fire Flame Detection	GAN and ViT classifier utilizing attention maps	GAN generates fire samples to train a ViT classifier, which detects fire and segments affected areas using attention maps.	Video streams from CCTV cameras; generated fire samples.	Fire detection and alerting system.
Emergency Warning Detection	GAN and ViT classifier	Utilizes GAN-generated samples of running individuals to train the ViT classifier for early emergency detection.	Video streams from CCTV cameras; samples of generated running individuals.	Detection of emergency warnings.
Recommending Real-time Path for Evacuation	Neural network-based counting combined with LPSs	Neural network counts individuals at each exit and suggests the least crowded route while LPS devices provide the shortest path.	Few annotated and labelled human head examples; paths to exits via LPS devices.	Recommended real-time evacuation routes.

8) CROWD FACIAL RACE, AGE, AND SEX RECOGNITION

To accomplish this task, a Deep Neural Network (DNN) model is used to recognize race, sex, and age in facial images. The model is trained on a set of facial examples generated by a GAN, with each image labeled with the person's race, sex, and age. We also employ advanced learning techniques, such as few-shot learning, which require minimal labeling. Once trained, the model can predict the race, sex, and age of any person detected in CCTV camera footage.

This algorithm is effective because it can learn complex patterns in data. The DNN model is capable of identifying facial features associated with different races, sexes, and ages. Additionally, the algorithm can handle occlusions, a common challenge in facial recognition. However, the accuracy of the model is sensitive to the quality of the training data. If the training data is not representative of real-world conditions, the algorithm's predictions may be inaccurate. Moreover, the model's accuracy could be biased towards certain groups, depending on the composition and balance of the training data. Overall, this algorithm is a powerful and effective tool for recognizing race, sex, and age in facial images from video streams.

B. CROWD E-HEALTH AND IOT COMPONENT

In this section, we introduce the crowd vision tasks. The tasks are as follows:

1) E-HEALTH SMART SYSTEM USING IOT DEVICES

IoT medical devices gather data from sensors and wirelessly transmit it to cloud platform services. The platform stores the data, provides medical insights based on patients' historical records, and enables medical staff to update, review, and test the data. Both medical staff and patients can access the platform from any mobile or stationary device with an internet connection.

2) IOT-BASED CROWD FLOW DIRECTION

Crowd flow direction using IoT devices is based on collecting geolocation data from pilgrims. This data is analyzed in the cloud system to determine the direction of the flow. The system classifies the flow into four directions: (1) red for north, (2) green for south, (3) yellow for east, and (4) orange for west. Additionally, a dashboard displays the real-time flow direction, which is monitored by an administrator to provide further assistance to pilgrims.

3) IOT-BASED CROWD ACCIDENT DETECTION

Crowd accident detection using IoT technology helps identify areas with high crowd density. This feature of our proposed framework assists the Hajj authority in making decisions to mitigate the impact of such situations. The framework alerts the authority through a notification panel during the Hajj event. It relies on real-time active sensors and geolocation-based sensors, with the cloud system processing the data gathered from the field. The system then sends alerts to the administrator in case of any unusual situations involving pilgrims or visitors. The framework uses fuzzy logic, which classifies the alerts into three levels of danger: low, medium, and high.

4) ANALYZING HEALTH-RELATED INFORMATION TO AID REAL-TIME DECISION MAKING

Analyzing health-related information, particularly patient data, has the potential to transform real-time decision-making in crowds, facilitating more informed and timely healthcare interventions. This study focuses on utilizing patient data to support real-time decision-making in crowds, leveraging the collective intelligence and experiences of individuals.

The methodology involves gathering patient data from various sources, including EHRs, wearable devices, and health monitoring applications. The collected data are thoroughly preprocessed to ensure quality and consistency, which includes data cleaning, normalization, and addressing missing values. Feature extraction techniques are applied to identify key information from the patient data, such as demographic details, medical history, symptoms, and treatment outcomes. These features provide valuable insights into the patients' health status and can be used to support real-time decision-making.

To enable real-time decision making, machine learning algorithms are utilized, specifically employing techniques such as classification, regression, and clustering. These algorithms analyze patient data to detect patterns, predict health outcomes, and offer recommendations for personalized interventions. One potential approach is the Long Short-Term Memory (LSTM) classifier, which is highly effective in capturing sequential dependencies and processing time-series data, making it well-suited for analyzing health-related information. The input can include pilgrims' symptoms, medical reports, demographics, treatment history, and other relevant data.

Using the collective knowledge within the patient data, this study aims to empower healthcare professionals and individuals to make informed decisions based on real-time insights. The results of this research have the potential to improve healthcare care delivery, improve patient outcomes, and advance the development of intelligent decision support systems in crowd-based healthcare contexts.

5) DISEASE SURVEILLANCE

This task involves systematically collecting, examining, and analyzing health data from the pilgrim crowd to prevent and control disease outbreaks. Support Vector Machine (SVM) and Random Forest (RF) models are used to predict disease outbreaks by analyzing aggregated data, including reported symptoms in specific areas, weather conditions, and activity patterns of the crowd. The outcome is a surveillance system that employs alarm detection for early interventions and effective planning to mitigate potential disease outbreaks.

6) HEALTHCARE RESOURCE ALLOCATION

This task focuses on determining the optimal allocation of healthcare resources, including supplies, medical instruments, and vaccines, among pilgrims, prioritizing those most in need. Machine learning techniques are used to predict the healthcare requirements and maximize the efficiency of resource distribution. Key features include demographic data of pilgrims with health conditions, geographic data, historical aid usage, and disease prevalence. An optimized allocation strategy is the expected outcome that ensures the effective use of limited healthcare resources and enables rapid response during medical crises.

C. CROWD EVACUATION COMPONENT

In this section, we introduce the crowd evacuation tasks. The tasks are as follows:

1) EARLY FIRE FLAME DETECTION

The objective of early fire detection is to identify any fire within a crowd scene. To achieve this, a GAN generates fire samples to train the ViT classifier for fire detection. Attention maps help segment the fire area, enabling the ViT to focus on fire flames during classification. Video data is captured by CCTV cameras to train and test the classifier. The fire flame detection system should be linked to an alarm system to notify people and security authorities. A limitation of this task is that the range of CCTV cameras may not cover all areas, potentially missing fires in certain locations.

2) EMERGENCY WARNING DETECTION

Similar to early fire detection, the goal of this task is to detect emergency warnings in a large-scale crowd in real-time. Running individuals may indicate an emergency situation. The GAN is used to generate samples of running people, which are then used to train a ViT classifier for early emergency detection. Video streams from CCTV cameras are used to acquire data for training and testing the classifier.

3) RECOMMENDING REAL-TIME PATH FOR EVACUATION

In emergent situations such as fires, criminal attacks, or terrorist incidents, large-scale indoor crowds face the potential danger of accidents, including injuries and fatalities. These situations force individuals to move quickly, evacuating simultaneously and heading toward the nearest exits. This mass movement can lead to individuals pushing through the crowd, which may result in harmful consequences. Therefore, ensuring the safety of crowd evacuation and the methods employed is critical to avoiding potential dangers and ensuring a safe flow of people. In this task, a method is proposed to recommend a real-time safe evacuation route for indoor crowds, based on crowd counting and exit directions. A neural network-based people counter and a Local Positioning System (LPS) devices are used, integrated with an app to suggest real-time evacuation paths to the least crowded gates.

4) EMERGENCY WARNING DETECTION

Similar to early fire detection, the objective of this task is to identify emergency warnings in a large-scale crowd in real-time. Running individuals may signal an emergency situation. The GAN generates samples of running people, which are used to train a ViT classifier for early emergency detection. Video streams from CCTV cameras are used for data acquisition to train and test the classifier.

D. COMPONENT OF AVAILABILITY AND RELIABILITY OF CLOUD-IOT FOR LARGE-SCALE CROWD APPLICATIONS

In this section, we introduce the tasks of availability and reliability of cloud-IoT for large-scale crowd applications. The tasks are as follows:

1) MANAGEMENT SYSTEM AND LOCAL LOAD BALANCER

The management system and local load balancer are crucial components of a cloud-IoT framework. The management system oversees the availability of local resources, while the load balancer allocates incoming tasks to operational nodes based on their computational needs. Task scheduling is determined by task classification and resource requirements. Local servers are ideal for handling high-priority tasks with low to medium computational demands, such as temperature monitoring and control, in a smart city system. This is because local servers provide several advantages over cloud-based servers for these types of tasks:

- Reduced latency: Local servers are generally positioned closer to the IoT devices they monitor and manage, leading to lower latency. This is crucial for high-priority tasks that demand real-time execution, such as temperature regulation.
- Increased reliability: Local servers are less susceptible to outages compared to cloud-based servers, making them more reliable for critical tasks.
- Enhanced security: Local servers are less prone to security breaches compared to cloud-based servers, making them more suitable for tasks that involve handling sensitive data.
- Cost efficiency: Local servers can be more cost-effective than cloud-based servers, particularly in large-scale systems.

2) TASK FAILURE ANALYSIS

Modern applications, such as smart cities, home automation, and eHealth, demand a new approach to enhancing cloud application reliability and availability. The complexity and scale of cloud environments make cloud services, both hardware and software, prone to failures. This framework analyzes and characterizes the behavior of failed and successfully completed jobs using crowd application traces. The goal of failure analysis and prediction is to examine the dynamic behavior of cloud applications in order to improve performance and minimize task failures.

3) TASK FAILURE PREDICTION

The proposed framework aims to predict job/task failures in the cloud with high accuracy using machine learning classification algorithms. By doing so, it helps reduce wasted cloud resources and enhances cloud infrastructure utilization. Our failure prediction model forecasts the termination status of submitted tasks before execution, which aids in improving resource utilization and cloud application efficiency. We evaluate our model using crowd application workloads and test various machine learning and deep learning models on crowd application traces to identify the most accurate one. This model can improve the reliability and availability of cloud services by predicting job failures, developing scheduling algorithms, adjusting priority policies, or limiting task resubmissions.

The failure prediction model design process consists of four key phases:

- Data Collection and Storage: The cloud management system monitors application metrics and requested resources (memory, CPU, disk space) and stores the data in cloud storage for preprocessing and analysis.
- Data Preprocessing and Filtration: This critical step enhances the quality of the failure prediction model by cleaning and organizing the data.
- Machine Learning: A machine learning algorithm is applied to the cloud workload traces to predict job or task failures.

- Decision-Making: Based on the prediction results, the cloud management system makes decisions. If the job is predicted to succeed, it is submitted and scheduled to available nodes. If a failure is predicted, failure mitigation techniques are applied.

4) RESOURCE ALLOCATION

We categorize availability levels into high, medium, and low to optimize cloud costs. This provides customers with various options to select the appropriate level of fault tolerance based on their application needs.

IV. IMPLEMENTATION DETAILS

This framework will be implemented using the Google Colaboratory (Colab) programming environment with the Python programming language. Google Colab is a free, cloud-based Jupyter notebook environment, making it ideal for developing and running advanced machine learning models. Additionally, Colab provides access to high-performance GPUs, which are crucial for processing large volumes of video streams in the proposed framework.

The project is anticipated to span two years for completion. During the first year, efforts will focus on developing the framework's components, implementing their tasks, and integrating them into a cohesive system. The second year will be dedicated to system evaluation and deployment for end-users.

V. DISCUSSION

The Integrated Intelligent Crowd Control and Management framework (IICCM) presents a significant advancement in the management of large-scale gatherings, particularly during the Hajj pilgrimage, which faces unique and critical challenges due to the sheer volume and diversity of attendees. By leveraging cutting-edge technologies such as AI, CV, and the IoT, the proposed framework offers a holistic solution to ensure safety, efficiency, and enhanced decision-making during crowd events.

One of the core strengths of the IICCM framework is its ability to integrate various disciplines and technologies into a unified framework. The combination of crowd vision, IoT-enabled health monitoring, real-time evacuation systems, and cloud-based availability and reliability ensures that the framework is both comprehensive and adaptive to the dynamic nature of crowd management. For instance, the framework's crowd vision component effectively monitors crowd density, movement patterns, and abnormal behaviors, offering actionable insights to security and health authorities. This integration exemplifies how interdisciplinary solutions can tackle complex problems in real-world scenarios.

The framework's emphasis on addressing Hajj-specific challenges, such as identifying missing persons, managing health emergencies, and ensuring orderly evacuations, demonstrates its practical relevance. The incorporation of IoT devices, such as wearables for health monitoring and geolocation sensors for tracking movement, is particularly notable. These tools not only provide real-time data but also facilitate timely interventions to prevent potential health crises or safety breaches. The ability to predict crowd dynamics using advanced models, such as GANs and ViTs, further enhances the framework's capacity to mitigate risks proactively.

Despite its strengths, the proposed framework faces several challenges. One major concern is its reliance on IoT and cloud computing, which introduces potential issues related to data security, privacy, and system latency. The transmission of real-time crowd data to cloud servers increases the risk of unauthorized access, data breaches, and cyberattacks, necessitating robust encryption and authentication mechanisms. Privacy concerns also arise as continuous monitoring may capture sensitive information, requiring strict compliance with data protection regulations and ethical guidelines. Additionally, dependence on cloud computing can lead to latency issues, particularly in high-density environments where rapid decision-making is critical for crowd safety. While this research does not specifically address these challenges, they offer valuable directions for future work, such as integrating edge computing for real-time processing, enhancing cybersecurity measures, and developing privacy-preserving AI models.

Ensuring the scalability of the framework during peak periods, such as Hajj, necessitates a robust infrastructure and fault-tolerant mechanisms. Additionally, the context-specific nature of certain tasks, such as crowd abnormal behavior detection, may restrict the framework's applicability in other environments unless significant customization and model retraining are undertaken.

Although the framework is tailored for Hajj, its principles and methodologies have wider applications for managing large-scale events globally, such as sports competitions, concerts, and political rallies. By showcasing its effectiveness in one of the most complex crowd management scenarios, the IICCM framework establishes a benchmark for the development and implementation of intelligent crowd management systems worldwide.

VI. CONCLUSION

This work presents the IICCM framework, which integrates advancements in computer vision (CV), artificial intelligence (AI), and Internet of Things (IoT) technologies to enhance the safety and security of large crowds. CV enables real-time individual identification and tracking, while AI analyzes crowd behavior to anticipate potential risks. IoT gathers environmental data to optimize crowd movement, minimize congestion, and facilitate assistance. Additionally, the framework aids in emergency evacuation planning by assessing crowd dynamics and determining the safest, most efficient evacuation routes.

While the framework is designed for a wide range of events, Hajj—an exceptionally large and complex annual pilgrimage—serves as the ultimate stress test for the IICCM system. Managing millions of participants from diverse cultural and linguistic backgrounds in a dynamic environment highlights the robustness of this solution. By successfully addressing the unique challenges of Hajj, the IICCM framework demonstrates its potential as a scalable, adaptable model for improving crowd safety and management at large-scale events worldwide. This work offers valuable insights for decision-makers considering the adoption of advanced crowd control technologies.

The proposed framework holds the potential to significantly enhance the safety and security of Hajj event. It also aims to improve the efficiency of crowd management while reducing the risk of incidents such as stampedes.

Beyond safety and security, the framework could enhance the overall crowd experience. For instance, it could provide pilgrims with real-time information about the location of essential facilities such as restrooms, food vendors, and medical stations. Additionally, it could offer personalized recommendations for transportation and accommodation, further improving the experience at large events.

To enhance the effectiveness of the proposed framework, future iterations of the IICCM framework could explore additional features, such as incorporating multilingual support for real-time communication with diverse participants and utilizing blockchain technology for secure data sharing among stakeholders. Additionally, integrating feedback mechanisms from end-users, including pilgrims and security personnel, could help refine the framework to better meet evolving needs. Expanding the dataset to cover a broader range of crowd scenarios would also enhance the generalizability of the models used in the framework.

REFERENCES

[1] W. Ertel, *Introduction to Artificial Intelligence*. Cham, Switzerland: Springer, 2024.

[2] B. Zhan, D. Monekosso, P. Remagnino, S. A. Velastín, and L.-Q. Xu, "Crowd analysis: A survey," *Mach. Vis. Appl.*, vol. 19, nos. 5–6, pp. 345–357, Apr. 2008.

[3] T. Li, H. Chang, M. Wang, B. Ni, R. Hong, and S. Yan, "Crowded scene analysis: A survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 3, pp. 367–386, Mar. 2015.

[4] M. S. Zitouni, H. Bhaskar, J. Dias, and M. E. Al-Mualla, "Advances and trends in visual crowd analysis: A systematic survey and evaluation of crowd modelling techniques," *Neurocomputing*, vol. 186, pp. 139–159, Apr. 2016.

[5] B. Ganga, B. Lata, and K. Venugopal, "Object detection and crowd analysis using deep learning techniques: Comprehensive review and future directions," *Neurocomputing*, vol. 597, Sep. 2024, Art. no. 127932.

[6] T. Chavdarova and F. Fleuret, "Deep multi-camera people detection," in *Proc. 16th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2017, pp. 848–853.

[7] T. Chavdarova, P. Baqué, S. Bouquet, A. Maksai, C. Jose, T. Bagautdinov, L. Lettry, P. Fua, L. Van Gool, and F. Fleuret, "WILDTRACK: A multi-camera HD dataset for dense unscripted pedestrian detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5030–5039.

[8] Y. Zhang, C. Wang, X. Wang, W. Liu, and W. Zeng, "VoxelTrack: Multi-person 3D human pose estimation and tracking in the wild," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 2, pp. 2613–2626, Feb. 2023.

[9] J. Holler, V. Tsiatsis, C. Mulligan, S. Karnouskos, S. Avesand, and D. Boyle, *Internet of Things*. New York, NY, USA: Academic, 2014.

[10] C. Stergiou, K. E. Psannis, B.-G. Kim, and B. Gupta, "Secure integration of IoT and cloud computing," *Future Gener. Comput. Syst.*, vol. 78, pp. 964–975, Jan. 2018.

[11] A. Javed, A. Malhi, T. Kinnunen, and K. Främling, "Scalable IoT platform for heterogeneous devices in smart environments," *IEEE Access*, vol. 8, pp. 211973–211985, 2020.

[12] K. Aruna and G. Pradeep, "Performance and scalability improvement using IoT-based edge computing container technologies," *Social Netw. Comput. Sci.*, vol. 1, no. 2, p. 91, Mar. 2020.

[13] V. S. Naresh, S. S. Pericherla, P. S. R. Murty, and S. Reddi, "Internet of Things in healthcare: Architecture, applications, challenges, and solutions," *Comput. Syst. Sci. Eng.*, vol. 35, no. 6, pp. 411–421, Jan. 2020.

[14] A. Sharma and R. B. Battula, "The Internet of Things solutions for transportation," in *AI and IoT for Sustainable Development in Emerging Countries: Challenges and Opportunities*. Cham, Switzerland: Springer, 2022, pp. 291–324.

[15] N. H. Motlagh, M. Mohammadrezaei, J. Hunt, and B. Zakeri, "Internet of Things (IoT) and the energy sector," *Energies*, vol. 13, no. 2, p. 494, Jan. 2020.

[16] A. Ksentini, M. Jebalia, and S. Tabbane, "IoT/cloud-enabled smart services: A review on QoS requirements in fog environment and a proposed approach based on priority classification technique," *Int. J. Commun. Syst.*, vol. 34, no. 2, p. 4269, Jan. 2021.

[17] P. Borra, "Comparison and analysis of leading cloud service providers (AWS, Azure and GCP)," *Int. J. Adv. Res. Eng. Technol. (IJARET)*, vol. 15, pp. 266–278, May 2024.

[18] S. Adibi, A. Rajabifard, D. Shojaei, and N. Wickramasinghe, "Enhancing healthcare through sensor-enabled digital twins in smart environments: A comprehensive analysis," *Sensors*, vol. 24, no. 9, p. 2793, Apr. 2024.

[19] P. Eslami, S. R. N. Kalhori, and M. Taheriyan, "eHealth solutions to fight against COVID-19: A scoping review of applications," *Med. J. Islamic Republic Iran*, vol. 35, p. 43, Apr. 2021.

[20] X. Chen, W. Yao, Y. Li, D. Liang, H. Zheng, S. Shakil, S. Wang, and T. Sun, "IG-GCN: Empowering e-health services for Alzheimer's disease prediction," *IEEE Trans. Consum. Electron.*, vol. 70, no. 4, pp. 6442–6451, Nov. 2024.

[21] S. K. Mohapatra and M. N. Mohanty, "Big data classification with IoT-based application for e-health care," in *Cognitive Big Data Intelligence With a Metaheuristic Approach*. Amsterdam, The Netherlands: Elsevier, 2022, pp. 147–172.

[22] I. J. B. Do Nascimento, A. B. Pizarro, J. M. Almeida, N. Azzopardi-Muscat, M. A. Gonçalves, M. Björklund, and D. Novillo-Ortiz, "Infodemics and health misinformation: A systematic review of reviews," *Bull. World Health Org.*, vol. 100, no. 9, p. 544, 2022.

[23] H. Hasannejadasl, C. Roumen, Y. Smit, A. Dekker, and R. Fijten, "Health literacy and eHealth: Challenges and strategies," *JCO Clin. Cancer Informat.*, vol. 6, Nov. 2022, Art. no. e2200005.

[24] I. Keshta and A. Odeh, "Security and privacy of electronic health records: Concerns and challenges," *Egyptian Informat. J.*, vol. 22, no. 2, pp. 177–183, Jul. 2021.

[25] J. Shahid, R. Ahmad, A. K. Kiani, T. Ahmad, S. Saeed, and A. M. Almuhaideb, "Data protection and privacy of the Internet of Healthcare Things (IoHTs)," *Appl. Sci.*, vol. 12, no. 4, p. 1927, Feb. 2022.

[26] M. Zhou, H. Dong, P. A. Ioannou, Y. Zhao, and F.-Y. Wang, "Guided crowd evacuation: Approaches and challenges," *IEEE/CAA J. Autom. Sinica*, vol. 6, no. 5, pp. 1081–1094, Sep. 2019.

[27] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020.

[28] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," 2017, *arXiv:1710.10196*.

[29] A. Brock, J. Donahue, and K. Simonyan, "Large scale GAN training for high fidelity natural image synthesis," 2018, *arXiv:1809.11096*.

[30] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4396–4405.

[31] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of StyleGAN," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 8107–8116.

[32] R. Abdal, Y. Qin, and P. Wonka, "Image2StyleGAN: How to embed images into the StyleGAN latent space?" in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 4431–4440.

[33] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.

[34] C. R. Chen, Q. Fan, and R. Panda, "CrossViT: Cross-attention multi-scale vision transformer for image classification," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 347–356.

[35] G. Yang, H. Tang, M. Ding, N. Sebe, and E. Ricci, "Transformer-based attention networks for continuous pixel-wise prediction," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 16249–16259.

[36] X. Li, Y. Hou, P. Wang, Z. Gao, M. Xu, and W. Li, "Trear: Transformer-based RGB-D egocentric action recognition," *IEEE Trans. Cognit. Develop. Syst.*, vol. 14, no. 1, pp. 246–252, Mar. 2022.

[37] D. A. Coccomini, N. Messina, C. Gennaro, and F. Falchi, "Combining EfficientNet and vision transformers for video deepfake detection," in *Proc. Int. Conf. image Anal. Process.* Cham, Switzerland: Springer, Jan. 2022, pp. 219–229.

[38] X. Pan, Z. Xia, S. Song, L. E. Li, and G. Huang, "3D object detection with pointformer," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 7459–7468.

[39] H. Larochelle, D. Erhan, and Y. Bengio, "Zero-data learning of new tasks," in *Proc. AAAI*, Jul. 2008, pp. 646–651.

[40] L. Fei-Fei, R. Fergus, and P. Perona, "One-shot learning of object categories," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 4, pp. 594–611, Apr. 2006.

[41] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2020, pp. 18661–18673.

[42] A. Pentina and C. H. Lampert, "Lifelong learning with non-IID tasks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015.

[43] E. Mlybari, A. Hedar, G. Elsayed, E. Yosry, A. E. Abdelhakim, M. Elbissy, and A. Gadelrab, "Systems and methodologies for performing intelligent perception based real-time counting," U.S. Patent 15 060 173, Sep. 8, 2016.

[44] K. R. Ahuja and N. N. Charniya, "A survey of recent advances in crowd density estimation using image processing," in *Proc. Int. Conf. Commun. Electron. Syst. (ICCES)*, Jul. 2019, pp. 1207–1213.

[45] X. Ding, F. He, Z. Lin, Y. Wang, H. Guo, and Y. Huang, "Crowd density estimation using fusion of multi-layer features," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 4776–4787, Aug. 2021.

[46] L. Chen, G. Wang, and G. Hou, "Multi-scale and multi-column convolutional neural network for crowd density estimation," *Multimedia Tools Appl.*, vol. 80, no. 5, pp. 6661–6674, Feb. 2021.

[47] M. S. Saleem, M. J. Khan, K. Khurshid, and M. S. Hanif, "Crowd density estimation in still images using multiple local features and boosting regression ensemble," *Neural Comput. Appl.*, vol. 32, no. 21, pp. 16445–16454, Nov. 2020.

[48] L. Zhu, C. Li, Z. Yang, K. Yuan, and S. Wang, "Crowd density estimation based on classification activation map and patch density level," *Neural Comput. Appl.*, vol. 32, no. 9, pp. 5105–5116, May 2020.

[49] T. Alafif, Z. Hailat, M. Aslan, and X. Chen, "On detecting partially occluded faces with pose variations," in *Proc. 14th Int. Symp. Pervasive Syst., Algorithms Netw. 11th Int. Conf. Frontier Comput. Sci. Technol. 3rd Int. Symp. Creative Comput. (ISPAN-FCST-ISCC)*, Jun. 2017, pp. 28–37.

[50] A. Nadeem, M. Ashraf, N. Qadeer, K. Rizwan, A. Mehmood, A. AlZahrani, F. Noor, and Q. H. Abbasi, "Tracking missing person in large crowd gathering using intelligent video surveillance," *Sensors*, vol. 22, no. 14, p. 5270, Jul. 2022.

[51] L. Wang and A. A. Siddique, "Facial recognition system using LBPH face recognizer for anti-theft and surveillance application based on drone technology," *Meas. Control*, vol. 53, nos. 7–8, pp. 1070–1077, Aug. 2020.

[52] A. E. Abdel-Hakim and M. El-Saban, "FRPCA: Fast robust principal component analysis for online observations," in *Proc. 21st Int. Conf. Pattern Recognit. (ICPR)*, Nov. 2012, pp. 413–416.

[53] J. K. Y. Lau, A. Kaheel, M. El-Saban, M. Shawky, M. Gonzalez, A. El Baz, T. Deif, and A. A.-H. Aly, "Using facial data for device authentication or subject identification," U.S. Patent 9 652 663, May 16, 2017.

[54] A. E. Abdel-Hakim and M. El-Saban, "Face authentication using graph-based low-rank representation of facial local structures for mobile vision applications," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCV Workshops)*, Nov. 2011, pp. 40–47.

[55] T. Alafif, Z. Hailat, M. Aslan, and X. Chen, "On classifying facial races with partial occlusions and pose variations," in *Proc. 16th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2017, pp. 679–684.

[56] R. P. Holder and J. Tapamo, "Using facial expression recognition for crowd monitoring," in *Proc. 8th Pacific-Rim Symp. Image Video Technol.*, Wuhan, China. Cham, Switzerland: Springer, Jan. 2018, pp. 463–476.

[57] E. Mlybari, E. Yosry, A.-R. Hedar, A. E. Abdel-Hakim, and G. Elsayed, "Smart surface-mounted hybrid sensor system, method, and apparatus for counting," U.S. Patent 9 672 462, Jun. 6, 2017.

[58] R. Alotaibi, B. Alzahrani, R. Wang, T. Alafif, A. Barnawi, and L. Hu, "Performance comparison and analysis for large-scale crowd counting based on convolutional neural networks," *IEEE Access*, vol. 8, pp. 204425–204432, 2020.

[59] Y. Miao, J. Yang, B. Alzahrani, G. Lv, T. Alafif, A. Barnawi, and M. Chen, "Abnormal behavior learning based on edge computing toward a crowd monitoring system," *IEEE Netw.*, vol. 36, no. 3, pp. 90–96, May 2022.

[60] S. Wang, Z. Pu, Q. Li, and Y. Wang, "Estimating crowd density with edge intelligence based on lightweight convolutional neural networks," *Expert Syst. Appl.*, vol. 206, Nov. 2022, Art. no. 117823.

[61] T. Alafif, A. Hadi, M. Allahyani, B. Alzahrani, A. Alhothali, R. Alotaibi, and A. Barnawi, "Hybrid classifiers for spatio-temporal abnormal behavior detection, tracking, and recognition in massive Hajj crowds," *Electronics*, vol. 12, no. 5, p. 1165, Feb. 2023.

[62] T. Alafif, B. Alzahrani, Y. Cao, R. Alotaibi, A. Barnawi, and M. Chen, "Generative adversarial network based abnormal behavior detection in massive crowd videos: A Hajj case study," *J. Ambient Intell. Humanized Comput.*, vol. 13, no. 8, pp. 4077–4088, Aug. 2022.

[63] W. Raghupathi and V. Raghupathi, "Big data analytics in healthcare: Promise and potential," *Health Inf. Sci. Syst.*, vol. 2, no. 1, pp. 1–10, Dec. 2014.

[64] N. El Aboudi and L. Benhlima, "Big data management for healthcare systems: Architecture, requirements, and implementation," *Adv. Bioinf.*, vol. 2018, pp. 1–10, Jun. 2018.

[65] J. Liu, Y. Chen, and Y. Chen, "Emergency and disaster management-crowd evacuation research," *J. Ind. Inf. Integr.*, vol. 21, Mar. 2021, Art. no. 100191.

[66] Y. Tong and N. W. F. Bode, "Simulation investigation on crowd evacuation strategies for helping vulnerable pedestrians at different stages of egress," *Int. J. Disaster Risk Reduction*, vol. 84, Jan. 2023, Art. no. 103479.

[67] J. M. Chertkoff, R. H. Kushigian, and M. A. McCool Jr., "Interdependent exiting: The effects of group size, time limit, and gender on the coordination of exiting," *J. Environ. Psychol.*, vol. 16, no. 2, pp. 109–121, Jun. 1996.

[68] Y. Ikeda and M. Inoue, "An evacuation route planning for safety route guidance system after natural disaster using multi-objective genetic algorithm," *Proc. Comput. Sci.*, vol. 96, pp. 1323–1331, Jan. 2016.

[69] L. Chittaro and D. Nadalutti, "A mobile RFID-based system for supporting evacuation of buildings," in *Proc. 2nd Int. Workshop Mobile Inf. Technol. Emergency Response*, Bonn, Germany, Jan. 2009, pp. 22–31.

[70] Y. Inoue, A. Sashima, T. Ikeda, and K. Kurumatani, "Indoor emergency evacuation service on autonomous navigation system using mobile phone," in *Proc. 2nd Int. Symp. Universal Commun.*, Dec. 2008, pp. 79–85.

[71] L.-W. Chen, J.-H. Cheng, and Y.-C. Tseng, "Optimal path planning with spatial–temporal mobility modeling for individual-based emergency guiding," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 45, no. 12, pp. 1491–1501, Dec. 2015.

[72] L. Chu, "A RFID-based hybrid building fire evacuation system on mobile phone," in *Proc. 6th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, Oct. 2010, pp. 155–158.

[73] L. Chu and S.-J. Wu, "A real-time fire evacuation system with cloud computing," *J. Converg. Inf. Technol.*, vol. 7, no. 7, pp. 208–215, Apr. 2012.

[74] Y. Iizuka and K. Iizuka, "Disaster evacuation assistance system based on multi-agent cooperation," in *Proc. 48th Hawaii Int. Conf. Syst. Sci.*, Jan. 2015, pp. 173–181.

[75] M. Alam, K. A. Shakil, and S. Sethi, "Analysis and clustering of workload in Google cluster trace based on resource usage," in *Proc. IEEE Int. Conf. Comput. Sci. Eng. (CSE) IEEE Int. Conf. Embedded Ubiquitous Comput. (EUC) 15th Int. Symp. Distrib. Comput. Appl. Bus. Eng. (DCABES)*, Aug. 2016, pp. 740–747.

[76] J. Gao, H. Wang, and H. Shen, "Task failure prediction in cloud data centers using deep learning," *IEEE Trans. Services Comput.*, vol. 15, no. 3, pp. 1411–1422, May 2022.

[77] M. S. Jassas and Q. H. Mahmoud, "Failure characterization and prediction of scheduling jobs in Google cluster traces," in *Proc. IEEE 10th GCC Conf. Exhib. (GCC)*, Apr. 2019, pp. 1–7.

[78] S. Analytics, "Strategy analytics: Internet of Things now numbers 22 billion devices but where is the revenue," *Retrieved August*, vol. 17, p. 2020, 2019.

[79] T. Muhammed, R. Mehmood, A. Albeshri, and I. Katib, "UbeHealth: A personalized ubiquitous cloud and edge-enabled networked healthcare system for smart cities," *IEEE Access*, vol. 6, pp. 32258–32285, 2018.

[80] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep learning face attributes in the wild," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 3730–3738.

[81] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database forstudying face recognition in unconstrained environments," in *Proc. Workshop Faces 'Real-Life'Images: Detection, Alignment, Recognit.*, 2008.

[82] W. Sultani, C. Chen, and M. Shah, "Real-world anomaly detection in surveillance videos," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6479–6488.

[83] L. Mahadevan, "Anomaly detectionincrowdedscenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Sep. 2010.

[84] N. Papanikolopoulos and V. Morellas, "Unusual crowd activity dataset of university of Minesota," Tech. Rep., 2015.

[85] R. E. Kalman, "A new approach to linear filtering and prediction problems," Tech. Rep., 1960.

[86] P. M. Djuric, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Miguez, "Particle filtering," *IEEE Signal Process. Mag.*, vol. 20, no. 5, pp. 19–38, Sep. 2003.

TARIK ALAFIF received the master's degree from Gannon University, Erie, PA, USA, in 2011, and the Ph.D. degree in computer science from Wayne State University, Detroit, MI, USA, in 2017. Currently, he is an Associate Professor with the Computer Science Department, Jamoum University College, Umm Al-Qura University. His primary research interests include computer vision, machine learning, deep learning, large-scale data, and generative AI. He has published numerous peer-reviewed papers in leading journals and conferences within these domains.

MOHAMMAD JASSAS received the master's and Ph.D. degrees in software engineering from Ontario Tech University, Canada, in 2022. He is currently an Assistant Professor with the Computer Science Department, Umm Al-Qura University. His research centers on developing intelligent software systems, strongly emphasizing cloud computing applications, and big data analytics. He has a distinguished publication record in leading scientific journals. Through his research endeavors, he aims to contribute to the advancement of innovative technological solutions that effectively address the rapid advancements within the field of computer science.

ALAA E. ABDEL-HAKIM received the bachelor's degree (Hons.) and the master's degree in electrical engineering from Assiut University, Egypt, in 1996 and 2000, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Louisville, Louisville, KY, USA, in 2007. He has received numerous academic honors, including a research assistance fellowship from the CVIP Laboratory, University of Louisville, from 2002 to 2007, and a postdoctoral fellowship from the same laboratory, in 2007. He was an Assistant Professor with the School of Technology, Michigan Technological University, and a Postdoctoral Fellow with Microsoft Research, Advanced Technology Laboratory, Cairo office (ATL-Cairo), from 2010 to 2012. Currently, he is a Professor of computer engineering with Assiut University and Umm Al-Qura University, Saudi Arabia. His research interests include deep learning, object recognition, machine vision, and image processing. His work has been published in numerous high-impact journals and conferences. He has been serving as a reviewer for several conferences and journals for several publishers, e.g., IEEE, Elsevier, and IET.

GHADA ALFATTNI received the Ph.D. degree in computer science from Manchester University, U.K., in 2021. She is currently an Assistant Professor with the Computer Science Department, Jamoum University College, Umm Al-Qura University. Her current research interests include natural language processing, text mining, and semiautomated curation of knowledge from unstructured textual data. Her projects emphasize large-scale extraction and curation of biomedical information and clinical/epidemiological findings, leveraging rule-based, data-driven, and deep learning approaches, with a particular focus on the innovative applications of large language models.

HASSAN ALTHOBAITI received the Doctor of Science degree in information technology from Towson University, MD, USA, in December 2018. He is currently an Assistant Professor with the Computer Science Department, Jamoum University College, Umm Al-Qura University. His current research interests include error-free computing, human motion classification by micro-Doppler radar, computer vision, image processing, pattern classification, object tracking, digital signal processing algorithms and implementations, parallel processing, and computer architectures.

MOHAMMED IKRAM received the master's degree from the University of Technology Sydney, Australia, in 2015, and the Ph.D. degree from the Computer Science Department, University of Technology Sydney, in 2020. He is currently an Assistant Professor with the School of Computer Science, Al-Jamoum Collage, Umm Al-Qura University. His research interests include cloud computing, the Internet of Things, networking systems, and data science. He published several publications in these fields.

AMIRAH ALHARBI received the master's degree (Hons.) and the Ph.D. degree in artificial intelligence from Leeds University, U.K., in 2015 and 2020, respectively. She is currently an Associate Professor with the Department of Computer Science and Artificial Intelligence, Umm Al-Qura University. She managed and developed several projects on improving Hajj and Umrah services using advanced technologies, such as AI, the IoT, and video behavior analysis along with sentiment analysis and face expression recognition. She has published a number of research papers in journals and international conferences.

HUSSAM ALSHARIF received the B.Sc. degree from Pittsburg State University, USA, in 2010, the M.Sc. degree from Western Illinois University, USA, in 2012, and the Ph.D. degree from N. C. A&T State University, USA, in 2020. He is currently an Assistance Professor with the Computer Science Department, University College of Al Jamoum (JUC), Umm Al-Qura University (UQU), Makkah, Saudi Arabia. His research interests include medical image processing, data science, big data, data mining, machine learning, deep learning, computational biology, and bioinformatics. He is a reviewer of the number of international journals.

MAZIN ALSHAMRANI received the Ph.D. degree from the University of Surrey, U.K., in 2015, and the Executive Master of Business Administration (E.M.B.A.) degree from the Prince Mohammed Bin Salman College (MBSC), in 2023. He is currently an Associate Professor of computer engineering with the Scientific Information and Services Department, Hajj and Umrah Institute, Umm Al-Qura University, Makkah, Saudi Arabia. He has published numerous research papers in journals and international conferences. His research focuses on advanced technologies, artificial intelligence, deep learning, and smart city solutions for healthcare and Hajj and Umrah applications. He serves as an active reviewer for several international journals.

EBTISAM ALHARBI received the master's degree in information management, in 2018, and the Ph.D. degree in data management from Manchester University, U.K., in 2022. She was with eScience Lab, a multidisciplinary research group focused on the development of tools and practices that facilitate the management and analysis of complex scientific data. She is currently an Assistant Professor with the Department of Data Science, Umm Al-Qura University. She participates in some EU research projects, such as FAIRplus and ELIXIR. She has published a number of research papers in high-impact journals and international conferences, e.g., Nature. Her research interests focus on implementing FAIR data principles in pharmaceutical research and development.

TAHANI ALSUBAIT (Senior Member, IEEE) received the Ph.D. degree from the University of Manchester, U.K.

She is currently an Associate Professor of computer science and artificial intelligence with Umm Al-Qura University (UQU), Makkah, Saudi Arabia. After that, she was appointed to many administrative positions. She supervised over ten postgraduate students. Her research interests include research areas related to artificial intelligence and data analytics, and their applications in a wide range of fields such as education, medicine, finance, and others. She is an Editor-in-Chief of *Engineering Journal* at UQU.

ABDULLAH ALHAWSAWI received the bachelor's and master's degrees in computer science from Oklahoma City University, USA, and the Ph.D. degree in crowd dynamics and simulation from The University of Melbourne, Australia. He is currently an Assistant Professor with Umm Al-Qura University. He is the Head of the Department of Information and Scientific Services, Custodian of the Two Holy Mosques Institute, Hajj and Umrah Research Center, Not only that but also, he, used to be, the Vice Dean for Development and Quality in the Custodian of the Two Holy mosques Institute, Hajj and Umrah Research Center, Umm Al-Qura University, and the Head of the Crowd Unit, Umm Al-Qura University, Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, and the Deputy Director of the Transportation and Crowd Management Center, Umm Al-Qura University.

BADR ALSOLAMI is currently an Associate Professor with the Faculty of Islamic Architecture, Umm Al-Qura University. He is with the Ministry of Hajj and Umrah as a Consultant for international pilgrims affairs. He has over 20 years of experience in Hajj context. His research interests include crowd management and human behavior.

KHALID KHAYYAT received the B.Sc. degree from Umm Al-Qura University, Saudi Arabia, the master's degree from Colorado State University, USA, and the Ph.D. degree in electrical and computer engineering from the University of Victoria, Canada. He was the Vice Dean of the College of Computer and Information Systems for Educational Affairs and Development, from 2013 to 2017, and was also the Deputy Director of the Office of Management and Vision Realization of Initiatives for Executive Affairs, Umm Al-Qura University, from 2018 to 2020. He has been an Assistant Professor with Umm Al-Qura University, since 2011. His research interests include computer network performance modeling, wireless ad hoc networks, wireless sensor networks, wireless applications, the Internet of Things (IoT), and artificial intelligence.