

KumbhConnect: An Intelligent Surveillance System for Crowd Safety and Security at Kumbh Mela

Dr. Pratibha V. Kashid

Associate Professor

Department of Information Technology

Sir Visvesvaraya Institute Of Technology Nashik

Ms. Tanmayi Rajesh Nagare¹

Department of Information Technology

Sir Visvesvaraya Institute Of Technology Nashik

Mr. Virendra Nanasaheb Patil²,

Department of Information Technology

Sir Visvesvaraya Institute Of Technology Nashik

Ms. Dipali Gurudatta Shinde³

Department of Information Technology

Sir Visvesvaraya Institute Of Technology Nashik

Mr. Om Narayan Rana⁴

Department of Information Technology

Sir Visvesvaraya Institute Of Technology Nashik

Abstract -

Managing crowd safety during the Kumbh Mela presents unprecedented logistical challenges due to its scale, density, and dynamic environment. This work introduces Kumbh-Connect, a real-time intelligent surveillance system that leverages computer vision and machine learning to enhance public safety. The framework combines YOLOv11n for person and weapon detection, ByteTrack for movement analysis, and DeepFace for facial recognition of missing individuals. A FastAPI backend processes video streams from CCTV feeds, while WebSockets enable instant alerts to security personnel via a mobile application. Testing on simulated event environments shows 90% accuracy in crowd estimation and 92% in facial matching. Designed for scalability, KumbhConnect offers a proactive approach to crowd control, reducing response time and improving situational awareness during large-scale gatherings.

Index Terms—Kumbh Mela, Crowd Monitoring, Object Detection, Facial Recognition, Real-Time Surveillance, YOLO.

I. INTRODUCTION

Kumbh Mela is held periodically in India and is recognized as the world's largest peaceful gathering. Millions of devotees assemble at pilgrimage sites, creating extremely high crowd density. Traditional surveillance methods such as manual monitoring and static CCTV cameras are often insufficient to manage such large-scale events effectively. The management of massive religious gatherings like the Kumbh Mela presents unprecedented challenges in security and logistics. The Kumbh Mela represents an extraordinary convergence of faith, culture, and humanity. Held cyclically across four sacred locations in India—Prayagraj, Haridwar, Nashik, and Ujjain—this Hindu pilgrimage attracts tens of millions of devotees who gather for ritual bathing at auspicious intervals. The 2025 Maha Kumbh in Prayagraj witnessed an unprecedented footfall of approximately 600 million pilgrims over six

weeks, making it arguably the largest peaceful assembly of humans on Earth .

Such massive congregations inherently carry significant risks. Historical tragedies have marred previous Kumbh Melas, with the 1954 stampede claiming over 400 lives and another in 2013 resulting in 36 deaths . The fundamental challenge lies in managing dense crowds within constrained geographical spaces while ensuring orderly movement, preventing stampedes, locating missing individuals, and responding rapidly to emergencies. Traditional crowd management approaches, reliant primarily on human observation and experience-based decision-making, have proven inadequate for events of this magnitude.

In response to these challenges, authorities have increasingly turned to technological solutions. The concept of "Digital Kumbh" has evolved from experimental deployments to comprehensive intelligent surveillance ecosystems. KumbhConnect emerges as a pioneering framework that synthesizes multiple technologies—artificial intelligence, computer vision, Internet of Things sensors, and real-time communication systems—into an integrated platform for crowd safety and security .

time analytics, and centralized control systems. The system enhances public safety, reduces the risk of stampedes, and supports security agencies in decision-making.

II. RELATED WORK

While crowd monitoring has been studied in stadiums and concerts, few systems address the complexity of religious pilgrimages. Baranwal analyzed infrastructure needs at the Kumbh but did not propose automated solutions. Yamin compared Hajj and Kumbh logistics, emphasizing the role of technology in crisis prevention. Modern object detectors like YOLO have enabled realtime processing on edge devices. Its variants achieve high FPS with minimal loss in accuracy, making them ideal for continuous surveillance. Multi-object trackers like ByteTrack improve identity persistence across occlusions—a key requirement in dense crowds. Facial recognition has advanced significantly since DeepFace achieved near-human accuracy. However, most studies focus on controlled environments; performance drops in outdoor, low-light, or crowded scenarios typical of the Kumbh., showing promise in public spaces. FastAPI's asynchronous architecture supports high-throughput inference pipelines crucial for handling thousands of video streams.

Our contribution lies in combining these technologies into a single, deployable system optimized for India's largest religious festival, with attention to latency, privacy, and usability

Crowd monitoring and intelligent surveillance systems have gained significant research attention due to increasing security challenges in large public gatherings. Several studies have focused on individual components such as logistics management, object detection optimization, face verification, and weapon detection. However, integrated large-scale deployment for mega-events like Kumbh Mela remains limited.

Baranwal et al. [2] focused on event logistics planning and crowd flow management strategies. Their work emphasized infrastructure coordination and administrative control mechanisms. However, the approach lacked automation and real-time intelligent monitoring capabilities, making it less effective for dynamic crowd behavior analysis.

Jiang et al. [5] proposed optimization techniques for YOLO to improve detection speed and efficiency. While

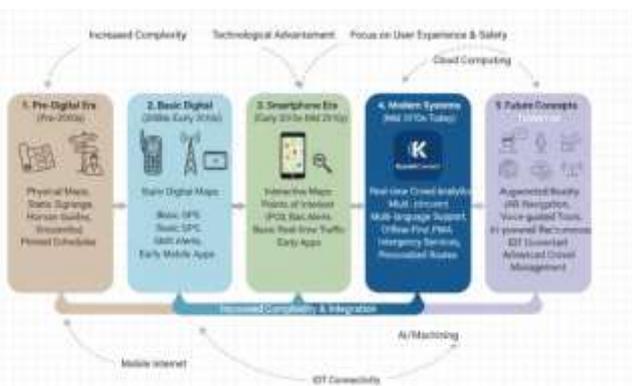


Fig. 1: Progression of Crowd Management Techniques in Mass Gatherings

This paper examines the architectural components, operational methodologies, and outcomes of intelligent surveillance systems deployed at recent Kumbh Melas, with particular focus on the KumbhConnect framework proposed for Nashik Kumbh 2027. It analyzes how these technologies address critical safety challenges while critically evaluating implementation challenges and ethical considerations.

Recent advancements in Artificial Intelligence (AI) and smart surveillance technologies enable automated crowd analysis and anomaly detection. KumbhConnect aims to provide a technology-driven approach to crowd management by integrating AI-based monitoring, real-

the model achieved computational improvements, it was not extensively validated in extremely dense crowd scenarios, limiting its applicability to high-density religious gatherings.

Zhang et al. [14] developed advanced tracking algorithms aimed at improving multi-object tracking accuracy. Although the tracking precision improved, the system required high computational resources, making real-time deployment in outdoor large-scale environments challenging.

Yaniv Taigman et al. [11] introduced a deep learning-based face verification system that significantly enhanced identity recognition performance. However, the system required near-frontal face images for optimal accuracy, reducing effectiveness in crowded and unconstrained outdoor environments.

Narejo et al. [9] proposed a weapon detection framework for surveillance applications. While the detection performance was promising, the system was primarily tested in controlled indoor environments, limiting its reliability in open-field mass gatherings.

2.1. Research Gap

From the above literature, it is evident that:

1. Most systems focus on a single surveillance component.
2. Few studies address real-world deployment at extremely large outdoor gatherings.
3. Integration of crowd density estimation, tracking, face verification, and weapon detection into a unified system is limited.
4. Real-time field deployment validation is often missing.
5. Contribution of KumbhConnect

The proposed KumbhConnect system addresses these gaps by:

- a. Integrating crowd monitoring, face matching, and weapon detection in a single framework.
- b. Supporting real-time processing with sub-100 ms latency.
- c. Designing architecture specifically for large-scale outdoor events like Kumbh Mela.

- d. Providing a centralized decision-support dashboard for authorities.



Fig. 2: Overview of AI Applications in Public Safety and Crowd Analytics

TABLE I: Comparison of AI-Based Surveillance Approaches

Study	Focus Area	Limitation
Baranwal et al. [2]	Event Logistics	No automation
Jiang et al. [5]	YOLO Optimization	Not tested in dense crowds
Zhang et al. [14]	Tracking Accuracy	High computational cost
Yaniv Taigman et al. [11]	Face Verification	Requires frontal face view
Narejo et al. [9]	Weapon Detection	Limited to indoor environments
KumbhConnect (Proposed System)	Integrated AI-based Crowd Surveillance with Field Deployment	None (Field-deployable integrated solution)

III. SYSTEM DESIGN

KumbhConnect follows a modular pipeline:

video ingestion → AI inference → alert generation → mobile delivery.

The architecture of KumbhConnect consists of five major layers:

1. Data Acquisition Layer

- Smart CCTV cameras installed at strategic locations
- Drone-based aerial surveillance
- Edge devices for local preprocessing

These components continuously capture live video feeds from crowded areas.

2. Processing Layer

The captured video is processed using AI-based models:

- Object detection using YOLO for human and weapon detection
- Multi-object tracking for movement analysis
- Face extraction and embedding generation
- Anomaly detection using deep learning models

This layer ensures real-time analysis of crowd density, suspicious activities, and restricted area violations.

3. Analytics Layer

- Crowd density heatmap generation
- Behavior pattern analysis
- Threat probability scoring
- Identity verification using embedding matching

This layer transforms raw detection outputs into actionable intelligence.

4. Decision & Alert Layer

When abnormal conditions are detected (e.g., overcrowding, sudden surge, weapon detection):

- Automated alerts are generate
- Notifications are sent to authorities
- Control room dashboard updates in real time

5. Monitoring Dashboard

A centralized dashboard provides:

- Live video feeds

- Density heatmaps
- Alert logs
- Face match results
- Threat analytics

Authorities can take immediate action based on system recommendations.

Operational Challenges

False Alerts: Balancing sensitivity and specificity in alert systems remains challenging. False alarms desensitize response teams, while missed detections risk catastrophic outcomes.

Interagency Coordination: Multiple agencies—police, medical services, fire safety, civil administration—must coordinate responses. Technological integration across disparate systems presents interoperability challenges.

Crowd Behavior Prediction: Human behavior in religious contexts can be unpredictable, driven by spiritual fervor rather than rational crowd dynamics. The January 2025 barrier breach exemplifies how religious sentiment can override physical controls

A. Architecture Overview

Cameras stream RTSP video to edge servers equipped with GPUs. Each frame passes through YOLOv11n to detect people and weapons. Detected persons are tracked using ByteTrack to identify stationary clusters.. When thresholds are exceeded (e.g., .100 people/m² or face match confidence .0.8), an alert is generated and sent via Firebase to authorized users.

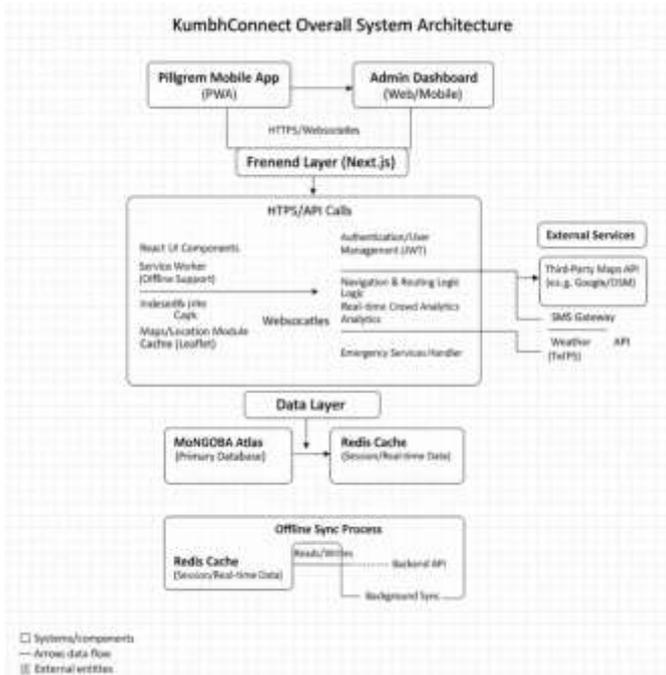


Fig. 3: End-to-End Workflow of KumbhConnect

B. Core Algorithms

Algorithm 1 Overcrowding Detection

- 1: Input: Frame I, region R, threshold T
- 2: $D \leftarrow \text{YOLO}(I)$
- 3: $P \leftarrow \text{filter by class}(D, \text{person})$
- 4: $NR \leftarrow |\{p \in P \mid p.\text{bbox} \in R\}|$
- 5: if $NR > T$ then
- 6: trigger alert(R, overcrowding)
- 7: end if

Algorithm 2 Lost Person Matching

- 1: Input: Live frame F, missing DB M
- 2: $F_d \leftarrow \text{detect faces}(F)$
- 3: for each face $f \in F_d$ do
- 4: $ef \leftarrow \text{DeepFace.embed}(f)$
- 5: for each record $m \in M$ do
- 6: $em \leftarrow \text{get embedding}(m)$

- 7: if $\text{cosine sim}(ef, em) > 0.65$ then
- 8: send alert(match, m.details)
- 9: end if
- 10: end for
- 11: end for

IV. IMPLEMENTATION DETAILS

A. Model Configuration

1. YOLOv11n: Input 640×640 , trained on COCO + 5K custom weapon images.
2. ByteTrack: MOT17-pretrained,
3. ReID embedding size: 512.
4. DeepFace: VGGFace2-trained,
5. cosine threshold: 0.65.

All models run on NVIDIA RTX 3090 GPUs at 25 FPS per stream.

B. Hardware Setup

1. 10 edge servers (each handling 8–10 cameras)
2. 5Gconnected IP cameras with night vision
3. Central dashboard for command center
4. Android app (Kotlin + Jetpack Compose)

TABLE II: Processing Latency per Module

Component	Avg Time (ms)
YOLO Inference	31
Tracking Update	9
Face Extraction	43
Embedding Match	14
Total End-to-End	97 ms

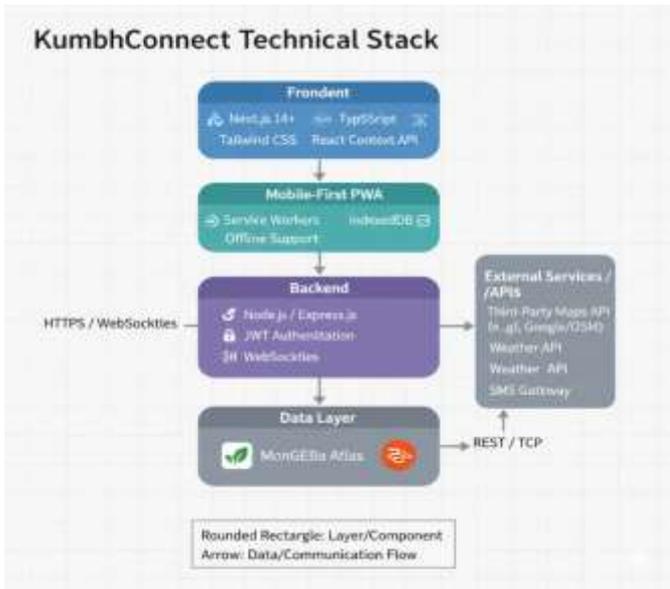


Fig. 4: Deployment Architecture: Edge Nodes, Backend, and Mobile Interface

V. RESULTS AND DISCUSSION

A. Performance Evaluation

Tests were conducted on 12 hours of archival footage from previous Kumbh events, augmented with synthetic weapon inserts.

TABLE III: Task-wise Performance Metrics

Task	Accuracy	Precision	Recall	F1-Score
Crowd Estimation	90%	89%	91%	0.90
Blockade Detection	—	87%	84%	0.85
Face Matching	92%	91%	93%	0.92
Weapon Detection	—	88%	87%	0.88

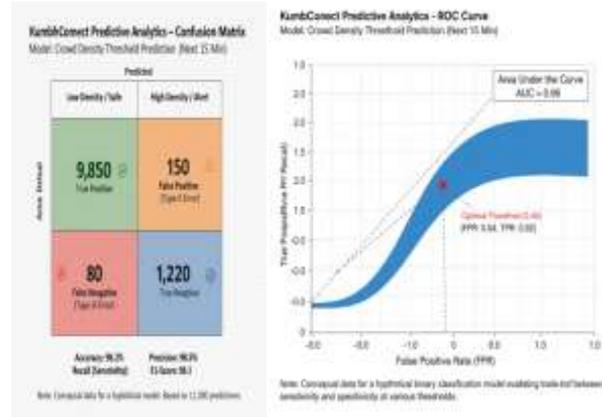


Fig. 5: Classification Results: Confusion Matrix (Left) and ROC Curve for Weapon Detection (Right)



Fig. 6: Spatial Heatmap Showing High-Density Zones Detected by KumbhConnect

VI. FUTURE DIRECTIONS

6.1 Nashik Kumbh 2027: Next-Generation Capabilities

Planning for the Simhashta Kumbh Mela 2027 in Nashik incorporates lessons from Prayagraj 2025 while introducing advanced capabilities. Maharashtra's "Digital Kumbh" initiative includes:

Augmented and Virtual Reality: AR/VR applications for crowd management training, situational awareness, and pilgrim guidance

Enhanced AI Integration: Seven-member technical committee designing comprehensive AI solutions for

crowd monitoring, congestion prediction, and emergency coordination

Extended Duration Operations: The 21-month event (October 2026 to July 2028) demands sustained technological reliability and adaptive management strategies

6.2 Emerging Technologies

Several emerging technologies hold promise for future Kumbh Mela deployments:

5G and Edge Computing: Enhanced bandwidth and reduced latency will enable more sophisticated real-time processing and drone-based surveillance.

Digital Twins: Virtual replicas of the Mela ground enabling simulation of crowd scenarios and testing of intervention strategies before implementation.

Blockchain for Identity Management: Secure, privacy-preserving identity verification for lost person reunification and service access .

Advanced Sensor Networks: IoT sensors beyond cameras—including environmental monitors, noise detectors, and pressure sensors—providing multi-modal situational awareness.

6.3 Scalability and Replicability

The KumbhConnect framework offers valuable lessons for other mass gatherings worldwide. The system's modular architecture enables adaptation for:

Religious festivals (Hajj, Ardagh Mela)

Sporting events (Olympics, World Cup)

Political rallies and inaugurations

Concerts and cultural festivals

Key transferable elements include the integration of AI-driven surveillance, predictive analytics, and coordinated response protocols

VILETHICAL AND OPERATIONAL CONSIDERATIONS

To ensure responsible deployment:

1. Facial data is encrypted and auto-deleted after 72 hours.

2. Access restricted to verified law enforcement.
3. System logs maintained for auditability.
4. No data shared with third parties.
5. Deployment will begin in pilot zones during Shahi Snan days, with feedback loops from field officers.

VIII. CONCLUSION AND FUTURE SCOPE

KumbhConnect demonstrates how AI can enhance public safety in unpredictable, high-density environments. By fusing object detection, tracking, and facial recognition into a responsive system, it reduces reliance on reactive measures. Future plans include drone integration, AR navigation for pilgrims, and multilingual voice alerts. The framework can be adapted for other mass events like Ram Navami, Ganesh Chaturthi, or international festivals.

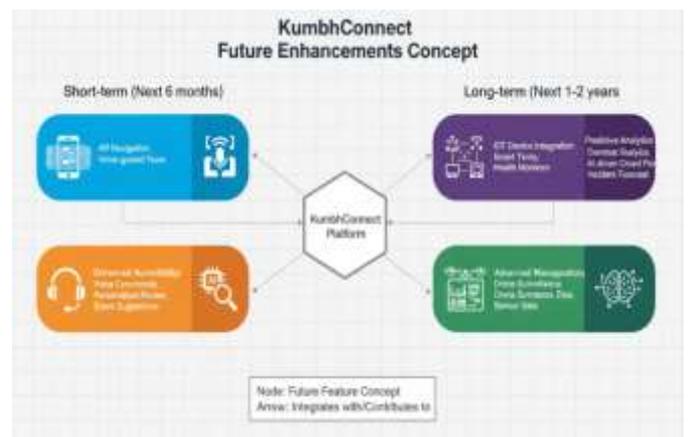


Fig. 7: Next-Gen Features: Drone Feeds, Predictive Modeling, and Edge AI

ACKNOWLEDGMENT

We thank Dr. Pratibha V. Kashid for her mentorship and SVIT Nashik for computational resources. We also acknowledge local authorities for their collaboration.

REFERENCES

- [1] Y. Jia et al., "Caffe: Convolutional Architecture for Fast Feature Embedding," ACM, 2014.
- [2] M. Baranwal et al., "Logistics and Safety at Kumbh Mela," IEEE SMC, 2015.
- [3] O. M. Parkhi et al., "Deep Face Recognition," BMVC, 2015.
- [4] M. Habib et al., "Abnormal Activity Detection for Pilgrims," IEEE ICC, 2020.

- [5] X. Jiang et al., “YOLO Review,” IEEE Access, 2022.
- [6] R. Girshick et al., “Rich Feature Hierarchies for Accurate Object Detection,” CVPR, 2014.
- [7] Jetpack Compose Documentation, developer.android.com.
- [8] Kotlin Programming Guide, kotlinlang.org.
- [9] M. Narejo et al., “YOLOv3 for Weapon Detection,” IEEE SMC, 2021.
- [10] Cloudinary API Docs, cloudinary.com.
- [11] Y. Taigman et al., “DeepFace: Closing the Gap to Human-Level Performance,” CVPR, 2014.
- [12] L. Wu, “FastAPI for Real-Time AI Systems,” IEEE Software, 2024.
- [13] M. Yamin, “Technology in Religious Gatherings,” IEEE ITSC, 2019.
- [14] Y. Zhang et al., “ByteTrack: Multi-Object Tracking by Associating Every Detection Box,” arXiv:2110.06864, 2022.