International Journal of Leading Research Publication (IJLRP)

E-ISSN: 2582-8010 ● Website: www.ijlrp.com ● Email: editor@ijlrp.com

"A Review on Real-Time Crowd Monitoring for Emergency Response and Public Safety Using Deep Learning"

Rupali D. Lodha[1], Dr.Mahesh R. Sanghavi[2]

[1]ME Computer Engineering, SNJB's KBJ COE Chandwad, Nashik, Maharashtra,423101
[2]Professor, SNJB's KBJ COE Chandwad, Nashik, Maharashtra,423101

Abstract:
Crowd monitoring is all about keeping people safe whenever a large number of people gather in one place, like during concerts, sports events, festivals, or in busy public areas like train stations. If big groups of people are not properly managed, it can cause dangerous situations such as stampedes, accidents, or even crimes that go unnoticed in the crowd. Modern crowd monitoring systems act like smart guardians. They keep an eye on the movement of people in real time, using cameras and sensors. These systems don't just count how many people are there—they also study how people are moving. For example, if people suddenly start running or pushing in one direction, the system can quickly identify that something might be wrong. Apart from watching the crowd as a whole, these systems can also look for specific individuals who may need help. For example, someone might collapse due to a medical emergency, or someone may behave in a suspicious way. The system can send alerts immediately so that security or medical staff can act quickly. To make this monitoring accurate and fast, especially on small and portable devices (like cameras with built-in processors), researchers use advanced computer models. These models are designed to be lightweight (so they don't slow down devices) and smart. By combining deep learning with attention-based techniques, the system becomes better at spotting unusual activities in a crowd.

Keywords: Crowd Monitoring, Public Safety, Real-Time Surveillance, Crowd Behaviour Analysis, Anomaly Detection, Edge Computing.

I. INTRODUCTION

Crowd monitoring is becoming more and more important as cities keep growing and large events are held frequently. Today, we see huge gatherings at concerts, sports matches, political rallies, festivals, and religious events. While these events bring people together, they also come with risks. A crowd that is not managed well can quickly become unsafe. History has shown us many examples where stampedes, fires, or sudden panic have caused serious injuries and even loss of life. Because of this, protecting security for the public in crowded regions is extremely difficult.

Modern technology is playing an important part in solving this issue. In the past, police and security officers had to keep an eye on a crowd just by using their eyes and experience.Humans can miss things even if they stay highly alert, particularly when thousands of people are moving simultaneously.This is where artificial intelligence(AI),cameras, and sensors come in helpful.These devices can watch the population continuously without growing tired and they are faster than people at identifying changes or problems. For example,the system will send an alarm in just a few seconds if someone collapses,a fire starts, or people begin to run all of a sudden. This quick action can stop accidents from developing into disasters.

The way that crowd monitoring operates has been totally modified in recent years by modern innovations like computer vision and deep learning. These intelligent technologies are able to "understand" what is happening in the situation rather than simply recording videos. They are able to track thousands of people simultaneously, odd behaviour, and access how quickly people are moving. Unlike traditional monitoring, which required humans to check each camera feed manually, AI can process massive amounts of data from many cameras and sensors together, and do it much faster with better accuracy.

The main purpose of this review paper is to bring together all the important research that has been done in this area. It will explain the different methods used to monitor crowds, discuss what works well and what problems still exist, and highlight the challenges that need to be solved in the future. By doing this, the paper aims to provide a clear and simple picture of how intelligent crowd monitoring has developed so far, and what steps can make it even better in keeping people safe.

II. BACKGROUND & MOTIVATION

Traditionally, crowd management relied on human supervisors watching CCTV cameras or physically present officers guiding people. While being helpful, these techniques are limited by the human attention duration, tiredness and reaction time. Hundreds of people can not be watched in real time by a single officer, and an important warning indicator can go missed. Sensors based techniques have also been applied, including pressure sensors, RFID tags, and motion detectors.These can detect movement or judge crowd density, but they frequently don't provide accurate information on individuals behaviour. They could not work well for large events like marathons or festivals, and they are extremely costly to use in vast ares. The development of computer vision has opened up new possibilities. In addition to counting people, AI-powered systems are able to analyse their movements and interactions. For example, a sudden stop in motion, odd running, or a person falling can all be automatically identified. This gives authorities useful information that is not available through traditional methods. Intelligent crowd monitoring is driven by the straightforward goals of reducing dangers, speeding up response time, and promoting safer public spaces. It is feasible to plan for issues before they become problems by integrating cameras, sensors and AI models.

Logesh Rajendran and R. Shyam Shankaran," Big Data Enabled Real-Time Crowd Surveillance Using AI and Deep Learning(2021) " Approximately 70% of incidents in India have occurred at huge crowds, such as religious celebrations, which have resulted in multiple horrible incidents. This paper explains how L&T Smart World used Artificial Intelligence during the Kumbh Mela 2019 to handle the world's largest religious gathering. The AI system used deep learning to count people in real-time, monitor how they were moving, and guide the flow of the crowd. Over 23 crore visitors in 50 days were managed safely. This project proves that AI can reduce risks and help authorities make quick decisions, especially in events with massive crowds.

Human Detection in Crowded Scenes Using Hybrid ResNet (2024) – Bandar M. Alghamdi. Detecting people in very crowded places is difficult because people often block each other from view. This paper presents a Hybrid ResNet model that is specially designed for this problem. It uses smart modules like cascade fusion (to spot small or hidden people) and activation modules (to improve accuracy). The system was tested on the CrowdHuman dataset and worked very well, even when people's postures changed or when they overlapped each other. The best part is that it's fast enough for real-time use, making it practical for real-world applications.

Wireless Sensor Network for Crowd Disaster Mitigation (2013) – Maneesha V. Ramesh et al., Stampedes and crowd disasters often happen suddenly. This paper suggests using smartphones as

sensors to prevent such tragedies. Phones can use their GPS, accelerometers, and light sensors to detect unusual movements (like sudden pushing or rushing). If danger is detected, the system immediately alerts a control center. This is a low-cost and easy-to-use solution, since almost everyone carries a smartphone. It shows how crowds themselves can provide useful data to predict and prevent disasters.
Real-time Face Detection and Tracking using PTZ Camera (2009) – Ajmal Mian, Normal CCTV cameras often fail to capture clear faces in crowds because they can't zoom or move properly. This paper introduced a system using a PTZ (Pan-Tilt-Zoom) camera, which can automatically move, tilt, and zoom to follow faces in real time. It was designed to overcome delays and motion blur, and tests showed that it could capture clear and accurate facial images even when people were moving. This was an early step towards smarter surveillance cameras

Smart Parking Navigator (IoT Fog-based) (2023) – Emma Qumsiyeh, Isam Ishaq. Finding a parking spot in crowded cities is frustrating. This paper presents a smart parking system that uses a Raspberry Pi camera and image processing to detect empty parking slots. The information is stored in a database and shown to drivers through a mobile app. This saves time, reduces traffic congestion, and is affordable enough for cities to use. It is a practical solution to one of the most common urban problems.
SSD Algorithm in People Flow Monitoring System (2020) – Feng Yizhou et al., Managing people in malls, bus stops, and tourist spots can be tough. This paper introduces a system that uses the SSD (Single Shot Multibox Detector) algorithm to detect people in real time. The system has two parts:

- A tiny computer that is embedded and keeps a watch on people.
- An online program that displays real time data.
- This enables management to monitor crowd levels, identify odd groups, and add quickly to improve visitor comfort and safety.

Husna Sultana et al., Infrared Multifunctional Detection for COVID-19 Protection(2024) Monitoring the health of the public has become an important topic during the COVID-19 pandemic. The IRMDT system, which combines Infrared sensors, deep learning, and computer vision, is presented in this work. It can detect faces, measure body temperature, track social distances, and decide when a person is wearing a mask. It is an effective tool for public health and safety because tests shown that it performed well in crowded situations.

Finding individuals In Real-World Crowd Analysis(SSD+DeepSORT)(2021) by Sachin Sharma and Piyush Juyal The monitoring of individuals in actual public spaces is the main topic of this paper. The system can follow people in video footage and record their current location on a real world map by utilizing SSD (for People Detection) and Deep SORT (for movement tracking). This helps authorities in identifying unusual behaviour, analysing crowd behaviour, and improves city planning and safety.
Qian Wang et al,(2016) Real Time Social Network Data With Differential Privacy. Social media users frequently post where they are, which is helpful for monitoring crowds movements. However, sharing this data openly creates privacy risks. This paper introduces RescueDP, a system that allows real-time crowd data to be shared while protecting personal privacy. It uses differential privacy techniques so that useful statistics (like how many people are in one area) can be shared, but without revealing individuals' exact information.

Detecting Rogue AP with Crowd Wisdom (2017) – Tongqing Zhou et al. In crowded places, people often connect to WiFi. But cybercriminals sometimes create fake WiFi hotspots (rogue APs) to steal user data. This paper proposes a method to detect these fake hotspots using crowdsensed signal data from people's devices. The system analyzes wireless signal patterns and detects fake access points in real time, without needing any special hardware. This improves cybersecurity in public areas.
Face Mask Detection Using NASNetMobile & CNN (2023) – Mohammed AbdulSattar et al. Wearing

face masks was very important during COVID-19. This paper presents a deep learning system that can detect if a person is wearing a mask in real time. Using NASNetMobile and CNN models, the system processes video from webcams and quickly classifies whether people have masks on or not. It achieved almost 99.8% accuracy, making it one of the most effective mask-detection systems during the pandemic.

DyMo: Dynamic Monitoring of LTE-Multicast Systems (2017) – Yigal Bejerano et al. Streaming video to large crowds (like at concerts or stadiums) is difficult because network feedback from users is very limited. This paper presents DyMo, a system that collects small amounts of feedback from users and uses statistical methods to detect signal issues. This allows the network to adjust in real time, ensuring smooth and reliable video streaming for large audiences. It is efficient, accurate, and designed for big events.

III. EXISTING APPROACHES TO CROWD MONITORING

One common way is vision-based monitoring, which uses CCTV cameras or even drone cameras. Special computer programs analyze the video to count people, see how crowded an area is, and study how people are moving. They can detect if a crowd is moving in one direction or if a lot of people are standing in one place. This works very well in open areas where everything will be clearly visible. But the system can not fastly understand the situation when people are closely together or when visibility is poor because of smoke, poor lighting, or bad weather.

Sensor-based monitoring uses devices like RFID tags, mobile phone Wi-Fi signals or pressure sensors on the ground. By using these devices we track how many people are present and how the people move. These works are useful in smaller places like offices and stadiums. It is very expensive and not practical to put sensors everywhere in large public areas like city squares or festivals. Also the issue of privacy of those devices, so that people may not carry devices to track them.

Next are machine learning techniques. These train algorithms to identify specific patterns in crowd behaviour using data gathered from cameras or sensors. For example, they can learn to identify odd behaviour, disaster dangers and restricted areas where people become stopped. Even though this looks promising, particularly huge and complicated datasets are typically difficult for earlier machine learning techniques to handle. When making quick decisions in real time, they can be slow and could not provide correct outcomes.

Lastly, deep learning is the most advanced and strong strategy. This makes use of complex algorithms that can automatically learn from photos and videos, such as CNNs (Convolutional Neural Networks) and RNNs (Recurrent Neural Networks). These models can identify complicated patterns on their own. Such as identifying unusual activity in a crowded area, so they don’t require human programming for every little detail. Attention mechanisms are used by more complex systems to assist the model “focus” just on the most important parts of a situation such as a person accidently collapsing in a crowd of people. By adding these methods into hybrid deep learning models, researchers can achieve high accuracy while protecting system speed for real-time monitoring.

IV. METHODOLOGY AND COMPARATIVE STUDY OF CROWD MONITORING SYSTEMS

1. Vision-Based Methods:

Vision-based technology is one of the most traditional and widely used methods for crowd monitoring. They use drones, CCTV, and Pan-Tilt-Zoom(PTZ) cameras to capture videos of crowded areas. By capturing video and using computer vision algorithms to analyse the footage, the device can “see” how people are moving in real time, identify crowded areas, and monitor crowd behaviour. These algorithms identify people, track movement, and estimate crowd density. For instance, the system could identify crowds in a stadium and point out areas that might be too crowded. Vision-based methods focused on visual features like forms, motion, and edges to explain crowd patterns.

Vision-based systems are lightweight because users do not need to carry an electronic device. They can continuously scan large areas and provide visual proof for further analysis. Installing cameras is often easier and less costly than creating complex sensor networks specially in areas where cameras are already in place. Poor performance in crowded areas, where people may overlap and make detection more difficult, is one of the challenges. Accuracy can be lowered by dim lighting, fog, rain, or darkness. Since videos record people's faces and movements, vision-based technologies may also raise privacy issues that need careful data management. In train stations, airports, shopping centers, and major public gatherings, vision-based techniques are frequently used. By combining AI and deep learning, modern systems increase accuracy and are able to identify odd patterns and motions even in difficult-to-reach places. The majority of the current crowd monitoring systems are built on these techniques.

2. Sensor-Based Methods:

Sensor-based approaches observe crowds using technologies other than cameras. These might be cell phone signals, wearing RFID tags, or floor pressure sensors. These systems use the sensors to identify movement patterns and physical interactions rather than video. Every sensor gathers information on a person's location, speed, or movement. For example, a worn RFID tag can track a person's activity, and a pressure sensor may identify when a big number of individuals walk in a specific region. A central system receives the data and uses pattern analysis to forecast traffic, odd movement, or possible dangers. Even in low-visibility situations like smoke, darkness, or fog, sensor-based systems can function. In compact interior spaces like stadiums or shopping centers, their ability to accurately identify movement and speed is helpful. Because sensors can initiate automated notification, these systems can also notify authorities more quickly than visual inspection alone. Installing sensors in large, open areas can be expansive and difficult. Many sensors require users to carry devices, which creates privacy concerns. Coverage is limited if not all participants have sensors, which may reduce overall effectiveness in public spaces.

3. Traditional Machine Learning Methods:

Current machine learning methods search trends in crowd behaviour by using historical data. They use algorithms such as decision trees, random forests, and support vector machines. These models can detect odd behaviour, identify crowd movement patterns, and alert dangers. All information which is collected from movies, sensors, and other sources is used to train machine learning models. Crowd density, speed, and direction all features are learned to be identified by models. After training, the system can predict unexpected events such as problems or possible stampedes.

All methods are very easy to use and offer useful predictions when data is limited. Also need a smaller amount of processing power than deep learning models and can be used on simple systems for small scale monitoring. Normal machine learning fails to handle large datasets, especially videos from multiple cameras. Handmade characteristics are also required, advanced decisions about the importance of data points to make by humans.

4. Deep Learning Methods:

Deep learning uses networks to automatically extract features from data without requiring input from humans. Recurrent neural networks (RNNs) and convolutional neural networks (CNNs) are most frequently used for crowd monitoring. These models can count crowds, detect people in crowds, track movement, and process images or videos. These algorithms are able to detect sudden increases or unexpected crowd behaviour and provide real-time alerts.

Deep learning models are good for handling complex patterns and larger datasets than normal machine learning. They can identify people in crowded areas, at hidden views, or crowded spaces. Deep learning models are also challenging to use on mobile devices because of their high processing requirements. To achieve high accuracy during training, they also need a large amount of labeled data. Deep learning is used in many modern monitoring systems, such as monitoring public events, stadiums, and metro stations.

5. Attention-Based Models:

Focus-based models expand deep learning techniques. They allow the system to focus on the most important part of an image or video. This is helpful in crowded areas. The model reduces to high-risk areas, such as groups of people moving randomly. By focusing on these important areas, the system may detect problems faster and more accurately than conventional CNNs or RNNs.

Attention processes shorten processing times and improve accuracy. Even in crowded or disorderly crowds, where traditional models might overlook minute changes, they work well. Because of their effectiveness, these models can also be used effectively with edge devices and real-time systems. Considering their strength, attention-based models can still be computationally challenging, and skill is needed to build the attention mechanism. If the camera or sensor cannot see the important places, they can also have trouble. Events with large crowds, including music festivals or religious gatherings, employ attention-based techniques. These technologies improve safety by alerting authorities about important movement paths so they can take action before accidents or stampedes happen.

6. Hybrid Models:

Hybrid models capture the advantages of several approaches by combining them. To focus on key areas, a hybrid system might combine sensors for more data, CNNs for detection, RNNs for tracking movement, and attention methods. The system combines the processing of sensor and camera data. Attention modules concentrate on high-risk areas, CNNs identify individuals, and RNNs examine sequences. Sensor data can provide alarms with more confirmation and assist in confirming anomalies.

Hybrid models offer accuracy, speed, and efficiency. They can work in real time, handle complex situations, and identify odd behaviour in crowded areas. These systems are more secure because they do not depend on a single data source. Installing multiple sensors and connecting them with video feeds can be expensive. Training these algorithms also requires huge and varied datasets. Hybrid systems work best for sporting events, and the Kumbh Mela. They can count crowds, keep an eye on movements, spot odd activity, and can alert authorities. These models are most advanced and reliable for modern crowd safety systems.

Table 1: Comparative Study

Methods	**Strengths**	**Limitations**	**Real-world Application**
Vision-Based	Non-intrusive; can monitor large areas; continuous 24/7	Struggles in dense crowds; poor lighting or weather affects	Train stations, airports, malls,

	monitoring; provides visual proof	accuracy; privacy concerns	stadiums, and large events
Sensor-Based	Works in low visibility; detects movement and speed accurately; fast alerts	Expensive for large areas; requires people to carry devices; coverage may be limited	Malls, airports, stadiums, and controlled events
Traditional Machine Learning	Easier to implement; works with small datasets; requires less computing power	Struggles with large, complex datasets; needs handcrafted features; slower for real-time	Small-scale monitoring; early research systems; anomaly detection in limited spaces
Deep Learning	High accuracy; handles large datasets; automatic feature learning; detects dense crowds	Requires powerful hardware; needs large training data; computationally intensive	Stadiums, metro stations, public events, and real-time crowd analysis
Attention-Based	Focuses on critical areas; improves accuracy; efficient for dense crowds; faster than standard deep learning	Still computationally intensive; requires expertise in design; may fail if critical areas are not visible	High-density events, festivals, concerts, religious gatherings
Hybrid Models	Combines multiple techniques; highly accurate; robust and efficient; handles complex scenarios	Complex setup; expensive; requires large datasets and computational resources	Large-scale events (Kumbh Mela, Hajj, major sports events); city-wide crowd monitoring; integrated AI systems

V. KEY RESEARCH

There is a need for datasets like collections of photos, videos, or sensor readings that match actual crowds to develop and test crowd monitoring systems. Ai systems learn from these datasets in the same way that students do from textbooks and real-world examples. Video recordings of moving people describe the scene’s events. In training systems identify various crowd activity types. There are thousands of photos of crowds of various sizes from small groups to large gatherings. It is a special dataset as well. It helps teach AI how to calculate crowd density, or the number of people in a specific area.

The UCF-QNRF dataset is very large and complex. There are many complex scenes with many people packed into one area. This dataset is especially helpful to evaluate how well a system performs in dynamic settings with thousands of people moving at once, like festivals or rallies. The Mall dataset, which includes video footage shot inside a mall, is another option. This dataset is useful for training systems because malls are common places where people gather. While these datasets are very useful, they also come with some limitations. Most of them are recorded in controlled or predictable settings. In real life, crowds are often more complicated. For example, poor lighting at night, heavy rain, extreme

weather, or cultural events with different types of clothing and behavior can confuse AI systems. Models trained only on these "clean" datasets may not perform well in such unpredictable real-world scenarios.
Another big limitation is that most existing datasets only focus on visual data (like images and videos). They do not include information from other sources, such as IoT sensors, mobile signals, or wearable devices. If we had datasets that combined both visual and sensor data, systems would become much more reliable because they could cross-check information. For example, cameras could show movement, while sensors could confirm the number of people or their locations. This is why researchers argue that we need more diverse and multimodal datasets in the future. By including different environments, conditions, and data types, AI systems will become smarter and more adaptable, making them much better at handling real-world crowd monitoring challenges.

VI. NUMBER OF PUBLICATIONS ON CROWD MONITORING OVER THE YEARS

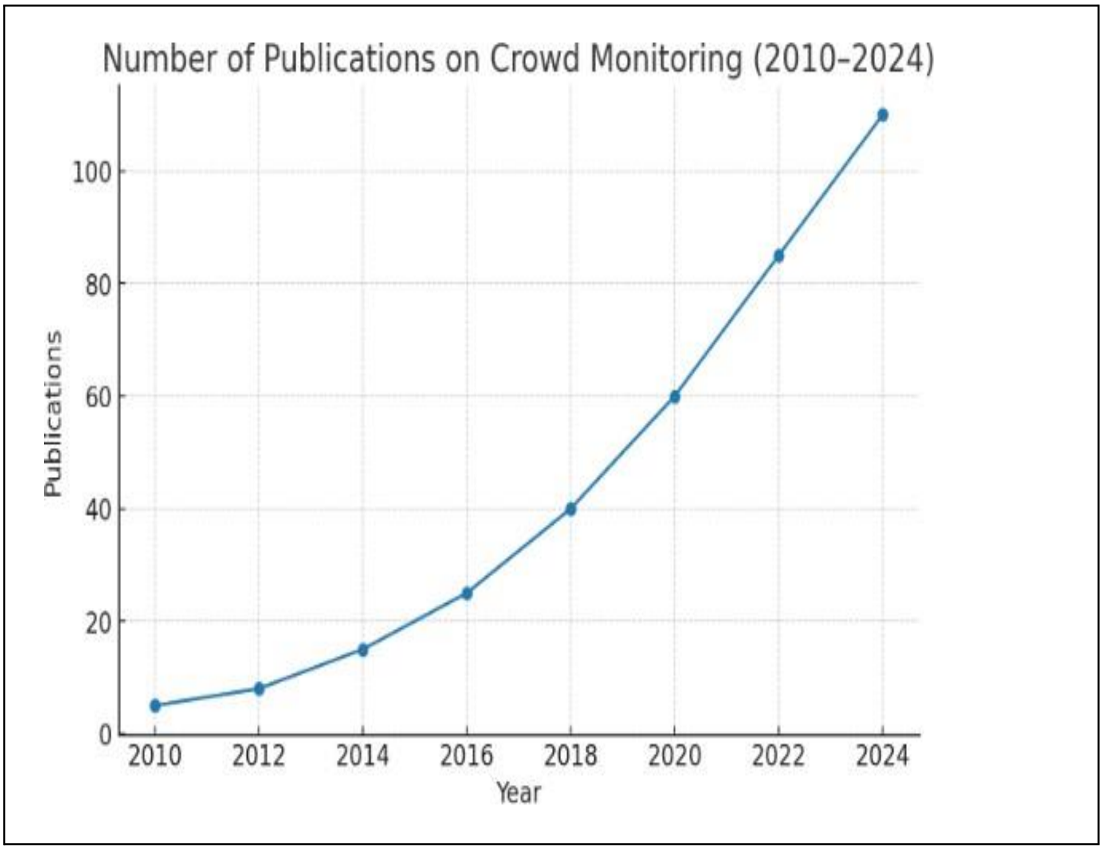

Figure 1: Publications on Crowd Monitoring

This graph tells the story of how interest in crowd monitoring research has grown over the years. There weren't many studies published on this subject in the early 2010s. Only a few researchers were investigating the use of technology for understanding and managing crowds at that time. However, as the years went by, an increasing number of researchers began to focus on this field. The growth became visible around 2016. Additionally, this was the time when deep learning, machine learning, and artificial intelligence began to gain ground. Researchers were able to approach crowd monitoring in new ways because of these new tools, which improved the analysis of massive volumes of video and sensor data.
The amount of scientific effort increased significantly after 2020. The annual number of papers published has rapidly increased. The COVID-19 epidemic, which brought attention to how crucial it is to monitor and control individuals in public areas for their health and safety, may be one cause of this.

The quick growth of smart cities, where crowd control is increasingly crucial, published in a single year. This is a clear indication that crowd monitoring has grown in popularity and importance as a research area. The consistent upward trend indicates that it is an issue of long-term significance rather than being simply a passing fad. All things considered, this graph indicates that crowd monitoring is no longer simply a concept for experiments. It is presently seen as a feasible solution to issues that arise in the real world, such as city planning, public safety, event management, and disaster response. In the upcoming years, we could see much more research as technology advances .

VII. COMPARISON OF METHODS USED IN CROWD MONITORING

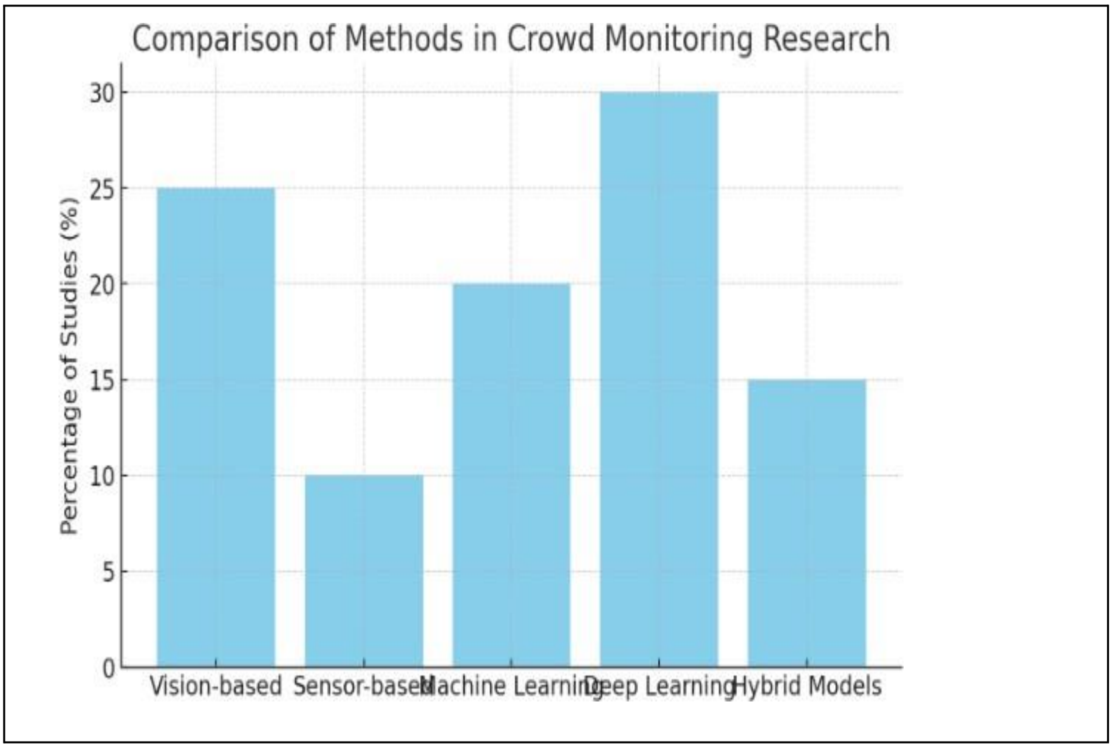

Figure 2: Comparison of Methods

The primary techniques used by researchers for crowd monitoring are contrasted in this graph. While many strategies have been explored throughout the years, some have gained more traction than others. Using cameras and computer vision to investigate crowds is the first strategy, known as vision-based approaches. When CCTV cameras are so widely available in public spaces. By analysing video footage, these systems are able to detect the number of people present as well as their movements. Because of this, a large percentage of research studies apply vision-based methodologies.

The next point is sensor-based methods, which use devices such as personal RFID tags, motion detectors, and pressure sensors that are based on the ground. These methods may be more effective in smaller, more controlled environments, such as stadiums or shopping malls, but more challenging to implement in large public spaces.

Machine learning techniques are also used in many studies. These involve training models on crowd data to find patterns, when people are moving normally versus when they might be in danger, like a bottleneck or stampede. When working with large and complex data sets, traditional machine learning is useful but has limitations.

In recent years, deep learning has become the most widely used and successful method. Because deep learning can learn directly from raw photos and videos, it is more effective than traditional methods at handling complex and large crowd situations. Also explain why deep learning has generated research in this field and is thought it is a more effective technique in it. Last but not least, hybrid models combine different approaches, like deep learning with sensor data or machine learning and computer vision..

VIII. ACCURACY COMPARISON OF DIFFERENT DEEP LEARNING MODELS

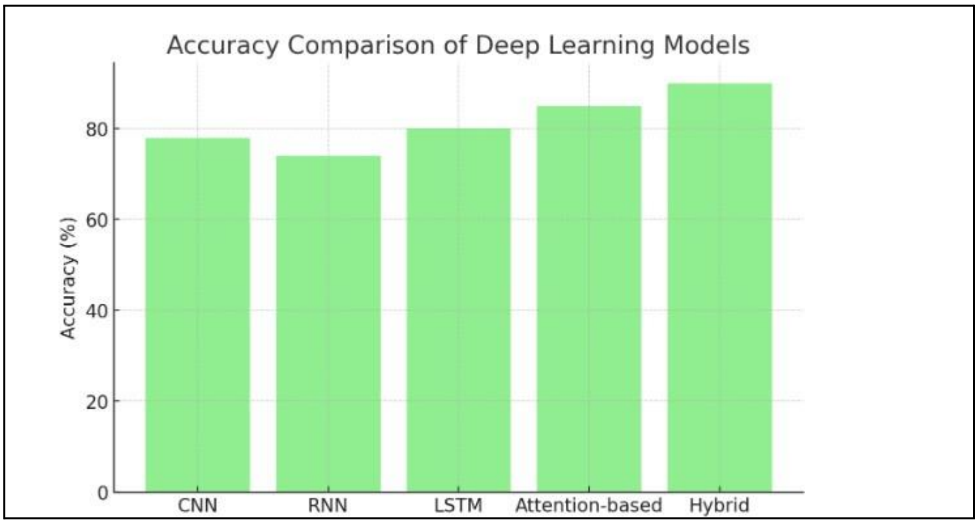

Figure 3: Accuracy Comparison of Mode

The first category includes convolutional neural networks, or CNNs, which are used in image analysis. They are very good at recognizing forms, objects, and movement in video frames. For crowd monitoring, CNNs can calculate crowd density and can identify large flow patterns. They work well, but when it comes to understanding extremely complicated or dynamic actions in big, crowded crowds, their accuracy is rather limited.

Next are recurrent neural networks. They are made to process data that varies over time. RNNs may examine frame sequences to identify odd behaviour patterns, including sudden rushed or panicked motions, because crowds are always moving. Although they work well, they are still not as good as more recent versions. The Long Short-Term Memory (LSTM) model is a more complicated variant of RNNs. LSTMs are more accurate than CNNs and RNNs because they can maintain significant patterns over longer time periods. They can, for example, monitor the slow shifts in crowd behaviour and spot dangers before they become serious.

The attention-based models follow. Because they do not consider every component of the input identically, these are more intelligent. Rather, they “pay attention” to the scene’s most crucial features. For example, in a crowd video, attention-based models concentrate on anomalous motions, such as a person falling or a sudden surge in one region, rather than evaluating each individual equally. Because of this, they are more precise and effective than conventional CNN, RNN, or LSTM models.Lastly, it is demonstrated that the hybrid models get the maximum accuracy. By combining CNNs for image identification, LSTMs for time-based analysis, and attention mechanisms for concentrating on important aspects, these models combine the advantages of several approaches. Hybrid models are more robust and strong since they combine several methods at once. They function better in real-world settings with sizable, erratic, and complex crowds.

Finally, this study indicates that whereas CNNs, RNNs, and LSTMs established the groundwork for deep learning in crowd monitoring, the future lies in attention-based and hybrid models. They are more capable of addressing real-world difficulties, enabling higher precision, and providing better support for safety and security in busy areas.

REFERENCES:

1. L. Rajendran and R. S. Shankaran, "Bigdata Enabled Real-time Crowd Surveillance Using Artificial Intelligence and Deep Learning," *Proc. IEEE*, Mar. 2021.
2. B. M. Alghamdi, "Human Detection in Crowded Scenes Using Hybrid ResNet," *IEEE Access*, Jan. 2024.
3. M. V. Ramesh, S. Anjitha, and P. Rekha, "A Novel Wireless Sensor Network Architecture for Crowd Disaster Mitigation," *Proc. IEEE*, Mar. 2013.
4. A. Mian, "Realtime Face Detection and Tracking Using a Single Pan, Tilt, Zoom Camera," *Proc. IEEE*, Jan. 2009.
5. E. Qumsiyeh and I. Ishaq, "Smart Parking Navigator: An IoT Fog-Based System for Realtime Car Parking Detection," *IEEE Internet of Things Journal*, Jul. 2023.
6. F. Yizhou, C. Junyan, X. Tianmin, C. Rongfeng, L. Xinyu, and T. Yongqing, "Application of the SSD Algorithm in a People Flow Monitoring System," *Proc. IEEE*, Mar. 2020.
7. H. Sultana, Shivani, R. K. Singh, and P. B. R., "The Integrated Infrared Multifunctional Detection and Tracking System for COVID-19 Protection," *Proc. IEEE*, Jul. 2024.
8. P. Juyal and S. Sharma, "Locating People in Real-World for Assisting Crowd Behaviour Analysis Using SSD and Deep SORT Algorithm," *IEEE Access*, May 2021.
9. Q. Wang, Y. Zhang, X. Lu, Z. Wang, Z. Qin, and K. Ren, "Real-Time and Spatio-Temporal Crowd-Sourced Social Network Data Publishing with Differential Privacy," *Proc. IEEE INFOCOM*, Aug. 2016.
10. T. Zhou, Z. Cai, B. Xiao, Y. Chen, and M. Xu, "Detecting Rogue AP with the Crowd Wisdom,"
11. *IEEE Trans. Mobile Computing*, Jul. 2017.
12. M. A. A. Abdullah and L. M. Ibrahim, "Face Mask Detection During COVID-19 Pandemic Using NASNetMobile and CNN Deep Learning Algorithms," *Proc. IEEE*, Dec. 2023.
13. Y. Bejerano, C. Raman, C.-N. Yu, V. Gupta, C. Gutterman, and T. Young, "DyMo: Dynamic Monitoring of Large Scale LTE-Multicast Systems," *IEEE Trans. Mobile Computing*, Oct. 2017.
14. W. Mansouri et al., "Deep Convolutional Neural Network-based Enhanced Crowd Density Monitoring for Intelligent Urban Planning on Smart Cities," *Scientific Reports*, vol. 15, no. 5759, Feb. 2025.
15. G. Gao, "A Survey of Deep Learning Methods for Density Estimation and Crowd Counting,"
16. *Machine Learning and Knowledge Extraction*, vol. 2025-024, Jan. 2025.
17. K. Rezaee, "A Survey on Deep Learning-Based Real-Time Crowd Anomaly Detection,"
18. *Personal and Ubiquitous Computing*, Jan. 2024.
19. A. M. Alasmari, "Recent Trends in Crowd Management Using Deep Learning," *Smart Robotics and Applications*, 2024.
20. M. F. M. Badauraudine et al., "Detection and Tracking of People in a Dense Crowd through Deep Learning Approach—A Systematic Literature Review," *Information Research Communications*, vol. 1, no. 2, pp. 65–73, Jan. 2025.
21. Z. Chen, "Dense-stream YOLOv8n: A Lightweight Framework for Efficient Crowd Monitoring," *Scientific Reports*, 2025.
22. A. Luo et al., "Hybrid Graph Neural Networks for Crowd Counting," *arXiv preprint*, Jan. 2020.
23. M. A. Khan, H. Menouar, and R. Hamila, "Revisiting Crowd Counting: State-of-the-Art, Trends, and Future Perspectives," *arXiv preprint*, Sep. 2022.

24. M. A. Khan et al., “Visual Crowd Analysis: Open Research Problems,” *arXiv preprint*, Aug. 2023.
25. M. H. Sharif, L. Jiao, and C. W. Omlin, “Deep Crowd Anomaly Detection: State-of-the-Art, Challenges, and Future Research Directions,” *arXiv preprint*, Oct. 2022.
26. A. Patwal, “Crowd Counting Analysis Using Deep Learning,” *Procedia Computer Science*, 2023.
27. “Crowd Control and Monitoring Using Deep Learning,” *ResearchGate*, Jan. 12, 2025.
28. “Federated Learning for Crowd Counting in Smart Surveillance Systems,” *IEEE Internet of Things Journal*, 2024.
29. “Joint CNN and Transformer Network via Weakly Supervised Learning for Efficient Crowd Counting,” *ECCV*, 2020s.
30. “CrowdFormer: Weakly-supervised Crowd Counting with Improved Generalizability,” *IJCAI*, 2020s.
31. “Scene-Adaptive Attention Network for Crowd Counting,” *GitHub repository*, 2023.
32. “Crowd Transformer Network,” *GitHub repository*, 2023.
33. “Scale-Aware Attention Network for Crowd Counting,” *GitHub repository*, 2023.