iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: XI Month of publication: November 2025

DOI: https://doi.org/10.22214/ijraset.2025.75607

Advanced Surveillance and Detection Systems Using Deep Learning

M L Sharma[1], Sunil Kumar[2], Soumi Ghosh[3], Samar Alam[4], Sadiya Firdos[5], Kartik Joshi[6]
[1, 2, 3]Faculty, Maharaja Agrasen Institute of Technology, Delhi
[4, 5,6] Research Scholar, Maharaja Agrasen Institute of Technology, Delhi

Abstract: The integration of deep learning technologies into surveillance systems represents a transformative advancement in public safety and security management. This research paper provides a comprehensive analysis of how AI-powered surveillance systems are revolutionizing security infrastructure globally. Through systematic review of empirical evidence spanning four decades, this paper evaluates the effectiveness of systems including CCTV, thermal imaging, facial recognition, and smart policing platforms. The findings indicate that while deep learning-enhanced surveillance demonstrates significant potential in improving security outcomes, deployment must be carefully balanced with privacy concerns, ethical considerations, and governance frameworks. This paper synthesizes insights from over 80 evaluation studies worldwide, providing evidence-based recommendations for policymakers, security professionals, and technology developers.

Keywords: Deep Learning,Convolutional neural networks(CNNs),Recurrent neural networks,smart surveillance ,Edge AI,YOLO (You Only Look Once)

I. INTRODUCTION

A. Background and Context

The landscape of security and surveillance has undergone remarkable transformation over the past two decades, driven by rapid advancements in artificial intelligence, machine learning, and computer vision. Modern surveillance systems have evolved from simple recording devices into sophisticated intelligent platforms capable of real-time analysis, pattern recognition, and predictive decision-making.

Closed circuit television cameras have become omnipresent in modern urban environments. Estimates suggest approximately 245 million surveillance cameras were installed globally by 2014, with continued exponential growth. In the United Kingdom, CCTV systems increased from approximately 100 installations in 1990 to over four million within two decades. In the United States, 49 percent of local police departments report using CCTV, increasing to 87 percent for agencies serving populations exceeding 250,000 residents.

However, camera presence alone does not guarantee security effectiveness. Traditional systems faced significant limitations, primarily their dependence on human operators to monitor feeds, identify threats, and coordinate responses. Human operators are inherently limited by attention span, fatigue, cognitive biases, and the sheer volume of data generated. A single urban surveillance system can generate terabytes of video data daily, making comprehensive human monitoring practically impossible.

Deep learning integration addresses these fundamental limitations. Deep learning algorithms can process vast amounts of video data in real time, detecting anomalies, recognizing faces, identifying objects, tracking movements, and predicting potential security incidents before they occur.

B. Deep Learning Fundamentals in Surveillance

Deep learning architectures suitable for surveillance primarily include Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and hybrid architectures.

CNNs excel at processing visual data through layered structures that automatically learn hierarchical feature representations.

The You Only Look Once (YOLO) family enables real-time object detection by processing images in a single forward pass. YOLO architectures can identify and localize multiple objects within video frames at speeds exceeding 30 frames per second. YOLOv5 has demonstrated exceptional performance in detecting humans in thermal imagery, achieving mean average precision scores exceeding 0.995.

RNNs and variants including Long Short-Term Memory networks excel at processing sequential data, analyzing temporal patterns and tracking object movements across video frames. Their ability to maintain memory of previous states makes them valuable for trajectory prediction, activity recognition, and anomaly detection.

C. Research Objectives and Scope

This paper examines several critical questions: What specific AI surveillance technologies are governments and organizations deploying? How effective are these systems in achieving security objectives? What technical requirements and challenges exist? What ethical considerations and privacy concerns arise? How do different surveillance modalities compare?

The scope encompasses multiple dimensions across diverse geographic contexts, from advanced democracies to emerging economies. Technologically, it covers CCTV enhanced with AI analytics, thermal imaging, facial recognition, smart city infrastructure, and predictive policing. Methodologically, it synthesizes findings from over 80 empirical studies, systematic reviews, and policy analyses covering developments from 2010 to 2025.

II. LITERATURE REVIEW AND THEORETICAL FRAMEWORK

A. Historical Development

Academic surveillance research has evolved considerably over four decades. Early research in the 1980s-1990s focused on social and political implications, examining privacy, civil liberties, and state overreach potential. The first systematic CCTV effectiveness evaluations began in the late 1990s, primarily focused on British deployments using quasi-experimental designs.

Welsh and Farrington's pioneering systematic review examined 22 evaluations, finding modest crime reduction effects, particularly for vehicle crimes. Subsequent updates tracked expanding evidence: the 2009 update included 44 evaluations confirming 16 percent overall crime reduction, driven by 51 percent reduction in car park crimes. The 2019 update synthesized 80 studies worldwide, representing an 82 percent evidence increase enabling more nuanced analysis of moderating factors.

Facial recognition technology emergence in the 2010s sparked new research waves. Early work focused on technical performance; as technology improved and deployments expanded, research increasingly addressed algorithmic bias, discriminatory impacts, and privacy implications. Most recently, research has examined global AI surveillance proliferation. The AI Global Surveillance Index documented AI surveillance in 176 countries, revealing 75 countries actively using these technologies.

B. Theoretical Frameworks

Several frameworks explain how surveillance systems may prevent crime. Situational crime prevention theory posits crime reduction through increasing effort required, increasing detection risks, reducing rewards, removing excuses, and reducing provocations. Surveillance cameras primarily operate through risk increase mechanisms, creating perception among potential offenders that activities will be observed with consequences.Rational choice theory suggests offenders make calculated decisions weighing costs against benefits. Surveillance alters this calculus by increasing perceived costs while potentially decreasing benefits. However, this assumes rationality not characterizing all offenders, particularly those acting impulsively.

The routine activities perspective emphasizes convergence of motivated offenders, suitable targets, and absence of capable guardianship. Surveillance cameras enhance guardianship, creating "eyes on the street" perception. Deep learning makes this guardianship more capable by enabling active threat detection rather than passive recording.

III. SURVEILLANCE TECHNOLOGY TYPES AND IMPLEMENTATIONS

A. Smart City and Safe City Platforms

Smart city platforms represent comprehensive surveillance ecosystems integrating multiple sensor types, data sources, and analytical capabilities into unified management systems. These platforms leverage IoT infrastructure, cloud computing, big data analytics, and AI to create intelligent urban environments capable of monitoring, analyzing, and responding to security threats in real time.

Huawei has been particularly active promoting safe city concepts, marketing comprehensive solutions that "predict, prevent, and reduce crime." Typical deployments include HD video surveillance networks, facial recognition systems, vehicle tracking, incident command centers, big data analytics, mobile integration, and secured cloud infrastructure.

The Kenya Safe City project illustrates this integrated approach. Huawei deployed 1,800 HD cameras and 200 traffic surveillance systems across Nairobi, connected to a national police command center supporting over 9,000 officers and 195 police stations. During Pope Francis's 2015 visit, where eight million people participated in welcome activities, the system demonstrated improved policing efficiency and increased detention rates.

Integration of multiple data sources in smart city platforms creates particular privacy concerns. When camera feeds combine with license plate databases, social media monitoring, financial transaction records, and telecommunications data, the result is extraordinarily detailed pictures of citizens' lives and movements.

B. Facial Recognition Systems

Facial recognition has emerged as one of the most powerful and controversial surveillance tools. Modern systems employ deep learning, particularly CNNs, to create mathematical facial feature representations compared against databases containing millions of identities.

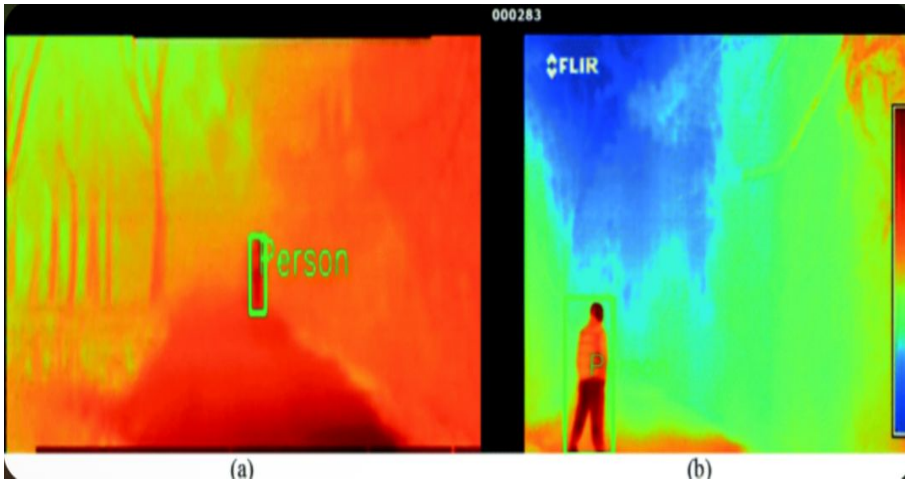
Technical performance has improved dramatically. US National Institute for Standards and Technology evaluations showed facial recognition improved twentyfold between 2014 and 2018, with failure rates dropping from 4.0 percent to 0.2 percent under optimal conditions. Under ideal circumstances, modern systems achieve accuracy exceeding 99.9 percent.

However, performance degrades significantly under real-world conditions. Poor lighting, non-frontal poses, partial occlusion, motion blur, low resolution, and aging substantially reduce accuracy. A UK Metropolitan Police evaluation found an 81 percent false positive rate, meaning 81 percent of system-flagged matches were incorrect.

Demographic bias represents another critical concern. Multiple studies document substantially worse performance on women and minorities compared to white males, arising from biased training data and potential algorithmic biases. When deployed in criminal justice contexts, these biases can perpetuate and amplify existing discriminatory patterns.

C. Thermal Imaging Surveillance

Thermal imaging detects infrared radiation based on object temperature, enabling surveillance where visible light cameras fail. This makes thermal cameras valuable for nighttime surveillance, monitoring through smoke or fog, detecting concealed people, and identifying individuals minimizing visual signatures.



Modern thermal cameras operate in longwave infrared spectrum (8-14 micrometers), corresponding to peak thermal emissions from human body temperature. Advanced systems combine thermal sensors with image processing algorithms to detect human presence at 10-20 meter distances.

Deep learning integration has enhanced thermal imaging capabilities significantly. Researchers successfully applied YOLO architectures to detect humans in thermal imagery, achieving exceptional performance. Training on specialized thermal datasets enables models to reliably identify people regardless of clothing, time, or weather.

Thermal imaging offers advantages over visible light surveillance: works in complete darkness, can detect hiding/camouflaged humans, is less affected by fog or smoke, and provides less personal appearance detail, potentially offering partial privacy advantages.

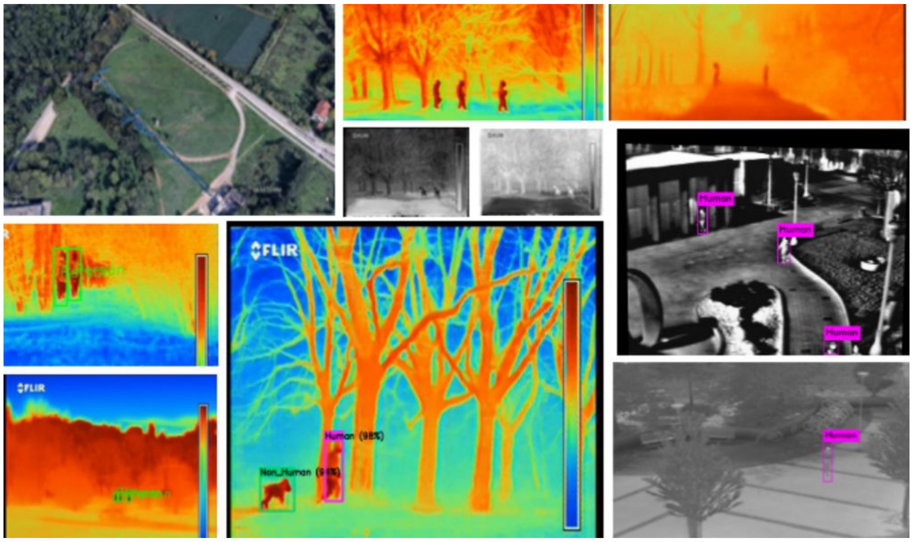
D. Smart Policing and Predictive Analytics

Smart policing leverages big data analytics and machine learning to guide law enforcement resource allocation, predict crime hotspots, identify at-risk individuals, and support investigations. Systems aggregate data from diverse sources including historical crime records, arrests, service calls, social media, license plate readers, and surveillance cameras.

Predictive policing platforms like PredPol claim to predict where crimes will occur with remarkable precision, generating forecasts at 500x500 square foot scales. Algorithms train on 2-5 years of historical crime data, learning spatial and temporal patterns to forecast similar future crimes.

However, predictive policing attracts substantial criticism. A fundamental problem: algorithms trained on historical data perpetuate embedded patterns, including enforcement biases. Neighborhoods subject to intensive past policing show more recorded crimes, leading algorithms to predict more future crime there, justifying continued intensive policing in self-reinforcing feedback loops.

The "dirty data" problem exacerbates concerns. Historical crime datasets reflect not objective reality but enforcement decisions about where to patrol, whom to stop, and what to record. If past enforcement was biased, algorithms reproduce and potentially amplify those biases.



Moreover, effectiveness remains questionable. Multiple studies found modest or no measurable impact when departments adopted these systems. Mechanisms through which predictions supposedly prevent crime remain unclear.

IV. EFFECTIVENESS EVIDENCE FROM SYSTEMATIC REVIEWS

A. *Meta-Analytic Findings on Crime Reduction*

The most comprehensive evidence comes from systematic reviews aggregating multiple studies. The 2019 Welsh and Farrington CCTV review represents the gold standard, synthesizing 80 evaluations using quasi-experimental or experimental designs with before-and-after measurements in treatment and control areas.

Overall meta-analytic findings indicate CCTV is associated with modest but statistically significant crime reduction. Across 76 studies providing sufficient data, the odds ratio was 1.141, indicating approximately 13 percent crime decrease in CCTV areas compared to controls ($p < 0.001$).

However, this overall effect masks substantial heterogeneity. Geographic setting emerges as critical. CCTV achieved largest effects in car parks with 37 percent crime reduction (OR 1.588), benefiting from constrained environments, specific detectable target crimes, high coverage, and active monitoring. Residential areas showed 12 percent reductions (OR 1.133). City and town centers showed no significant overall effect despite being the most common deployment setting.

Crime type significantly influences effectiveness. CCTV showed largest effect on drug crime with 20 percent reduction (OR 1.249), and 14 percent reductions for property and vehicle crime. Violent crime showed no significant effect, possibly reflecting that violent incidents often occur impulsively or privately, or that potential offenders are less deterred.

Operational characteristics matter substantially. Schemes with active monitoring showed 15 percent reductions (OR 1.172), while passively monitored systems showed no significant effect (OR 1.015). This underscores that cameras alone don't prevent crime; effectiveness depends on integration into responsive security operations. Complementary interventions alongside CCTV also matter. Schemes deploying multiple interventions (improved lighting, security personnel, signage, police operations) achieved 34 percent reduction (OR 1.513). Systems with single or no complementary interventions showed no significant effects.

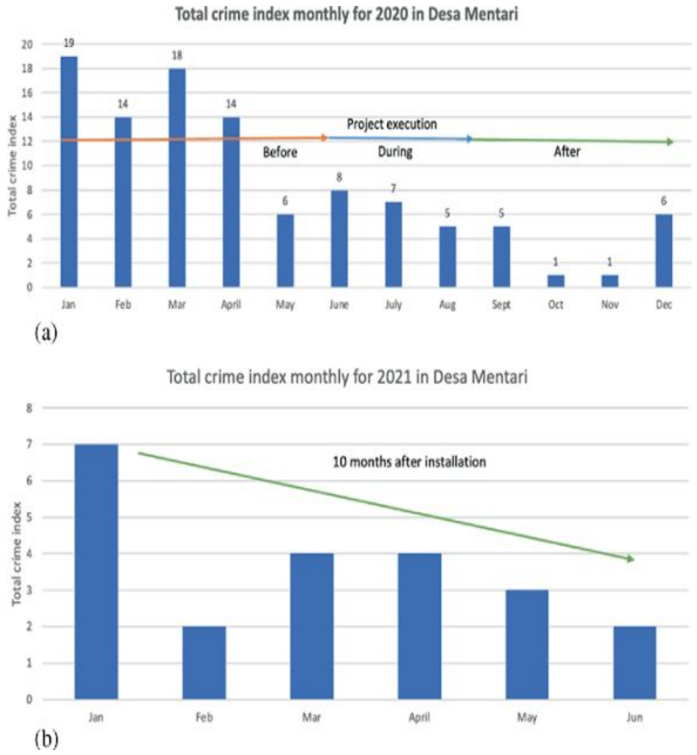
B. *Displacement and Diffusion Effects*

An important consideration is whether crime reductions merely reflect displacement to nearby unsurveilled locations rather than genuine prevention. Among 50 studies examining adjacent buffer areas, only 6 found evidence of crime increases in areas bordering CCTV zones, with 3 finding both displacement and diffusion.

More common is the opposite pattern: diffusion of crime prevention benefits, where crime decreases extend beyond directly surveilled areas. Fifteen studies found such diffusion effects, possibly occurring through offender uncertainty about coverage boundaries, visible cameras signaling general enforcement, or police patrols extending beyond cameras.

C. Investigative and Case Clearance Benefits

Beyond deterrence, surveillance supports investigations by providing identifying evidence. Several studies demonstrate CCTV footage significantly increases case clearance rates. In Milwaukee, overall clearances were 14 percent higher at CCTV-covered locations. In Australian rail networks, clearance rates with CCTV footage were 18 percent higher than matched cases without video.



(a)

(b)

British Transport Police found CCTV captured 45 percent of rail network crimes, with investigators judging footage useful in 65 percent of those cases. Having useful CCTV evidence increased case clearance likelihood from approximately 20 percent to 50 percent.

Real-time monitoring may enable on-scene apprehension. In Newark, incidents detected through active CCTV monitoring resulted in on-scene enforcement at 33.1 percent rates versus 17.0 percent for 911-reported crimes.

D. Governance Quality Role

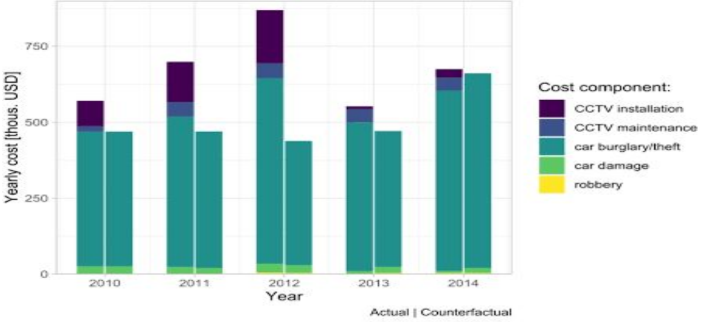
Research on global AI surveillance deployment shows governance quality strongly influences whether technologies are used lawfully or abused for repression. Analysis indicates 51 percent of advanced liberal democracies deployed AI surveillance versus 37 percent of closed autocracies and 41 percent of electoral autocracies. This demonstrates regime type alone doesn't predict adoption; democracies are actually more likely to deploy these technologies.

However, deployment manner differs fundamentally. Democratic countries with strong governance implement surveillance within regulatory frameworks including judicial authorization, independent oversight, data protection standards, public transparency, and abuse remedies. Authoritarian governments frequently use surveillance to suppress opposition, monitor minorities, restrict expression, control civil society, and intimidate dissidents.

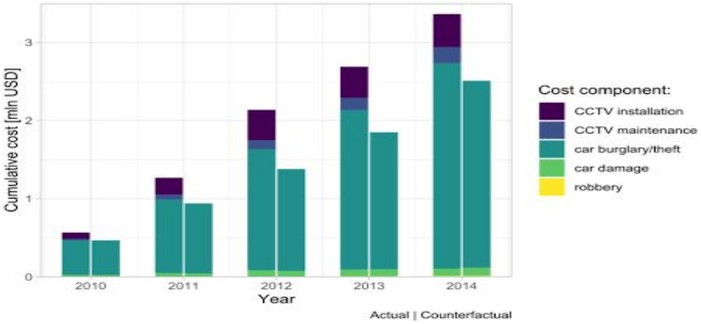
V. CHALLENGES AND LIMITATIONS

A. Technical Challenges

Despite advances, deep learning surveillance faces significant limitations. Accuracy under non-ideal conditions remains challenging. While facial recognition achieves near-perfect accuracy under controlled conditions, performance degrades substantially in real-world messiness with poor lighting, non-frontal poses, occlusion, motion blur, low resolution, and distance.



a) yearly



b) cumulatively

Algorithmic limitations constrain capabilities. While detecting persons or vehicles is straightforward, complex tasks like assessing intent, predicting behavior, or distinguishing normal from suspicious activity remain challenging. The "long tail" problem, where rare high-interest events occur too infrequently for adequate training representation, limits novel threat detection.

Adversarial attacks represent fundamental vulnerabilities. Researchers demonstrated carefully crafted input perturbations, often imperceptible to humans, can cause misclassification or detection failure. Adversaries could potentially exploit these through confusion-designed clothing patterns, recognition-foiling makeup, or false-detection light projections.

Data requirements for training effective models are substantial. Deep learning typically requires thousands or millions of labeled examples. Obtaining sufficient training data for specialized surveillance tasks can be challenging, particularly for rare events. Training data must be representative of deployment conditions; models trained on one population may perform poorly in different contexts.

B. Ethical and Social Challenges

Ethical challenges extend beyond privacy to encompass fairness, accountability, transparency, and social justice. Algorithmic bias creates serious fairness concerns. Facial recognition shows significantly higher error rates for women and darker-skinned people compared to white males. When deployed in criminal justice, these biases perpetuate and amplify existing discriminatory patterns.Lack of transparency in algorithmic decision-making creates accountability challenges. Deep learning models are often opaque "black boxes" providing predictions without clear reasoning explanation. When systems influence consequential security or law enforcement decisions, inability to understand reasoning is problematic.

Informed consent is often absent in public surveillance contexts. Individuals subjected to public space surveillance typically cannot decline or even know surveillance is occurring. The shift from occasional observation to comprehensive persistent tracking fundamentally changes public space nature.

Function creep erodes democratic control. Technology deployed for specific justified purposes may be retained and repurposed for broader application without renewed public debate or explicit legislative authorization.

Chilling effects on free expression and assembly may be the most severe long-term harm. When people know their presence at protests, religious services, medical facilities, or political events may be recorded, tracked, and retained, they may self-censor. This occurs even without actual punishment, through mere surveillance awareness.

C. Practical Implementation Challenges

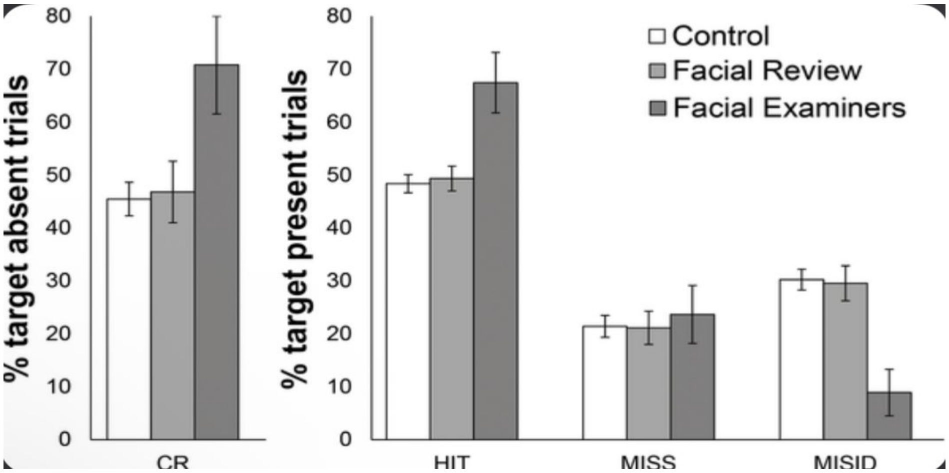
Organizations face numerous practical challenges. Cost represents a fundamental constraint requiring substantial investment in cameras, infrastructure, storage, AI hardware, software licenses, and maintenance. Many jurisdictions lack resources for comprehensive deployments.

Personnel training and capacity building are often inadequate. Operators must understand system capabilities and limitations, know how to interpret alerts and footage, and have clear response procedures. Without proper training, excellent technical systems may deliver poor results.

Maintenance and reliability present ongoing challenges. Equipment can fail due to environmental factors, vandalism, or wear. Networks may be unreliable. Storage systems can reach capacity. Software updates may introduce bugs. Without disciplined maintenance, system performance deteriorates.

Data management at scale poses significant challenges. Modern surveillance generates massive data volumes requiring sophisticated infrastructure for storing, securing, indexing, and retrieving. Many organizations lack technical capacity for effective surveillance data management.

D. Research Limitations



The existing research base has important limitations. Study designs are predominantly quasi-experimental rather than randomized trials,making them more susceptible to bias. Publication bias may skew evidence toward positive findings. Limited implementation detail information hampers understanding of what makes surveillance effective. Short follow-up periods may not capture long-term impacts. Limited evidence exists on newer technologies. Geographic research concentration limits generalizability. Lack of outcome measure standardization complicates synthesis.

VI. FUTURE DIRECTIONS AND RECOMMENDATIONS

A. Technological Developments

Future surveillance trajectory points toward several directions. Continued deep learning improvements promise increased accuracy for challenging tasks. Emerging architectures including transformers and graph neural networks may enable learning from smaller datasets and better generalization. Edge computing and embedded AI will increasingly enable camera-based processing rather than centralized analysis, offering reduced bandwidth requirements, lower latency, enhanced privacy, and continued network interruption operation. Multimodal fusion combining diverse sensor types will create more robust systems leveraging complementary strengths. Federated learning may address privacy concerns by enabling local model training while keeping raw data decentralized. Explainable AI techniques can make models more interpretable and auditable. Privacy-preserving technologies offer potential to reduce surveillance harms while maintaining security benefits.

B. Policy and Governance Recommendations

1) Establish Strong Legal Frameworks: Legislation should mandate necessity and proportionality assessments before deployment, addressing algorithmic bias testing, transparency requirements, and data retention limits.

2) Implement Independent Oversight: Independent oversight bodies with technical expertise and audit authority must be established, promoting transparency and accountability for discriminatory or inaccurate outcomes.

3) Mandate Privacy Impact Assessments: PIAs should be mandatory prior to deployment and regularly updated, systematically evaluating privacy risks and identifying mitigation strategies.

4) Prioritize Complementary Interventions: Policy must shift from viewing CCTV as standalone deterrent. Investment should prioritize active monitoring and multiple complementary proactive policing interventions to maximize crime reduction.

5) Strengthen Public-Private Partnerships: Deep collaboration between AI developers and law enforcement is necessary to create tailored, effective, user-friendly solutions addressing specific operational needs in rights-respecting ways.

6) Focus on Training and Capacity Building: Comprehensive training for law enforcement is essential to ensure effective use, interpretation, and oversight of complex AI tools while adhering to ethical standards.

VII. CONCLUSION

The integration of deep learning into surveillance represents a pivotal moment in public safety, offering unprecedented capabilities to enhance deterrence and drastically improve case clearance. The evidence overwhelmingly supports that AI-enhanced surveillance is not plug-and-play; effectiveness is directly proportional to operational strategy strength and governance framework integrity.

To fully realize this technology's protective potential while safeguarding democratic principles, a global coordinated call to action is required:

1) Invest in explainable AI and edge computing to address transparency and latency concerns

2) Mandate algorithmic bias audits and implement rigorous legal and ethical frameworks matching technological advancement pace

3) Prioritize interagency collaboration and training to ensure effective and responsible use by human operators

By committing to ethical development, proactive deployment, and robust oversight, deep learning can be transformed from a controversial surveillance tool into an indispensable, rights-respecting asset in the fight for a safer society.

REFERENCES

[1] Ashby, M. P. J. (2017). The value of CCTV surveillance cameras as an investigative tool. European Journal on Criminal Policy and Research, 23(3), 441-459.

[2] Caplan, J. M., Kennedy, L. W., & Petrossian, G. (2011). Police-monitored CCTV cameras in Newark, NJ: A quasi-experimental test. Journal of Experimental Criminology, 7(3), 255-274.

[3] Gerell, M. (2016). Hot spot policing with actively monitored CCTV cameras. International Criminal Justice Review, 26, 187-201.

[4] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

[5] Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys, 54(6), 1-35.

[6] Piza, E. L., Welsh, B. C., Farrington, D. P., & Thomas, A. L. (2019). CCTV surveillance for crime prevention: A 40-year systematic review with meta-analysis. Criminology & Public Policy, 18(1), 135-159.

[7] Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. IEEE CVPR, 1701-1708.

[8] Welsh, B. C., & Farrington, D. P. (2009). Public area CCTV and crime prevention: An updated systematic review. Justice Quarterly, 26(4), 716-745.