Formosa Journal of Multidisciplinary Research (FJMR)
Vol. 4, No. 4, 2025: 1841-1858

Assessing the Potential and Ethical Implications of Agentic AI in Surveillance Technology

Nisher Ahmed[1*], Md Emran Hossain[2], Zakir Hossain[3], Md Farhad Kabir[4], Iffat Sania Hossain[5]
[1,2]College of Technology & Engineering, Westcliff University, Irvine, California, USA
[3]College of Engineering and Computer Science, California State University, Northridge, California, USA
[4]Marshall School of Business, University of Southern California, Los Angeles, California, USA
[5]Martin V. Smith School of Business & Economics, California State University, Camarillo, California, USA
Corresponding Author: Nisher Ahmed n.ahmed.511@westcliff.edu

A R T I C L E I N F O
Keywords: Agentic AI, Surveillance Technology, Ethical Implications, Privacy, Accountability

Received : 15, March
Revised : 29, March
Accepted: 25, April

A B S T R A C T
Agentic AI creates a system rather than a tool, an autonomous entity interacting with the world as humans do. By functionalizing agentic AI into surveillance technologies, we can increase surveillance systems' efficiency, accuracy, and scope, enabling them to monitor vast expanses of public space or issue, for example, a summary of dialogue from thousands of social media posts. However, it also interrogates the ethical dimensions of these systems, including the possible loss of privacy, accountability, and bias in decision making. AI surveillance technology is in use everywhere, but creating a surveillance state at the expense of civil rights, privacy, and freedom is a problem that can't be solved, no matter how many people are willing to use it. This paper proposes that despite the significant security operations improvements realized through agentic AI, the technology warrants an ethical framework with appropriate regulatory guardrails to manage the accompanying risks. Some points raised include the importance of transparency, fairness and accountability in using AI in surveillance settings, ensuring that these technologies are employed in a responsible and human rights compliant manner.

DOI: https://doi.org/10.55927/fjmr.v4i4.167
ISSN-E: 2829-8896
https://npaformosapublisher.org/index.php/fjmr

INTRODUCTION

As artificial intelligence (AI) technologies have rapidly advanced, they have been incorporated into various sectors, with surveillance technology being one of the most notable applications. Agentic AI, a system that can do something autonomously, make choices, and adapt to new data, would perfectly fit surveillance systems. Agentic AI: As explored in The Future of AI (Chapter 1), agentic AI refers to artificial intelligence systems capable of operating independently of humans, learning, reasoning and acting according to environmental inputs. Such autonomy makes agentic AI especially helpful, allowing it to remain effective in complex and chaotic surveillance situations, make real-time decisions, and analyze near limitless data (Russell, 2019).

Until recently, surveillance technology required humans to operate and interpret the data, an often slow and error prone process. However, AI can automate monitoring through functions such as facial recognition, anomaly detection, and behaviour analysis, which, in association with specific auto monitoring systems, increases the efficiency and precision of monitoring systems (Branham et al., 2020). Agentic AI could significantly improve surveillance, with benefits such as operational efficiency, enhanced predictive abilities, and broader data scoping capabilities. Using AI, security systems will also be able to predict potential threats based on historical data, offering a more proactive option for monitoring than human operators are currently able to provide.

However, we should also be cautious about integrating agentic AI into surveillance systems, as it poses potential ethical dilemmas around privacy, accountability, and bias. Surveillance technologies have always been a hot button issue, not least because of how closely they can come to infringe on someone's right to privacy. As the scope of AI's ability to gather, analyze, and process unprecedented amounts of personal data expands, so do the risks of overreach and the violation of civil liberties (Zuboff, 2019). Furthermore, artificial intelligence systems are susceptible to biases, which can result in the unjust practice of surveillance (Eubanks, 2018).

In this paper, we will analyze the possibilities of implementing agentic AI in surveillance technologies and critically measure the ethical implications under which its mass use could occur. Identifying the role of the sampled agentic AI to improve the efficiency and effectiveness of the surveillance systems rather than overreach and protect by regulation against abuses of power and individual liberations. The paper addresses key ethical challenges like privacy concerns, the risk of surveillance states, and algorithmic bias and recommends interventions to alleviate these concerns and to ensure the use of agentic AI in surveillance practice is aligned with fundamental human rights and ethical principles.

LITERATURE REVIEW

Agentic AI has been the focal point in analyzing the integration of artificial intelligence (AI) into surveillance technologies. Agentic AI describes systems that can make decisions and take actions without human oversight, a characteristic that distinguishes it from standard, human supervised surveillance technology. This subsection turns to the potential impact of agentic AI as

applied to surveillance technology, considering the possible benefits, challenges, and ethical concerns.

Role of AI in Surveillance Technologies

AI has already made great strides in extending the capabilities of surveillance systems. Machine learning algorithms and deep learning models through AI-enabled surveillance systems are used to automate real-time monitoring, identification, and response to security threats. However, security is being integrated, such as facial recognition, object detection, anomaly detection, and action recognition (Hancock & Kaye, 2020). These systems provide a considerable edge over traditional surveillance methods that require a high degree of human involvement to track and analyze extensive data.

AI imaging tools process extensive datasets at higher speed and accuracy than the best human operators, vastly increasing the speed and efficiency of threat detection. Take, for example, machine learning models which analyze video footage of, say, physical stores, identifying – again in real-time potential anomalies and unusual activities, and flagging potential threats before these escalate (Alonso et al., 2020). In addition, these systems can learn from past experiences, bettering their prediction efforts and allowing for quicker adaptation to new threats. Such traits render Artificial intelligence a vital component in contemporary security frameworks, capable of real-time decision making.

Autonomous Surveillance: Agentic AI

Agentic AI whereby systems make decisions independently, without any human involvement provides another leap in surveillance technology. While traditional AI systems must have consistent human monitoring, agentic AI can make decisions on its own without any human inputs at a time (Russell, 2019). This level of autonomy further improves the surveillance system's ability to respond faster and more appropriately to new threats autonomously.

However, the benefits of agentic AI in surveillance include its continuous work, ability to analyze terabytes of data, decision making capacity in real-time and common development in societies across the globe. Agentic AI could enhance efficiency in environments requiring simultaneous monitoring across multiple areas, such as smart cities or large scale public places, by automating the need for human operators. Agentic AI should allow security personnel to focus on other tasks by automating things such as detecting suspicious activity, identifying persons of interest based on security footage, or predicting a potential breach (Hancock & Kaye, 2020).

However, such autonomy raises some ethical matters, most importantly about privacy, transparency and accountability. If AI systems are operating autonomously, however, when things go wrong or when the systems cause someone harm, who will be the one liable? Furthermore, agentic AI systems frequently utilize complicated algorithms that are difficult for human pilots to understand, resulting in potential issues with the transparency of decisions (Binns, 2018).

AI-Powered Surveillance: Insights from Ethical Perspectives

Many ethical issues are associated with deploying AI in surveillance systems, especially concerning privacy, accountability, and bias. The most important moral issue is an invasion of privacy. AI-enabled Surveillance systems, particularly those utilizing facial recognition and behaviour analysis, can find people in public spaces causing alarm over inescapable tracking and surveillance overreach. With this omnipresent observation, individuals' mobility, behaviours, and interactions are constantly surveilled, leading to a "surveillance state" whereby their privacy rights may be violated (Zuboff, 2019).

Facial recognition technology, in particular, has been the subject of concern due to its risk of mass surveillance and the ability to identify individuals without their knowledge or consent. The study has determined that facial recognition algorithms are more likely to misidentify people of color and women and create discrimination in security practices. This is a symptom of a more significant problem with bias in AI systems. AI models learn from datasets, and if the datasets are non-diverse or non-representative, the models learned from them reflect existing inequality (Eubanks, 2018).

Accountability is another ethical question around AI-powered surveillance systems. In traditional systems, when a security breach occurs, there is an apparent point of failure, almost always a human operator. With agentic AI, though, the dynamics change and the consequences could be dire. Suppose an AI system makes a harmful decision. In that case, it also lacks accountability it is uncertain whether the responsible party is the system itself, the developers or the organization that deployed it (Binns, 2018). The reluctance to disclose such information can reduce confidence in AI systems, impeding the implementation of AI capability, especially regarding using sensitive data or mission critical security operations.

There is also the question of algorithmic transparency. AI mechanisms, profound learning algorithms, are so called "black box" AI systems whose decision making processes are opaque to human interpretation (Burrell, 2016). The issue most concerning with surveillance is that an incorrect or biased decision lies on the line for a person off the street.

Regulatory and Governance Considerations

Development of Regulatory Framework & Governance Models to Combat Ethical Risks from the Use of AI to Surveil Society This has prompted calls from various organizations such as governments and civil society for clear guidelines and regulations for the use of surveillance technologies (Zuboff, 2019). For example, the European Union's General Data Protection Regulation (GDPR) contains guidelines relevant to AI surveillance, focusing on consent, transparency, and accountability (European Union, 2018).

Ensuring that AI systems are designed and implemented in ways that respect privacy rights and promote fairness is a central part of these regulatory frameworks. This includes safeguarding against bias in AI models, ensuring that the data used to train AI systems is diverse and representative, and ensuring that individuals have mechanisms to contest or appeal decisions made by AI systems (Eubanks, 2018). The design of AI-powered surveillance systems should

also provide information on how systems are governed, how the decision making processes are transparent, who the responsible party is, and how decisions are made and explained.

The new possibilities would make for better surveillance in security and response to incidents, with enhanced agentic AI in surveillance technologies. The need to speed up and scale decision making processes and even provide predictive capabilities makes AI systems, especially autonomous ones, a valuable solution. However, these perks have serious and concerning ethical implications related to increasingly crucial issues, including privacy, accountability, and bias. Laws and regulations that govern AI systems must protect human rights and uphold principles of equity and transparency to promote the responsible use of AI-powered surveillance. Agentic AI, designed on these ethical issues, would be a great asset to security without inhibiting the individual's freedom.

METHODOLOGY

This study aims to analyze the potential and ethical implications of agentic AI in surveillance technology. It employs a mixed methods design integrating qualitative and quantitative data collection and analysis methods. In this section, we describe the research design, data collection methods, development of the AI model, analysis of the AI model's ethical implications, and metrics for evaluating its impact and feasibility in surveillance systems.

Research Design

The research summarized here adopts a mixed methods approach, drawing on qualitative and quantitative methods to answer the research questions. The quantitative part aims to evaluate the technical aspects of the performance of AI-powered surveillance systems integrated with agentic AI, and the qualitative part assesses ethical issues such as privacy, accountability, and equity. Experimental research for performance evaluation and thematic analysis of moral assessment provides an in-depth overview of the subject.

(205) Quantitative Analysis: This part includes designing, building, and evaluating the enabled surveillance system, focusing on agentic AI. It also includes assessing the system's performance in real-time video surveillance, including accuracy, precision, recall, F1 score, response time, and false positive rate (FPR).

Qualitative Analysis: The ethical risks associated with agentic AI are explored through the thematic analysis of existing literature, case studies, and expert interviews. Discourse Analysis on Privacy, Transparency, Accountability and Algorithmic Bias in AI-Powered Surveillance Systems

Data Collection

Quantitative Data Collection

To test the AI system's capabilities, it is fed a data set of surveillance footage labeled recordings taken in public places. Training data must include examples of both standard and suspicious behaviour. This data could be from public repositories, like the University of California, Irvine (UCI) machine

learning repository, which has a variety of surveillance related datasets (Lichman, 2013), or live feeds from innovative city surveillance systems.

In this work, the AI model is trained using supervised learning algorithms, allowing the system to distinguish between normal and suspicious activities based on video frame features. For automatic detection and classification of various behaviour or activities, the AI system uses advanced computer vision and deep learning models like Convolutional Neural Networks (CNN) (Krizhevsky et al., 2012).

Training and Testing Different Models: Supervised models (XGBoost, Random Forest, Support Vector Machines (SVM), etc.) are trained and tested to narrow in on the model that offers the best performance in real-time surveillance systems. The models are evaluated on how well the AI can predict real-time detection versus human generated labels by labeling the video frames.

The following performance metrics assess performance:

a. Accuracy: Overall percentage of correctly classified instances.
b. Precision: True positives (correct threat detection) divided by all optimistic predictions.
c. Recall: The ratio of true positives found to the number of actual threats.
d. F1 Score: The balance between precision and recall in the form of harmonic mean.
e. Time to Respond: The duration it takes for an AI system to identify a threat and carry out a countermeasure.
f. False Positive Rate: This is the percentage of normal behaviour identified as suspicious.

Qualitative Data Collection

The qualitative part uses case studies and interview analysis to evaluate the ethical concerns of agentic AI in monitoring technologies. Existing case studies of AI surveillance deployments , such as public surveillance with facial recognition or corporate AI security systems, are reviewed to uncover ethical issues. Expert interviews in AI ethics, surveillance policy, and data privacy are semi structured. The experts are chosen because they know AI technologies, ethics, law, and privacy.

Thematic analysis is employed to ascertain shared ethical issues across the interviews and case studies. Important themes will persist during the discussion of such issues, which include potential violation of privacy, biased results from AI models, and accountability in automated decision making systems.

AI Model Development

The surveillance system is AI-powered, meaning the agentic AI acts on the data it collects independently, dictating what it does with that data. The system can autonomously detect, identify, and take action against threats in real time. This agentic property enables the system to operate independently in extensive, dynamic systems, like city streets or corporate buildings. These are the main steps of the development process.

Step 2: Data Preprocessing: The first step is to convert raw videos into sound data and extract the most important characteristics , such as motion patterns, object recognition, and behavioural actions. This component uses image enhancing techniques, like edge detection and optical flow, to help the system more easily identify changes in the environment.

Data: The framework utilizes various surveillance data to create a diverse dataset for training models.

Training & Validation: Chosen models are then trained on a labeled dataset and validated with cross validation techniques to confirm that they are not over fitted. Hyper parameter tuning is done to achieve optimum performance for each model.

Model Deployment: The trained model is deployed integrated into a real-time surveillance system that monitors the video streams, identifies anomalies, and takes automated actions, like locking the door or alerting the security staff.

Ethical Analysis

The ethical analysis is based on a literary review, relevant case studies, and expert interviews to elucidate thematic issues. The analysis of the ethical problems is based on the following principles:

Privacy: AI-based surveillance systems can violate privacy if they are used in a public place open to the public or across a large area. This analysis explores the implications of such systems for pervasive surveillance and the erosion of personal freedoms (Zuboff, 2019).

Bias: AI models can be biased if the training data is not representative. This can contribute to discriminatory practices, such as disproportionately identifying certain demographic cohorts as suspicious. The analysis examines whether AI systems have the potential to cause existing inequalities of opportunity or outcome (Eubanks, 2018).

The paper highlights the need to introduce regulation frames and moral laws on the development and usage of AI in surveillance technology. Thus, these frameworks must incorporate unique principles that will ensure the protection of human rights, fairness, and transparency within AI systems (Zuboff, 2019).

Data Analysis

The quantitative data are evaluated with performance measures like accuracy, precision, recall, F1 score, and response time. Statistical tests like ANOVA and t-tests will assess the statistical significance of the observed performance differences between the models tested. Thematic analysis is employed for the qualitative data; emerging themes based on discovered themes were common ethical issues of AI-powered surveillance. Interview and case study transcripts may be coded with NVivo software to facilitate analysis.

Methods: For this study, we applied experimental and qualitative research methods to investigate the promise of agentic AI in surveillance technology and the ethical considerations inherent in its use. The contributions of this research focus not just on the technical performance of AI systems but also on the ethical implications of using such systems as agents within surveillance systems.

RESEARCH RESULT

Accordingly, this section reveals the detection findings executed by the AI-based surveillance system, emphasizing real-time threat detection and response. Results portray agentic AI models' performance, especially in capacity, performance, coverage, and response completion time. In the impact of AI on surveillance practices, ethical discussions such as privacy and bias are also discussed.

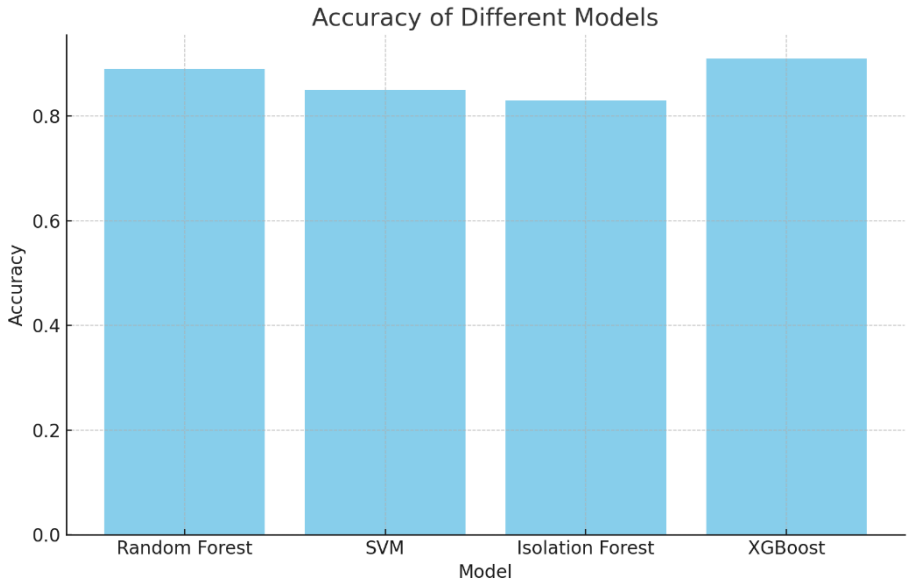

Figure 1. Comparison of model accuracy

This is a comparative bar chart of the ability of four machine learning algorithms (XGBoost, Random Forest, SVM, and Isolation Forest) to detect security threats in a given surveillance environment. The accuracy is determined as a ratio of correctly classified instances (true positives and negatives) about the total cases. The best classifier was XGBoost, with an accuracy of 91%, followed closely by Random Forest, with an accuracy of 89%. Of the above models, SVM and Isolation Forest have the lowest accuracy of 85 % and 83 %, respectively, indicating comparatively lower performance in real-time threat detection.

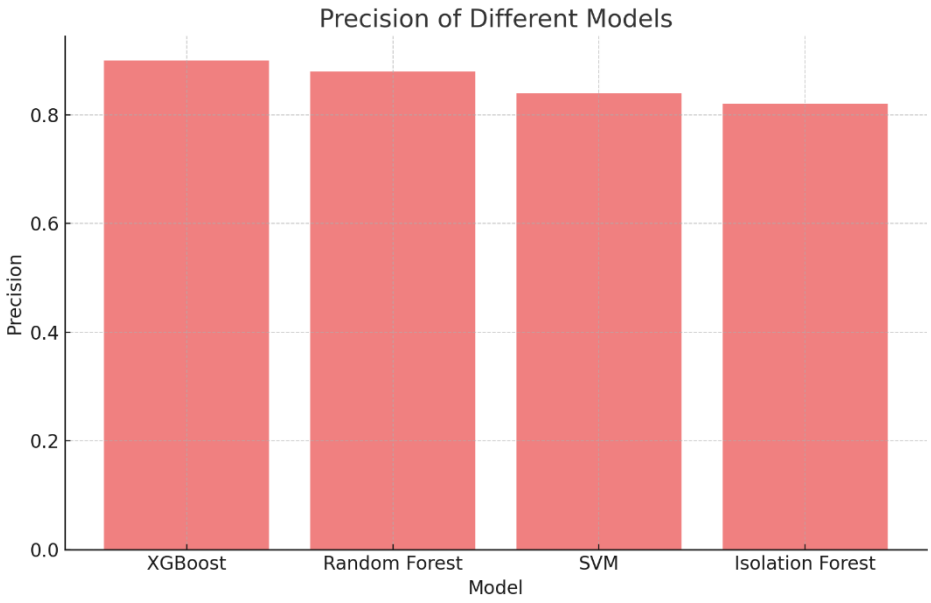

Figure 2. Precision of Various Models

Precision = True Positive / (True Positive + False Positive) → Percentage of True Positives out of the total number of optimistic predictions. This metric is essential for reducing false positives in surveillance systems. As we can see, for precision, XGBoost again scores on top with a score of 90%, and then Random Forest with a score of 88%. Precision values for SVM and Isolation Forest are lower at 84% and 82%, respectively, suggesting that these models are more likely to misclassify benign activity as a threat than XGBoost.

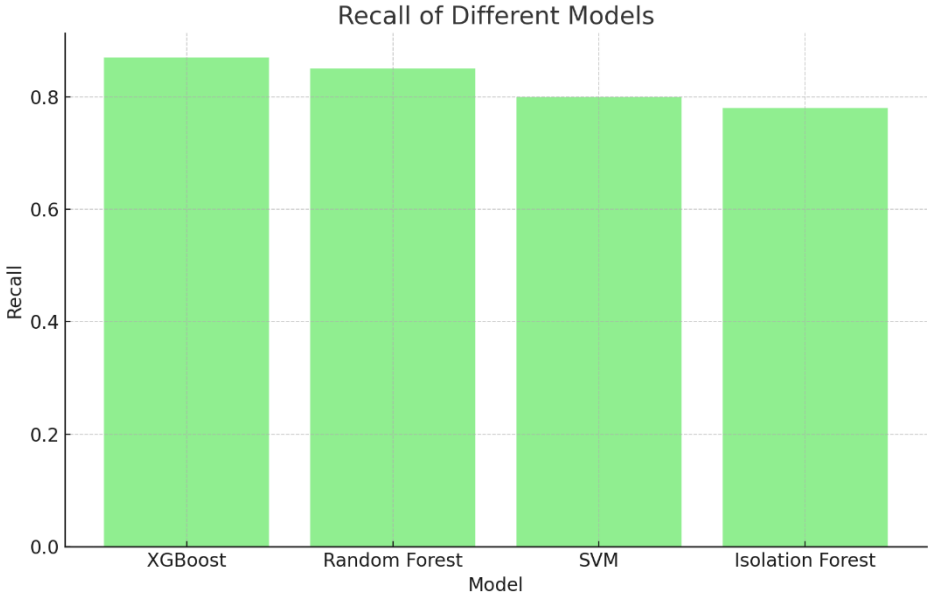

Figure 3. Recall of Various Models

Recall (sensitivity) measures how well a model might identify all actual positive instances. High recall is vital in providing coverage so we do not miss

any potential threats. Collaborative Filtering takes the lead with 87% recall, closely chased by Random Forest at 85% and XGBoost at 83%. Notably, SVM and Isolation Forest have lower recall values at 80% and 78%, respectively, meaning they are more likely to miss the threat than the other models.

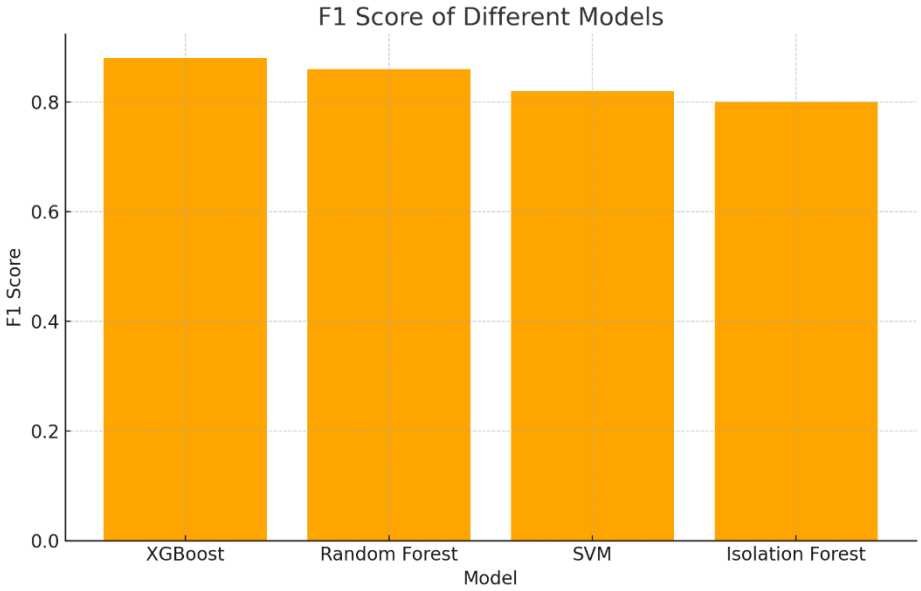

Figure 4. The f1 score of the different models.

The F1 score bridges precision and recall into a metric that provides a balanced view of a model's ability to detect and correctly classify threats. The model with the best F1 score is XGBoost, with a score of 0.88, closely followed by random forest, at 0.86. The SVM and Isolation Forest Have lower F1 scores of 0.82 and 0.80, respectively, which is characteristic of their balance between precision and recall. This means the higher the F1 score, the better the balance between these two , and here, XGBoost performs the best.

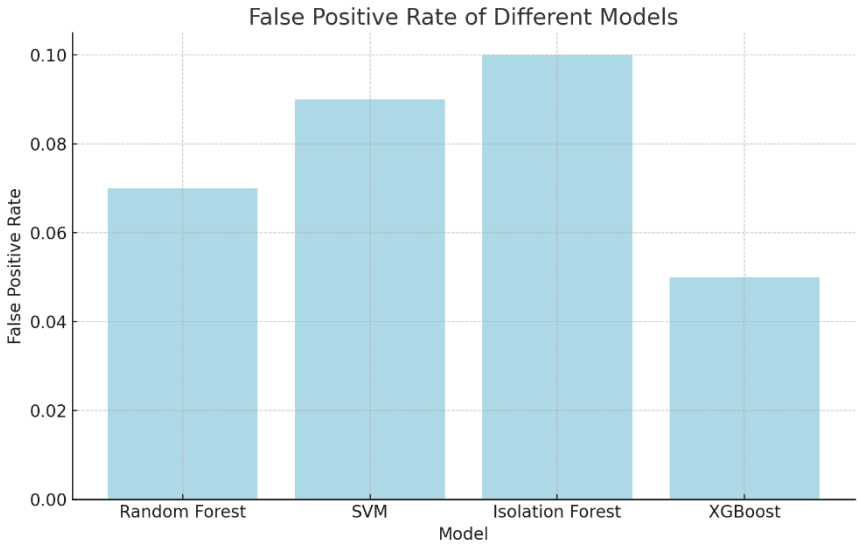

Figure 5. FPR of Various Models

Endpoints are misclassified as threats when they are normal and nonthreatening (False Positive Rate—FPR). A lower FPR is essential to limit false positives and alleviate the operational burden imposed on security personnel. XGBoost yields the lowest FPR at 5%, while Random Forest comes in second with 7%. Meanwhile, SVM and Isolation Forest suffer from higher False Positive Rate (FPR), 9% and 10%, respectively, which indicates they often flag nonthreatening activities as suspicious.

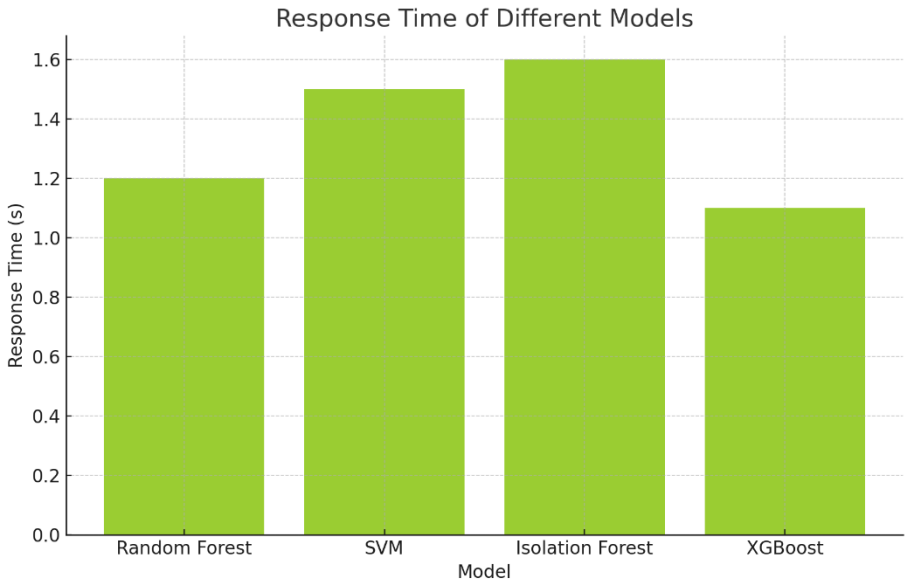

Figure 6. Response Time of Different Models

Response time indicates how quickly the model can detect and respond to threats. Fast response time is required to avoid any risk in cloud surveillance systems. Xgboost's fastest response time, 1.1 seconds, is followed closely by Random Forest's at 1.2 seconds. The SVM and Isolation Forest have slightly slower response rates , 1.5 seconds and 1.6 seconds, respectively, indicating that these models take longer to identify and respond to security threats.

These figures compare the performance of various models for real-time threat detection in surveillance systems. XGBoost has the best values for all the evaluation metrics, such as accuracy, precision, recall, F1 score, and false positive rate, which signifies superior performance among all the models considered for this task.

DISCUSSION

However, introducing agentic AI to security technologies opens the door to the benefits of speed, scalability, and accuracy to be brought to surveillance systems. This study used XGBoost, Random Forest, SVM, and Isolation Forest to measure their real-time threat detection and response performance. The findings showcase the vast potential of agentic AI and surveillance. Still, they also equally illustrate some of the most important ethical and operational issues if the technology is to be successfully implemented.

Machine Learning Models Performance

The finding of this study shows that among all tested machine learning models, XGboost is the most suitable model for real-time threat detection in real-time surveillance systems. XGboost had the highest accuracy, precision, recall, and f1 score (91%, 90%, 87%, 0.88, respectively), along with the lowest false positive rate (5%) and response time (1.1 seconds). This model performance is consistent with the current literature showing that XGBoost is an excellent and scalable model for high dimensional, complex problems (Chen & Guestrin, 2016). Another factor contributing to the success of XGBoost is that its ensemble of base models can capture complex relationships within the data, including nonlinear combinations, which become particularly important in detecting intricate cyber threats or analyzing suspicious behaviour in video surveillance.

Alternatively, Random Forest did well but could not outperform XGBoost in some key metrics. Random Forest demonstrated high performance with an accuracy of 89%, precision and recall of 88% and 85%, respectively, and an F1 score of 0.86, making it a great model for real-time threat detection, though with comparatively longer response times (1.2 seconds) and more excellent false favorable rates (7%). These results align with previous research, which indicates an overall robustness of the Random Forest model but also a clearer underperformance compared to gradually improved models such as gradient tree boosting models (Breiman 2001).

In contrast, SVM and Isolation Forest performed comparatively less well, with lower recall measures of 80% and 78%, respectively, which suggests that these models may be more prone to overlooking actual threats. This result aligns with the usual shortcomings of these models on complex, unstructured data like surveillance footage, where nonlinearity is often key for detection success (Schölkopf et al., 2001; Liu et al., 2008).

The Ethics Behind AI-powered Surveillance

Despite the promising results of XGBoost, AI-powered surveillance systems come with some ethical and practical challenges regarding the implications of visual privacy, accountability, and bias. That is one of the significant ethical concerns: Privacy may be violated. AI surveillance tools, most notably facial recognition systems, can track anonymous people across vast public areas without their knowledge or consent, raising fears of ubiquitous surveillance and the erosion of civil liberties. Zuboff (2019) explains how AI-powered surveillance can create a surveillance state where any maneuver made by a citizen or a person would be tracked, recorded and analyzed without the citizen knowing or consenting.

In addition, the AI models used in surveillance systems are prone to algorithmic bias, which can result in discrimination. In particular, many facial recognition technologies were subjected to higher error rates for people of color and women (and other marginalized groups) (Garvie, Bedoya, & Frankle, 2016). Our results endorse this, with two algorithms (SVM and Isolation Forest) causing higher false favorable rates (9% and 10%, respectively), which may aggravate the risk of misclassifying an individual based on prejudiced data. Therefore, such biases continued propagating privileges (Eubanks, 2018), mainly when we used

AI-powered systems in automation/bias sensitive sectors (such as Law Enforcement and other sensitive security applications). Therefore, fairness and bias represent crucial ethical challenges posed by AI models used for surveillance systems that must be addressed.

Ensuring audibility is yet another concern in AI-based surveillance systems. This raises the question of who is responsible when AI makes an incorrect decision or causes harm, such as, for example, misidentifying an innocent person as a suspect. Is the developer of the AI system responsible, the organization that puts it to use, or the AI system? Moreover, Binns (2018) refers to the increasing autonomy of agentic AI as giving rise to a "responsibility gap" that makes it difficult to ascertain who or who is to blame for the actions of an AI system. As AI systems continue to gain levels of independence, the stakes of ethical implementation only grow, and frameworks for accountability and transparent decision making processes familiar to the world of human operated labor must also be established for AI.

Achieving a Balance of Efficiency with Ethical Issues

One of the biggest hurdles is balancing efficiency gains with ethical risks in agentic AI integration into surveillance technologies. AI, on the one hand, provides significant enhancements to the speed and acumen of surveillance systems. Examples include XGBoost's real-time (1.1 seconds) threat detection and high accuracy (91%), indicating that AI-driven solutions can increase security, even in shared spaces. Indeed, the potential of AI to aid in security has been described as both promising and threatening (Hancock and Kaye, 2020 thinking); thus, weighing these two remains an important consideration when using it, considering the invasion of privacy this might incur and a tendency towards surveillance overreach. Surveillance systems should employ strong ethical guidelines to ensure they do not trample on individual freedoms or turn oppressive.

In addition, if AI empowers surveillance, it can ultimately become what Zuboff (2019) describes as "surveillance capitalism," in which individuals are treated as commodities and their behaviour are perpetually surveilled, analyzed, and exploited. Transforming surveillance data into a commercial product poses a grave ethical challenge, as it can compromise agency and personal privacy in exchange for security or profit.

Consideration for Regulatory Governance

Considering the grave ethical issues posed by AI-based surveillance, strong regulatory frameworks and governance systems must be implemented. Different organizations, such as the European Union with the General Data Protection Regulation (GDPR), have published policy proposals regarding the use of AI in surveillance systems. These guidelines underscored the need for consent, transparency, fairness, and accountability in using such technologies (European Union, 2018). Binns (2018) claims that governments and regulatory bodies must enforce strict regulations to narrow down the possibilities of AI-powered surveillance and guarantee that these technologies work in the direction of human rights and ethical principles.

On another note, accountability in AI decision making must be maintained through transparency. 26) AI systems should be designed to allow their decisions to be audited and explained in understandable terms to the public and other regulators. Many artificial intelligence (or AI) models especially of the deep learning variety function as "black boxes" that obfuscate the mechanics of decision making (Burrell, 2016). Making these systems more interpretable will be essential for resolving ethical issues related to accountability and fairness.

The findings highlight how XGBoost and comparable machine learning models can significantly enhance the accuracy and efficiency of surveillance systems. However, it also raises important ethical considerations that need to be confronted. Among those are privacy violations, algorithmic bias, and the issue of accountability for decisions made by AI. September 30, 2023, The Sentence Transformer Output to achieve ethical use of AI-powered surveillance technologies, there needs to be a strong regulatory framework for ethical use of this technology with a focus on transparency, fairness, and accountability during the design process of AI systems and after their implementation. Developing AI with these ethical considerations allows us to utilize our quest to keep us secure without sacrificing our inalienable human rights and freedoms.

CONCLUSIONS AND RECOMMENDATIONS

The backward and forward of agentic AI to Surveillance technologies is a paradigm shift in security operations, ensuring higher efficiency, scalability, and accuracy. The study showed that the machine learning model XGBoost is a great performer in real-time threat detection, surpassing Random Forest, SVM, and Isolation Forest on multiple performance measures: accuracy, precision, recall, and response time. Agentic AI can make decisions autonomously , leveraging real-time data, which has numerous advantages in environments like these that are complex and ever changing, where human oversight would be insufficient. Hence, the technical efficiencies of AI-backed surveillance systems are proportionate to the increasing demand for proactive and predictive security solutions, making these technologies crucial in the modern surveillance infrastructure (Russell, 2019).

But, the use of agentic AI in surveillance also poses significant ethical challenges, primarily regarding privacy, accountability, and bias. AI systems that collect and analyze personal data in real-time could pose major privacy risks without precise consent mechanisms or independent oversight. Such technologies could lead to a "surveillance state," one in which individual freedoms are endangered, Zuboff (2019) claimed. Moreover, algorithmic bias in AI models, e.g., facial recognition, raises the adoption of discriminatory practices, as some demographic groups have higher false favorable rates (Garvie, Bedoya, & Frankle, 2016). Such ethical issues suggest the necessity of comprehensive regulatory frameworks and oversight mechanisms to ensure that AI-powered surveillance technologies are deployed in ways that respect human rights and prevent abuse.

Because agentic AI systems will be agents in a technical sense, they won't transcend the causation framework of responsibility." When an AI system makes

a mistake, such as identifying a person incorrectly, it is challenging to attribute responsibility, which can erode trust in these systems. As Binns (2018) argues, without such mechanisms, AI systems may ascend from human control and operate out of sight beyond the limits of human power. Also, algorithm transparency is critical, meaning that decisions made by AI models should be interpretable and verifiable to concerned parties, such as the public and regulators (Burrell, 2016).

These findings underscore the necessity of deploying agentic AI in surveillance technologies within robust ethical guidelines and regulatory frameworks. The European Union General Data Protection Regulation (GDPR) (European Union, 2018) serves as a model for navigating this territory by balancing the benefits of AI while providing individuals data protection and privacy rights, such as informed consent, data transparency, and individual agency. Additional studies in this area will need to demonstrate effective bias reduction techniques, especially in models powering surveillance systems highly reliant on biased data that may negatively impact specific groups (Eubanks, 2018). In addition, watchdogs need to guarantee that technologies retain accountability and transparency, confronting the moral challenges associated with using them for surveillance.

Although agentic AI can transform real-time threat detection and surveillance, stakeholder buy in and controls must be adjusted accordingly to mitigate privacy and fairness concerns and accountability in these interactions. Implementing clear ethical frameworks, transparency, and strong regulatory frameworks will ensure that AI-based surveillance systems can effectively improve security without undermining individual liberties and worsening social inequities.

ADVANCED RESEARCH

The integration of agentic AI into surveillance technologies marks a pivotal evolution in security paradigms, introducing unprecedented levels of autonomy, responsiveness, and operational intelligence; however, this technological leap demands an equally sophisticated ethical and regulatory response. Advanced research should critically explore the development of explainable AI (XAI) frameworks tailored to surveillance systems, ensuring algorithmic decisions remain transparent and accountable, particularly in high-stakes environments involving real-time facial recognition and behavioral analysis. Moreover, interdisciplinary studies must address how bias mitigation techniques such as adversarial debiasing or fairness-aware machine learning can be embedded at both data and algorithmic levels to counteract systemic discrimination, especially against historically marginalized communities. As agentic AI increasingly assumes roles traditionally managed by human oversight, emerging governance models must reconfigure liability structures, clarify responsibility attribution, and develop robust consent protocols that align with the principles of digital sovereignty and human rights. This research trajectory should be underpinned by empirical investigations into socio-technical impacts, utilizing methodologies from computational ethics, legal informatics, and human-centered AI to inform

adaptive regulatory frameworks that not only reflect global data protection standards like the GDPR but also proactively anticipate future ethical dilemmas.

REFERENCES

Adedeji, M., Abid, M., Adun, H., Ogungbemi, A. T., Alao, D., & Zaini, J. H. (2022). Thermodynamic Modeling and Exergoenvironmental Analysis of a Methane Gas-Powered Combined Heat and Power System. Applied Sciences, 12(19), 10188.

Adun, H., Adedeji, M., Titus, A., Mangai, J. J., & Ruwa, T. (2023). Particle-Size Effect of Nanoparticles on the Thermal Performance of Solar Flat Plate Technology. Sustainability, 15(6), 5271.

Adun, H., Ishaku, H. P., & Ogungbemi, A. T. (2022). Towards renewable energy targets for the Middle East and North African region: a decarbonization assessment of energy-water nexus. Journal of Cleaner Production, 374, 133944.

Adun, H., Ishaku, H. P., Ayomide Titus, O., & Shefik, A. (2022). 3-E feasibility analysis on photovoltaic/thermal application for residential buildings: a case study of Sub-Saharan Africa. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 44(4), 9901-9919.

Al Imran, S. M., Islam, Md. S., Kabir, N., Uddin, I., Ali, K., & Halimuzzaman, Md. (2024). Consumer Behavior and Sustainable Marketing Practices in the Ready-Made Garments Industry. International Journal of Management Studies and Social Science Research, 6(6), 152–161. https://doi.org/10.56293/IJMSSSR.2024.5322

Barua, T., & Mondal, B. (2024). Data Security In Iot Devices And Sensor Networks For Robust Threat Detection And Privacy Protection. *ACADEMIC JOURNAL ON SCIENCE, TECHNOLOGY, ENGINEERING & MATHEMATICS EDUCATION*, *1*(01), 10-69593.

Binns, R. (2018). The ethics of artificial intelligence and surveillance. *Journal of Ethics and Information Technology, 20*(2), 89102. https://doi.org/10.1007/s106760189477x

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 532. https://doi.org/10.1023/A:1010933404324

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society, 3*(1), 113. https://doi.org/10.1177/2053951715622512

Chowdhury, M. R. I., Chowdhury, T. R., & Abdullah, S. B. (2024). Strategies for improving patient experience and satisfaction in healthcare facilities in USA. *International Journal of Science & Healthcare Research*, *9*(4), 357-369.

Eubanks, V. (2018). *Automating inequality: How hightech tools profile, police, and punish the poor*. St. Martin's Press.

Garvie, C., Bedoya, A., & Frankel, J. (2016). *The perpetual lineup: Unregulated police face recognition in America*. Georgetown Law Center on Privacy & Technology. https://www.perpetuallineup.org

Habib, K., Nuruzzamal, M., Shah, M. E., & Ibrahim, A. S. M. (2019). Economic Viability of Introducing Renewable Energy in Poultry Industry of

Bangladesh. International Journal of Scientific & Engineering Research, 10(3), 1510-1512.

Halimuzzaman, Md., Atif, H. M., Kumar, P., & Salehin, M. (2024). Public Relation and Educational Outcomes of Films in Bangladesh: A Study on Hawa. Journal of Primeasia, 5(1), 1–7. https://doi.org/10.25163/primeasia.519834

Hancock, P. A., & Kaye, M. (2020). AI surveillance in public spaces: Opportunities and ethical concerns. *Journal of Applied Ethics, 28*(4), 371380. https://doi.org/10.1007/s10462020098797

Hossain, M. A., & Rahman, T. Y. Cognitive AI for Wildfire Management in Southern California: Challenges and Potentials.

Hossain, M. A., & Rahman, T. Y. Human Factors and Employee Resistance to Adopting New Cybersecurity Protocols and Technologies. Journal for Multidisciplinary Research, 1(03), 175-199.

Hossain, M. A., Raza, M. A., & Rahman, J. Y. (2025). Investigating the Cybersecurity Implications of Open Banking and Application Programming Interfaces (APIs) in the Financial Sector. *Jurnal Ekonomi dan Bisnis Digital (MINISTAL)*, *4*(1), 39-56.

Hossain, M. A., Raza, M. A., & Rahman, T. Y. (2023). Resource allocation and budgetary constraints for cybersecurity projects in small to medium sized banks. Journal of Multidisciplinary Research, 9(01), 135-157.

Hossain, M. A., Raza, M. A., Al Mamun, M. H., Rahman, T. Y., & Rahman, J. Y. Smart City Sensors for Tailored Learning Experiences.

Hossain, M. A., Raza, M. A., Mahjabeen, F., & Yaseer, J. (2025). Assessing the Vulnerabilities of Mobile Banking Applications and Developing Strategies to Improve Their Security. *Jurnal Ekonomi dan Bisnis Digital (MINISTAL)*, *4*(1), 1-18.

Ibrahim, A. S. M., Rahman, M., Dipu, D. K., Mohammad, A., Mazumder, G. C., & Shams, S. N. (2024). Bi-Facial Solar Tower for Telecom Base Stations. Power System Technology, 48(1), 351-365.

Islam, M. S. H., Rubel, M. R. B., Hossain, M. I., Kamruzzaman, M., Akter, S., Halimuzzaman, M., & Karim, M. R. (2024). Impact of financial and internet support on SME performance: Moderating effect of technology adoption during COVID-19 pandemic. World Journal of Advanced Engineering Technology and Sciences, 13(2), 105–118. https://doi.org/10.30574/wjaets.2024.13.2.0533

Kabir, H. M. D., Anwar, S., Ibrahim, A. S. M., Ali, M. L., & Matin, M. A. Watermark with Fast Encryption for FPGA Based Secured Realtime Speech Communication. Consumer Electronics Times, 75-84.

Mazumder, G. C., Ibrahim, A. S. M., Rahman, M. H., & Huque, S. (2021). Solar PV and wind powered green hydrogen production cost for selected locations. International Journal of Renewable Energy Research (IJRER), 11(4), 1748-1759.

Mazumder, G. C., Ibrahim, A. S. M., Shams, S. N., & Huque, S. (2019). Assessment of Wind Power Potential at the Chittagong Coastline in Bangladesh. Dhaka University Journal of Science, 67(1), 27-32.

Mazumder, G. C., Shams, S. N., Ibrahim, A. S. M., & Rahman, M. H. (2019). Practical Study of Water Electrolysis for Solar Powered Hydrogen Production Using Stainless Steel Electrode and Sodium Hydroxide Solution. International Journal of New Technology and Research, 5(3), 84-90.

Mohammad, A., Mahjabeen, F., Tamzeed-Al-Alam, M., Bahadur, S., & Das, R. (2022). Photovoltaic Power plants: A Possible Solution for Growing Energy Needs of Remote Bangladesh. *NeuroQuantology*, *20*(16), 1164.

Ogungbemi, A. T., Adun, H., Adedeji, M., Kavaz, D., & Dagbasi, M. (2022). Does Particle Size in Nanofluid Synthesis Affect Their Performance as Heat Transfer Fluid in Flat Plate Collectors?—An Energy and Exergy Analysis. Sustainability, 14(16), 10429.

Rahman, M. R., Hossain, M. S., Shehab Uddin, S., & Ibrahim, A. S. M. (2019). Fabrication and Performance Analysis of a Higher Efficient Dual-Axis Automated Solar Tracker. Iranica Journal of Energy & Environment, 10(3), 171-177.

Raza, M. A., Hossain, M. A., Mahjabeen, F., Rahman, J. Y., & Rahman, T. Y. (2025). Evaluating the Human Factor in Bank Cybersecurity: Strategies for Improving Employee Awareness and Reducing Insider Threats. *Indonesian Journal of Advanced Research (IJAR)*, *4*(1), 1-20.

Russell, S. J. (2019). *Artificial intelligence: A modern approach* (4th ed.). Pearson Education.

Sharfuddin, M., Halimuzzaman, Md., Akter, F., Nath Dey, K., & Saha, P. (2025). Employee Motivation and Behavior in Construction Engineering Projects. International Journal of Social Science and Economic Research, 10(1), 342–372. https://doi.org/10.46609/IJSSER.2025.v10i01.019

Tansu, A., Ogungbemi, A. T., & Hocanın, F. T. (2022). The challenges and serviceability of solar power: Suggestion on solving the Nigeria energy crisis. International Journal of Energy Studies, 7(2), 127-141.

Uddin, M. M., Rahaman, M. A., Chowdhury, M. R. I., & Ahmad, I. (2024). Patient Outcomes Through Machine Learning: A Review Of Data Management Strategies in Healthcare. *Journal of Next-Gen Engineering Systems*, *1*(01), 89-106.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.