Operator Machine Augmentation Resource Framework

Mohammed Ameen[1], Richard Stone[2], Majed Hariri[3], Faisal Binzagr[4]
Human-computer Interaction Department, Iowa State University, Ames, USA[1]
Industrial and Manufacturing Systems Engineering Department, Iowa State University, Ames, USA[2]
Human-computer Interaction Department, Iowa State University, Ames, USA[3]
Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh, KSA[4]

***Abstract*—The growing number of people gathering in public and the massive incidents that have occurred in recent years. It raises questions about public safety and security. This paper illustrates the technical implementation of the operator machine augmentation resource (OMAR) framework, which integrates advanced technologies, including a Computer Vision model and CCTV operators' training techniques, to address the limitations of traditional surveillance systems. The OMAR framework enhances the productivity of surveillance systems by facilitating operators' tasks and improving theirs. The framework's components, including Alert Triggers, a Computer Vision model, and human training, work together to create better output, and a more convincing system will improve the quality of security and reduce human effort. Although the OMAR framework represents a potentially significant step forward in surveillance security systems, it remains a theoretical model requiring further investigation and rigorous testing. Future work will focus on evaluating the effectiveness of the OMAR framework through an empirical study, examining its impact on various aspects of human performance and adaptations.**

***Keywords*—Crowd monitoring; public security; Operator Machine Augmentation Resource (OMAR) framework; CCTV operator; surveillance system; crowd monitoring systems**

I. Introduction

In today's increasingly complex and connected world, the need for vigilant and continuous monitoring is undeniable. Crowd monitoring systems, particularly CCTV, have long been integral to security strategies, but the human operators behind these systems often face challenges in maintaining performance over prolonged periods. Recognizing the limitations of human operators, recent advancements have increasingly integrated intelligent surveillance technologies, like computer vision models, to augment human capabilities [1].

Originally, surveillance relied solely on human operators, but the integration of artificial intelligence (AI) has significantly transformed the field. AI in CCTV systems enhances surveillance by automating the analysis of video data, thus reducing the workload on human operators and improving overall efficiency [2]. Despite AI's ability to enhance data processing and event detection, its effectiveness is maximized when used in conjunction with human oversight. Human operators are still crucial for handling complex decision-making processes and responding to unpredictable scenarios.

In surveillance contexts, the scope of monitoring extends beyond mere criminal activities to encompass all instances of abnormal behavior. Abnormal behavior in public settings, demonstrated by erratic or aggressive actions, can sometimes present more significant risks than conventional criminal acts, like the tragic incident that happened during the Hajj pilgrimage in Makkah in 2015 [3]. The tragic incident highlights the critical need for vigilant and responsive surveillance systems. The incident that occurred in Makkah in 2015 emphasizes the necessity for a comprehensive surveillance strategy that effectively integrates AI technology with the essential oversight provided by human operators. The surveillance strategy, which blends AI technology and human operators, aims to enhance the system's accuracy and adaptability. A hybrid surveillance strategy system reduces the likelihood of errors and increases the system's efficacy in managing diverse and complex security situations [4]. A hybrid system blend of AI and human input ensures that surveillance systems are adept at managing both regular and exceptional security challenges, making them indispensable in modern security strategies.

II. Objective

The integration of AI into surveillance systems represents a transformative leap in the security domain, enhancing both the efficiency and the effectiveness of monitoring operations. The objective section of the paper will explore AI technologies' multifaceted roles in augmenting human operators' capabilities within security frameworks. By utilizing advanced AI, surveillance operations can surpass traditional limitations, offering precision and swiftness in addressing security threats and incidents. AI transforms surveillance systems by introducing real-time data analysis, enhanced object recognition, and active monitoring, significantly boosting the accuracy and response times to potential threats. For example, AI-enabled cameras actively analyze footage to detect unusual activities, thus allowing for immediate intervention. Such hybrid systems can accurately differentiate between types of movements and objects, cutting down on false alarms and enabling a more focused approach to real threats [5]. The comparison between operators working with and without AI assistance highlights the substantial enhancements brought by technology. AI aids in quickening response times and increasing detection accuracy, thus supporting human operators in maintaining robust security standards effectively. AI's role is transformative, redefining the operational dynamics of surveillance tasks and setting new standards in security protocols [2]. Moreover, the development of specific AI algorithms tailored for critical and

large-scale environments, like the Holy Mosque, emphasizes the customized application of AI in surveillance. These algorithms are finely tuned to detect subtle and context-specific behavioral cues, thus enhancing the sensitivity and accuracy of surveillance in areas with significant cultural and religious importance. Additionally, the impact of AI on the workload and stress levels of operators is profound. By optimizing AI system designs for enhanced operator comfort and effectiveness, surveillance tasks become less brutal and more efficient, potentially improving job satisfaction and reducing staff turnover. The combination of AI and human operators underscores the human-centered approach in designing AI surveillance systems, ensuring they support rather than replace the human element [4]. The deployment of AI technologies in surveillance systems offers substantial enhancements in these systems' operational effectiveness and technical capabilities. Through sophisticated analytics and adaptive learning algorithms, AI enables a responsive surveillance mechanism that significantly aids human operators in executing their duties with unmatched precision and efficiency. This integration marks a significant advancement in surveillance technology, promising marked improvements in security, safety, and operational efficiency across various settings.

III. The Role of AI in CCTV Systems

A. AI in CCTV System

AI-enabled CCTV systems represent a significant leap forward in surveillance technology, offering capabilities for monitoring and threat detection. These intelligent systems are engineered to scrutinize video footage in real-time, utilizing deep learning techniques to discern patterns and behaviors indicative of potential security breaches or emergencies [6]. AI comprises sophisticated algorithms that can analyze body language, facial expressions, and movement trajectories, distinguishing between normal and suspicious behaviors with a high degree of accuracy [7]. This integration of AI into CCTV systems, with these features, allows for the automated detection of a wide array of activities, ranging from violence in crowded spaces to unauthorized access in restricted areas. Therefore, the integration of AI into CCTV systems enables automated detection of diverse activities, enhancing surveillance capabilities for proactive threat identification and security management. The real-time processing power of AI significantly enhances the responsiveness of surveillance systems. In scenarios where every second counts, such as detecting a left-behind package in a busy terminal or identifying an individual brandishing a weapon, AI-driven CCTV systems can alert human operators instantaneously [8]. For example, they can monitor the speed and direction of individuals in a crowd, flagging those moving against the flow or at an unusual pace, which could indicate a person in distress or someone with malicious intent [9]. This rapid identification enables quicker decision-making and potentially life-saving interventions [10]. The efficacy of these systems is not just theoretical; empirical research has demonstrated their robustness in various settings, including community spaces and transportation hubs, where they have been instrumental in maintaining public safety [11]. Therefore, leveraging AI-enhanced surveillance systems is adept at recognizing subtler nuances of human behavior.

B. Efficiency

The integration of advanced AI within security and surveillance frameworks has indeed brought about a significant transformation in the industry. This transformation can be observed in various aspects of security and surveillance systems, including but not limited to threat detection, monitoring, and response mechanisms [12]. A prime example of such innovation is the development of AI-powered Intelligent Surveillance tools that incorporate the YOLO (You Only Look Once) algorithm. This framework processes images in real-time, simultaneously identifying and classifying multiple objects [13]. It divides the image into a grid and predicts bounding boxes and probabilities for each grid cell. The predictions are then filtered through non-maximum suppression to provide the final detection outcomes [13]. This approach allows YOLO to detect objects with high precision rapidly, making it an ideal choice for surveillance in densely populated areas where efficiency is crucial. Consequently, AI-driven security technologies significantly bolster operational efficiency.

C. Accuracy

The advent of AI has led to substantial enhancements in the effectiveness of monitoring systems, particularly by minimizing erroneous alarms. Conventional security setups often face challenges with excessive false alerts, which may result in unwarranted responses and fatigue among security staff. AI algorithms are designed to learn from extensive data sets, allowing them to differentiate between real dangers and harmless irregularities with greater precision [14]. Employing advanced learning models, such as CNNs, enables these systems to analyze and assess video footage instantly, delivering prompt and precise evaluations [6]. Furthermore, AI-based surveillance mechanisms are not static; they evolve continually, thereby increasing their accuracy over time. This adaptive nature means that they become more familiar with the unique settings they oversee, leading to fewer false alarms. Consequently, incorporating AI into surveillance provides a more effective way of protecting public areas.

IV. Challenges Associated with Sole AI Operation

A. Lack of Contextual Understanding

The challenge of contextual understanding in AI, particularly in surveillance, is a significant hurdle researchers and practitioners are trying to overcome. AI systems, including those used in surveillance, excel at identifying patterns and correlations but often lack the depth of understanding necessary to accurately interpret complex human behaviors and contexts. This limitation can lead to misinterpretations and inappropriate responses, as AI may not discern the subtleties of cultural differences or non-threatening unusual behaviors [15]. For example, an AI surveillance system might flag a janitor standing still as a potential threat, failing to recognize the context that the individual is merely performing their job duties. As such, bridging the gap between pattern recognition and contextual understanding is paramount for the responsible deployment of AI in surveillance, ensuring accurate interpretation and appropriate responses to diverse human behaviors.

B. Privacy Concerns

The incorporation of AI into facial detection systems within CCTV networks presents significant privacy issues [16]. In numerous regions, regulations explicitly forbid the use of facial data in software applications. As noted by scholars and privacy advocates, the deployment of AI surveillance systems that operate without human intervention exacerbates these privacy concerns. A key issue is the potential for these systems to autonomously collect and analyze extensive amounts of personal data without sufficient oversight or consent, mainly when such data includes facial recognition information [17]. The demand for security systems that utilize AI while protecting human privacy is now more critical than ever.

V. Human-AI Synergy

A. Human-AI Collaboration

The relationship between human thinking and AI systems is a topic of continuing discussion. AI systems can make better choices than people in certain circumstances, but people can make superior decisions when judgment is needed [18]. In addition, people interpret information within a framework of social values, cultural influences, and unique life experiences, enabling them to make decisions that are both complex and context-dependent [19] [20]. Human involvement is particularly essential in scenarios requiring ethical sensitivity, empathy, and moral discernment. For example, while AI can detect trends and highlight irregularities in monitoring activities, the human operator can evaluate the purpose behind actions, recognize special situations, and make choices that uphold privacy and individual freedoms. Thus, academic discussion suggests AI should enhance human decision-making rather than take its place.

B. Human Training and Competency

1) System proficiency: The skill level of CCTV operators in leveraging AI-based tools plays a crucial role in the successful operation of surveillance systems. Skilled operators grasp the complexities of AI systems, adjusting settings and fine-tuning the technology to reduce false alarms and improve precision. Their knowledge ensures that AI processes the most pertinent data, eliminates irrelevant information, and zeroes in on significant occurrences [21]. Additionally, capable operators interpret AI alerts with context, considering aspects such as the time of day, location, and setting. Their judgment helps avoid unnecessary actions while promptly addressing legitimate threats. Skilled operators act as a link between AI data and practical measures, weighing urgency and potential risks to determine if an alert should be escalated [22]. Their expertise and insight contribute to the overall efficiency. Therefore, CCTV operators must be trained for proficiency.

2) Situational awareness and skill development: The ability to perceive and understand the situation is a crucial part of a CCTV operator's duties. It entails knowing the area under surveillance and being capable of analyzing the importance of specific actions or events. Training in situational awareness provides operators with the capability to evaluate the seriousness of dangers and foresee possible problems ahead of time, resulting in a security approach that is more preventative than reactive [23]. Additionally, cultivating skills that align with AI is an important aspect of training for CCTV operators. Even though AI can recognize a variety of behaviors and irregularities, there may still be subtle indications that require human judgment. Operators need to be skilled at spotting these subtle distinctions that AI may miss. Therefore, the synergy between human expertise and technological capabilities is crucial for an efficient surveillance approach.

C. Challenges without AI Assistance

The transition from manual video review to AI-backed video analytics has significantly accelerated the monitoring and detection of the targeted process, allowing operators to focus more on critical details rather than being distracted by others unnecessarily [24]. Secondly, human operators can lose focus or become distracted while monitoring camera feeds, leading to missing crucial details. Thirdly, another significant challenge is human reaction times are slower than AI algorithms, which can cause delayed responses and impact emergency interventions or crime prevention. Therefore, the integration of AI technology with CCTV systems mitigates human limitations.

VI. Operator Machine Augmentation Resource

The OMAR framework represents a significant advancement in enhancing operational efficacy for CCTV operators. It has been systematically engineered to optimize task workflows and elevate overall performance metrics within surveillance systems. This innovative framework integrates advanced methodologies supported by sophisticated artificial intelligence models, reflecting recent developments in AI-driven surveillance technologies [25]. Additionally, it includes comprehensive training tailored explicitly for operators, which has been shown to augment human performance in surveillance environments [26]. This holistic approach ensures that the Omar framework not only harnesses the latest technological advancements but also substantially enhances the skill set of the personnel involved, thereby increasing the number of hits and response times of surveillance operations.

A. Detection Model

The Detection Model is a critical component of the CCTV hybrid model, designed to analyze live video efficiently feeds by combining the strengths of artificial intelligence (AI) and human operators. The model leverages advanced machine learning techniques to automatically detect and annotate objects of interest in real time while also providing a seamless interface for human operators to review and validate the AI-generated annotations.

1) Technical architecture: The technical architecture of the Detection Model is composed of several key components that work together to enable efficient and accurate video analysis. The following diagram illustrates the interactions between these components (see Fig. 1).

The architecture consists of the following main components:

- **LiveFeed**: This component is responsible for capturing and transmitting live video data from CCTV

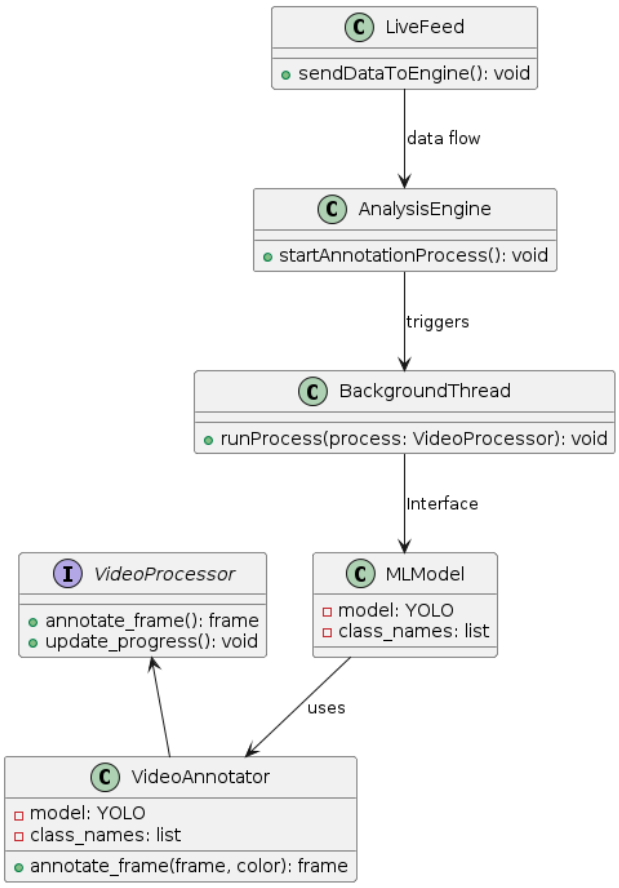

Fig. 1. Detection model's technical architecture.

cameras to the AnalysisEngine. It continuously sends video frames to the AnalysisEngine for real-time processing and annotation.

- **AnalysisEngine**: The AnalysisEngine is the central component that orchestrates the video analysis process. It receives live video data from the LiveFeed and triggers the annotation process by initiating a BackgroundThread.
- **VideoProcessor** (interface): The VideoProcessor interface defines the common methods that need to be implemented by the video processing components, such as annotating frames and updating progress. It serves as a blueprint for the MLModel and VideoAnnotator components.
- **MLModel**: The MLModel component implements the VideoProcessor interface. It utilizes a pre-trained machine learning model to automatically detect and annotate objects of interest in the video frames. The MLModel operates in the background, processing the video frames efficiently.
- **VideoAnnotator**: The VideoAnnotator component is responsible for providing a user interface for human operators to review and validate the annotations generated by the MLModel. It allows human operators to manually annotate frames, add additional annotations, or modify the existing annotations as needed. The VideoAnnotator uses the MLModel to obtain the initial annotations and then presents them to the human operator for review and refinement.
- **BackgroundThread**: The BackgroundThread component is responsible for running the video processing tasks in the background. It takes an instance of the VideoProcessor (MLModel) and executes the annotation process asynchronously, ensuring that the main application remains responsive.

2) User flow: The user flow of the Detection Model is designed to provide a seamless experience for both AI-assisted and human-operated video analysis. The following sequence diagram illustrates the user flow and interactions between the various components (see Fig. 2).

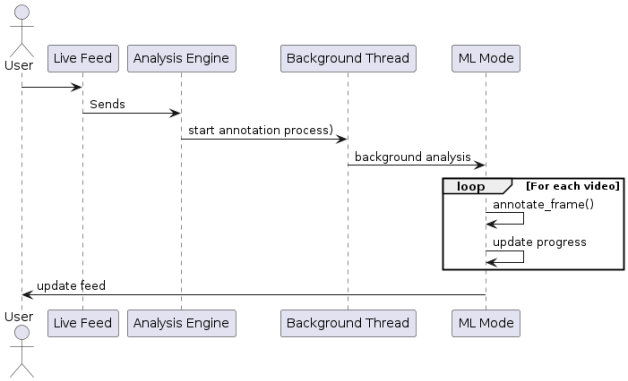

Fig. 2. Detection model's user flow.

The user flow consists of the following steps:

1) The User initiates the live video feed, which the LiveFeed component captures.
2) The LiveFeed sends the video data to the AnalysisEngine for processing.
3) The AnalysisEngine starts the annotation process by triggering the BackgroundThread.
4) The BackgroundThread utilizes either the MLModel or the VideoAnnotator, depending on the mode of operation, to perform the video analysis.
5) The MLModel annotates each video frame using the pre-trained YOLO model.
6) The MLModel updates the progress of the annotation process.
7) The annotated video frames are returned to the User interface for real-time display and further analysis.

The Detection Model's architecture and user flow demonstrate the effective integration of AI and human expertise in the CCTV hybrid model. By leveraging the strengths of both machine learning and human operators, the Detection Model can provide accurate and reliable video analysis, enhancing the overall effectiveness of the CCTV system in detecting and responding to objects of interest [22].

3) YOLO: OMAR framework uses YOLO model. It is a state-of-the-art object detection system renowned for its speed and accuracy. It processes images in a single pass,

predicting bounding boxes and class probabilities directly from full images in one evaluation, making it significantly faster than systems that propose regions and then classify them [27]. This efficiency enables YOLO to achieve high frame rates while maintaining good accuracy, making it particularly well-suited for real-time applications such as video surveillance and live video analysis [28] [29].

The OMAR framework signifies a state-of-the-art integration of computer vision and graphical user interface technologies tailored for video content analysis using the YOLO object detection model. The framework initializes a robust logging system to capture detailed debug-level data, which is essential for troubleshooting and development purposes. The core of the OMAR framework functionality resides within the analysis engine, which processes the frames coming from the live feed, transfers them to the YOLO model, and applies them to video frames to detect and annotate objects. Each frame is processed to draw bounding boxes, and label detected objects with class (behaviors) names, enhancing the raw video data with valuable contextual information. The annotations are performed frame-by-frame, with properties such as frame dimensions and frame rate meticulously extracted using OpenCV, thereby preserving the video's original specifications in the annotated output.

Furthermore, the OMAR framework is designed with a user-friendly interface using Tkinter. Tkinter is a standard GUI toolkit for Python. It enables the development of user-friendly graphical interfaces in Python applications. Tkinter is included in the Python standard library, making it widely accessible and straightforward for creating desktop applications [30]. An intuitive and straightforward user interface lets users observe annotated videos through a simple GUI and provides near-real-time updates. The OMAR framework design facilitates ease of use and supports asynchronous processing by employing threading, thus maintaining UI responsiveness. The framework is particularly suited for safety environments where processing large video datasets for object detection and tracking pilgrims is essential, offering a practical tool for CCTV operators to evaluate situations and perform actions directly on video.

B. Human Training

The performance of CCTV operators is markedly improved when they receive comprehensive training. Adequate training endows operators with the essential skills to proficiently monitor and analyze surveillance footage, enhancing their ability to detect and swiftly respond to suspicious activities accurately. It has been observed that CCTV operators exhibit superior performance when they clearly understand their objectives and some specific indicators [31]. The OMAR framework emphasizes key elements that are crucial for effective surveillance operations:

1) Work environment: The OMAR framework prioritizes comprehensive training to ensure that CCTV operators deeply understand their working environments, recognizing how such knowledge significantly enhances their performance and effectiveness. Moreover, experienced CCTV operators quickly identify crucial aspects of their tasks earlier than their less experienced peers, suggesting that targeted training substantially improves their operational capabilities [32]. Further studies have been conducted, delving into the daily challenges and practices of these operators, highlighting the critical need for a thorough grasp of their work environments to ensure effective surveillance [33]. The OMAR framework aims to comprehensively cover this aspect of training, ensuring that all operators gain a thorough knowledge of their work environments, thereby enhancing their ability to monitor and respond to incidents effectively.

2) Job roles: The OMAR framework is intended to elevate the competencies of CCTV operators, ensuring they comprehend the full scope of their professional duties. This understanding is crucial for boosting their performance and improving their decision-making skills and overall well-being. Operators can effectively execute their roles and responsibilities by implementing appropriate training strategies and harnessing their experience. This comprehensive approach fosters a well-rounded development, ultimately enhancing efficacy and accuracy.

3) Skills and competencies: The OMAR framework is designed to cultivate a comprehensive set of skills and competencies in CCTV operators, which are crucial for effective surveillance, crime prevention, and prompt response. This training enhances operators' proficiency in handling sophisticated surveillance technologies and focuses on developing their overall capabilities through targeted training programs, competency assessments, and skill development initiatives.

4) Nature of the place: The OMAR framework training program equips CCTV operators with a deep understanding of the environments they monitor, which is essential for effective surveillance. This contextual awareness enables operators to differentiate between normal and suspicious activities, optimizing their monitoring efforts and enhancing response times. Familiarity with the environment also helps reduce false alarms and improve the coordination of responses. Additionally, operators trained in this way can identify potential vulnerabilities and recommend proactive security measures. By incorporating this environmental knowledge into training programs, the OMAR framework ensures that CCTV operators are well-prepared, efficient, and practical, ultimately contributing to a safer and more secure environment.

VII. Conclusion and Future Work

The OMAR framework, as articulated in this paper, represents a comprehensive and innovative approach to surveillance monitoring systems. By harnessing advanced technologies, such as machine learning and specialized training techniques, the OMAR framework establishes a reliable and confidential monitoring system designed to enhance safety and security. The framework's components are intricately integrated to form a cohesive and efficient system that encourages optimal performance among CCTV operators. Specifically, the OMAR detection model identifies and alerts operators to abnormal behaviors, while the training component is designed to enhance human motivational factors for CCTV operators. Additionally, the Computer Vision Model enables continuous behavior recognition, improving user experience by accurately identifying anomalies within crowds. Although the OMAR framework marks a significant advancement in

surveillance and security system interventions, it is important to acknowledge that it remains a theoretical construct at this stage. Its potential applications and impacts necessitate further exploration and rigorous testing. This paper does not declare the achievement of specific outcomes; instead, it aims to delineate the implementation of the OMAR framework. Consequently, future research will concentrate on empirically evaluating the effectiveness of the OMAR framework. This research will investigate the framework's influence on various aspects of user experience to ascertain its efficacy in enhancing surveillance systems. Future studies will assess the OMAR framework's impact on human performance levels in comparison to traditional surveillance systems. Additionally, these studies will examine the framework's effects on performance, visual discrimination, cognitive load, trust, and confidence. In conclusion, the OMAR framework offers a promising solution to the challenges faced by security systems. Through rigorous testing and refinement, we aspire to contribute to a future where advanced artificial intelligence technologies and effective human training interventions are accessible, fostering a safer and more secure public environment.

REFERENCES

[1] M. Ameen and R. Stone, "Advancements in crowd-monitoring system: A comprehensive analysis of systematic approaches and automation algorithms: State-of-the-art," *arXiv preprint arXiv:2308.03907*, 2023.

[2] N. Dadashi, A. Stedmon, and T. Pridmore, "Semi-automated cctv surveillance: the effects of system confidence, system accuracy and task complexity on operator vigilance, reliance and workload." *Applied ergonomics*, vol. 44 5, pp. 730–8, 2013.

[3] P. Salamati and V. Rahimi-Movaghar, "Hajj stampede in mina, 2015: Need for intervention," *Archives of trauma research*, vol. 5, no. 2, 2016.

[4] J. De Bruyne, J. Joundi, J. Morton, N. Van Kets, G. Van Wallendael, D. Talsma, J. Saldien, L. De Marez, W. Durnez, and K. Bombeke, "Smooth operator: a virtual environment to prototype and analyse operator support in cctv surveillance rooms," in *International Conference on Human-Computer Interaction*. Springer, 2021, pp. 233–240.

[5] J. De Bruyne, J. Joundi, J. Morton, A. Zheleva, N. Van Kets, G. Van Wallendael, D. Talsma, J. Saldien, L. De Marez, W. Durnez *et al.*, "I spy with my ai: The effects of ai-based visual cueing on human operators' performance and cognitive load in cctv control rooms," *International Journal of Industrial Ergonomics*, vol. 95, p. 103444, 2023.

[6] G. Sreenu and S. Durai, "Intelligent video surveillance: a review through deep learning techniques for crowd analysis," *Journal of Big Data*, vol. 6, no. 1, pp. 1–27, 2019.

[7] J. Usha Rani and P. Raviraj, "Real-time human detection for intelligent video surveillance: an empirical research and in-depth review of its applications," *SN Computer Science*, vol. 4, no. 3, p. 258, 2023.

[8] J. Xu, "A deep learning approach to building an intelligent video surveillance system," *Multimedia Tools and Applications*, vol. 80, no. 4, pp. 5495–5515, 2021.

[9] C. Sindhuja, K. . Srinivasagan, and S. Kalaiselvi, "An efficient method for crowd event recognition based on motion patterns," *2014 International Conference on Recent Trends in Information Technology*, pp. 1–6, 2014.

[10] E. L. Piza, B. C. Welsh, D. P. Farrington, and A. L. Thomas, "Cctv surveillance for crime prevention: A 40-year systematic review with meta-analysis," *Criminology & public policy*, vol. 18, no. 1, pp. 135–159, 2019.

[11] S. Yao, B. R. Ardabili, A. D. Pazho, G. A. Noghre, C. Neff, and H. Tabkhi, "Integrating ai into cctv systems: A comprehensive evaluation of smart video surveillance in community space," *arXiv preprint arXiv:2312.02078*, 2023.

[12] A. Adefemi, E. A. Ukpoju, O. Adekoya, A. Abatan, and A. O. Adegbite, "Artificial intelligence in environmental health and public safety: A comprehensive review of usa strategies," *World Journal of Advanced Research and Reviews*, vol. 20, no. 3, pp. 1420–1434, 2023.

[13] G. Lavanya and S. D. Pande, "Enhancing real-time object detection with yolo algorithm," *EAI Endorsed Transactions on Internet of Things*, vol. 10, 2024.

[14] T. Saheb, "Ethically contentious aspects of artificial intelligence surveillance: a social science perspective," *AI and Ethics*, vol. 3, no. 2, pp. 369–379, 2023.

[15] K. C. Yam, T. Tan, J. C. Jackson, A. Shariff, and K. Gray, "Cultural differences in people's reactions and applications of robots, algorithms, and artificial intelligence," *Management and Organization Review*, vol. 19, no. 5, pp. 859–875, 2023.

[16] L. Sintonen, H. Turtiainen, A. Costin, T. Hamalainen, and T. Lahtinen, "Osrm-cctv: Open-source cctv-aware routing and navigation system for privacy, anonymity and safety (preprint)," *arXiv preprint arXiv:2108.09369*, 2021.

[17] D. Almeida, K. Shmarko, and E. Lomas, "The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: a comparative analysis of us, eu, and uk regulatory frameworks," *AI and Ethics*, vol. 2, no. 3, pp. 377–387, 2022.

[18] C. Eric, "What ai-driven decision making looks like," *Harv Bus Rev. https://hbr. org/2019/07/what-ai-driven-decision-making-looks-like*, 2019.

[19] O. Brdiczka, "Contextual ai: The next frontier of artificial intelligence," *Adobe. Retrieved Sept*, vol. 10, p. 2020, 2019.

[20] B. E. Weeks and D. S. Lane, "The ecology of incidental exposure to news in digital media environments," *Journalism*, vol. 21, no. 8, pp. 1119–1135, 2020.

[21] S. Ali, T. Abuhmed, S. El-Sappagh, K. Muhammad, J. M. Alonso-Moral, R. Confalonieri, R. Guidotti, J. Del Ser, N. Díaz-Rodríguez, and F. Herrera, "Explainable artificial intelligence (xai): What we know and what is left to attain trustworthy artificial intelligence," *Information fusion*, vol. 99, p. 101805, 2023.

[22] A. Aldoseri, K. N. Al-Khalifa, and A. M. Hamouda, "Re-thinking data strategy and integration for artificial intelligence: concepts, opportunities, and challenges," *Applied Sciences*, vol. 13, no. 12, p. 7082, 2023.

[23] J. Brands, T. Schwanen, and I. Van Aalst, "What are you looking at? visitors' perspectives on cctv in the night-time economy," *European urban and regional studies*, vol. 23, no. 1, pp. 23–39, 2016.

[24] H. M. Hodgetts, F. Vachon, C. Chamberland, and S. Tremblay, "See no evil: Cognitive challenges of security surveillance and monitoring," *Journal of applied research in memory and cognition*, vol. 6, no. 3, pp. 230–243, 2017.

[25] M. Potgieter and J. Van Niekerk, "Multi-agent augmented computer vision technologies to support human monitoring of secure computing facilities," *SAIEE Africa Research Journal*, vol. 104, no. 2, pp. 80–88, 2013.

[26] K. Petrini, P. McAleer, C. Neary, J. Gillard, and F. E. Pollick, "Experience in judging intent to harm modulates parahippocampal activity: An fmri study with experienced cctv operators," *Cortex*, vol. 57, pp. 74–91, 2014.

[27] J. Tao, H. Wang, X. Zhang, X. Li, and H. wei Yang, "An object detection system based on yolo in traffic scene," *2017 6th International Conference on Computer Science and Network Technology (ICCSNT)*, pp. 315–319, 2017.

[28] Y. Li, S. Li, H. Du, L. Chen, D. Zhang, and Y. Li, "Yolo-acn: Focusing on small target and occluded object detection," *IEEE Access*, vol. 8, pp. 227 288–227 303, 2020.

[29] Q. Guo, J. Liu, and M. Kaliuzhnyi, "Yolox-sar: High-precision object detection system based on visible and infrared sensors for sar remote sensing," *IEEE Sensors Journal*, vol. 22, pp. 17 243–17 253, 2022.

[30] D. Beniz and A. Espindola, "Using tkinter of python to create graphical user interface (gui) for scripts in lnls," pp. 56–58, 2017.

[31] C. J. Howard, T. Troscianko, I. D. Gilchrist, A. Behera, and D. C. Hogg, "Suspiciousness perception in dynamic scenes: a comparison of cctv operators and novices," *Frontiers in human neuroscience*, vol. 7, p. 441, 2013.

[32] E. M. Crowe, C. J. Howard, I. D. Gilchrist, and C. Kent, "Motion disrupts dynamic visual search for an orientation change," *Cognitive research: principles and implications*, vol. 6, no. 1, p. 47, 2021.

[33] B. Heebels and I. van Aalst, "Surveillance in practice: operators' collective interpretation of cctv images," *Surveillance & Society*, vol. 18, no. 3, pp. 312–327, 2020.