Hybrid Security Systems: Human and Automated Surveillance Approaches

Mohammed Ameen[1], Richard Stone[2], Ulrike Genschel[3], Fatima Mgaedeh[4]
Human-computer Interaction Department, Iowa State University, Ames, USA[1]
Industrial and Manufacturing Systems Engineering Department, Iowa State University, Ames, USA[2]
Department of Statistics, Iowa State University, Ames, USA[3]
Department of Industrial Engineering, Jordan University of Science and Technology, Irbid, Jordan[4]

***Abstract*—The study investigates the performance of hybrid security systems under different personnel training and artificial intelligence (AI) assistance conditions. The aim is to understand the system's impact on different scenarios that involve human operators and AI and to develop a predictive model for optimizing system performance. A human security information model was built to predict the performance of hybrid security systems. The system's performance metrics (response time, hits, misses, mistakes), cognitive load, visual discrimination, trust, and confidence were measured under different training and assistance conditions. Participants were divided into trained and non-trained groups, and each group performed surveillance tasks with and without AI assistance. Predictive modeling was performed using Linear Regression. The training significantly improved performance by reducing misses and mistakes and increasing hits, both with and without AI assistance. In the non-trained group, AI assistance boosted speed and hit accuracy but led to more mistakes. AI assessment reduced response time and misses for the trained group while increasing hits without affecting the mistake rate. Trust and confidence were higher with AI in the non-trained group, while AI reduced cognitive load in the trained group. The findings highlight the interactions between human operators, AI assistance, and training in hybrid surveillance systems. The predictive model can guide the design and implementation of these systems to optimize performance. Future studies should focus on strategies to enhance operator trust in AI-assisted systems and confidence, further optimizing the collaborative potential of hybrid surveillance frameworks.**

***Keywords*—Hybrid surveillance systems; human-AI interaction; operator training; predictive modeling; linear regression**

I. Introduction

The delicate balance between human judgment and artificial intelligence (AI) in surveillance is critical. While AI-powered automated systems have demonstrated remarkable capabilities in enhancing the efficiency and effectiveness of monitoring tasks, their inherent limitations necessitate the continued involvement of human operators [1]. Extensive research has demonstrated that deploying monitoring and surveillance devices, such as cameras and sensor-operated security systems, is crucial in reducing crime rates [2]. The demand for security services has evolved beyond monitoring criminal activities to encompass detecting and tracking abnormal behaviors [3]. Such behaviors, mainly in crowded or densely populated areas, can pose significant risks. These gatherings, often driven by religious, cultural, or social events, necessitate heightened security measures due to their importance and potential disruptions. The need for advanced security surveillance cameras has never been more critical. Integrating intelligent surveillance systems capable of identifying suspicious or abnormal behaviors is indispensable. Moreover, the effectiveness of these systems is greatly enhanced by the presence of qualified and trained personnel who can operate and interpret the data from these devices, thereby completing the security framework [4]. Ensuring public safety and security is paramount, surpassing all other considerations. In the absence of security, the fundamental components of life disintegrate. Motivated by this imperative, our project aims to elevate the quality of security surveillance systems by enhancing their ability to detect abnormal behaviors and assist personnel in their duties [5]. This approach provides crucial guidance and optimizes efforts, ensuring a more secure environment.

A. Surveillance Systems Evolution

Traditional security surveillance systems have evolved significantly, transitioning from basic physical security measures to sophisticated technological solutions. Initially, security systems primarily involved visual monitoring, which was adequate but limited by human capabilities and response times. Technology became essential as the need for more efficient and reliable security solutions grew. This evolution introduced electronic surveillance systems, which have become a cornerstone of modern security strategies. Among these advancements, Closed-Circuit Television (CCTV) systems emerged as a supportive technology, providing real-time monitoring capabilities and enhancing the overall effectiveness of security operations.

In the 1940s, CCTV first appeared, primarily gaining traction within security contexts. Germany pioneered installing the world's inaugural CCTV system [6] [7]. Subsequently, British law enforcement deployed CCTV during political demonstrations in central London. However, this early adoption faced significant challenges due to costs [8]. Since these initial implementations, CCTV technology has undergone substantial advancements. Improvements have encompassed enhanced visual quality, data storage, remote accessibility, and the integration of automated detection systems powered by artificial intelligence.

B. Automated Surveillance Technologies Implications on Security

Artificial Intelligence (AI) technologies significantly enhance Closed-Circuit Television (CCTV) systems by

introducing advanced features such as facial recognition, behavior analysis, and real-time threat detection. These capabilities allow for proactive surveillance, expanding the effectiveness of traditional CCTV setups. Automated surveillance can quickly and accurately analyze large amounts of data, identifying potential security threats that might be missed by human operators. This reduces the high cognitive load level on CCTV operators, allowing them to focus on critical incidents that require human judgment [9]. Thereby increasing their productivity and reducing the number of personnel needed simultaneously [10].

Despite the advancements and positive impacts of AI on surveillance systems, the human element remains crucial [9]. While AI can support routine monitoring tasks, it cannot entirely replace human intuition and expertise. Comprehensive training programs for CCTV operators are essential, focusing on how to handle various aspects of their work environment, job roles, skills and competencies, and the nature of the places they monitor [11]. Effective training ensures that operators can manage tasks, understand their responsibilities clearly, and develop the necessary competencies to perform their tasks efficiently [12]. By being well-prepared to deal with the complexities of the environments they oversee, operators can leverage both human intuition and advanced AI technologies to enhance their monitoring capabilities and respond more effectively to security threats.

C. Hybrid Security Systems

Hybrid surveillance systems offer numerous benefits by combining AI's precision and speed with human oversight's contextual understanding, positively affecting performance. These systems enhance accuracy and reduce false alarms by automating routine tasks and analyzing vast amounts of data in real-time, which lowers the workload on human operators and allows them to focus on more complex tasks, thereby improving overall performance [13]. The reduced cognitive load enables operators to maintain higher levels of alertness and efficiency [14]. Enhanced visual discrimination is achieved as AI quickly identifies patterns and anomalies, assisting human operators in detecting subtle differences that might be missed otherwise. This collaboration fosters greater confidence and trust in the system as operators can rely on AI to provide accurate initial assessments, ensuring quicker, more accurate, and contextually appropriate responses [15]. Ultimately, these improvements contribute to a more effective and reliable surveillance operation.

II. OMAR Framework

In this study, we test a new framework that will positively affect the security system. The Operator Machine Augmentation Resource (OMAR) framework is a comprehensive system designed to enhance the efficiency and effectiveness of CCTV surveillance operations. OMAR integrates advanced technologies such as a Computer Vision model, human training techniques, and alert triggers to address limitations in traditional surveillance systems. The framework improves the productivity of surveillance by facilitating operator tasks and reducing human effort, ultimately enhancing the quality of security. It includes components like a detection model using the YOLO (You Only Look Once) object detection system, which efficiently analyzes live video feeds for real-time object detection and annotation. OMAR training sessions are designed to cultivate a broad set of skills and competencies, thereby equipping CCTV operators with the necessary knowledge and expertise to effectively monitor and manage surveillance environments [16]. The rationale behind OMAR is to create a hybrid system that leverages both AI and human oversight, combining the strengths of each to achieve better accuracy, reduce false alarms, and improve overall surveillance efficacy.

III. Literature Review

It has become evident that the use of surveillance systems has increased dramatically in the past decade, mainly due to the computerization of some of these techniques to fight terrorism and other activities that lead to increased crime rates. These systems play critical roles in guaranteeing security and detecting and managing large crowds in different settings. Monitoring systems are generally categorized into two types: Vision-based and non-vision-based.

A. Monitoring System

1) Vision-based systems: Vision-based systems mainly rely on cameras, leveraging advanced image-processing technologies and computer vision models to ensure safety and security. These systems are extensively deployed in urban areas, business districts, commercial hubs, and transportation centers, aiming to mitigate insecurity and enhance public safety [17]. The integration of computer vision within these systems enables sophisticated functionalities such as facial recognition, behavior analysis, and anomaly detection, significantly improving their efficacy and reliability. By incorporating these elements, AI-driven monitoring systems provide a robust framework for proactive and reactive security measures, facilitating real-time monitoring and prompt response to potential threats.

2) Non-Vision-based systems: Other forms of monitoring mechanisms rely on other means to detect and observe parts of the physical space where vision-based surveillance is challenging or cannot be applied. These systems are particularly useful, especially when issues such as the absence of light or something obstructing sight make using cameras less effective. Popular non-vision-based system tools include Wi-Fi, Bluetooth, Radio frequency identification, RFID, and cellular networks [18]. Bluetooth is a short-range and low-cost wireless technology designed with features similar to Wi-Fi sets but with less coverage range [19]. It is commonly used in premises monitoring to track a person's or object's slow movement. Bluetooth technology, however, uses personal devices to track the movement and location of Bluetooth-enabled devices compared to Wi-Fi technology, which uses access points to monitor the movement and location of Wi-Fi devices. This technology can be used particularly well in crowded areas that are difficult to maintain order within, such as airports, malls, and stadiums. RFID, or Radio Frequency Identification Technology, is the system of using radio frequencies to verify the identity of an individual and or object tagged [19]. RFID systems can be of two types: one type does not have energy resources, and the second has energy resources inbuilt in them and can have better

signal transmission and sound range than the first ones, known as RFID tag passive or active accordingly [20]. Cellular network monitoring involves transmitting information between mobile devices and cells. This technology is crucial for tracking the location and movement of mobile users over vast geographical areas. Effective utilization of cellular networks for surveillance purposes requires the cooperation of multiple mobile network operators to ensure seamless data transmission and coverage [18] [21]. Additionally, advancements in 5G technology promise to further enhance the capabilities of cellular-based surveillance systems by providing higher data rates, lower latency, and more reliable connections.

Non-vision-based surveillance technologies require optimal conditions to function effectively. Moreover, the receivers associated with these systems are susceptible to deliberate interference and manipulation by individuals. Based on the facts presented, although both vision and non-vision systems are essential components of modern security systems, the vision-mode systems possess certain advantages noteworthy on the significance of real-time controls, and additional capabilities originating from AI technologies. With these capabilities, vision-based surveillance systems are more suitable for most applications, especially in areas where detailed monitoring is required and prompt action is sensitive, as in urban and highly crowded regions.

B. Automated Systems

To fully leverage the vision-based monitoring system technologies, it is crucial to integrate these systems with advanced artificial intelligence (AI), machine learning (ML), and deep learning (DL) methodologies. Recent advancements in AI have led to the development and deployment of various sophisticated techniques, each evaluated based on their efficacy in identifying anomalous behavior. This comparative analysis has revealed significant improvements in surveillance capabilities, underscoring AI's critical role in enhancing modern surveillance systems' accuracy and reliability. In recent years, a diverse exhibition of models and techniques has been extensively tested, including Generative Adversarial Networks (GANs), Convolutional Neural Networks (CNNs), Long Short Term Memory (LSTM) networks, Gaussian Mixture Models (GMMs), Support Vector Machines (SVMs), and Random Forests (RF) [22]. Numerous studies have employed these techniques to identify behaviors that could potentially disrupt crowds. In our research, we drew upon prior work that has categorized actions such as standing, sitting, sleeping, running, moving in opposite or different crowd directions, and non-pedestrian movements such as cars and wheelchairs as abnormal behaviors that could compromise the safety and flow of moving crowds [23] [24]. Although these methodologies have been proven effective in their respective functions, their effectiveness nevertheless has its drawbacks. These gaps imply that human oversight is still very relevant in order to come up with results that can meet all the parameters of precision [9]. Despite the critical role played by CCTV operators, humans need to improve their ability to monitor large crowds over extended periods effectively. This limitation arises from human cognitive capacity constraints, which deteriorate under prolonged monitoring tasks, leading to lowered performance. Consequently, there is a pressing need for an auxiliary element, such as AI, to augment human capabilities. AI can significantly enhance the efficiency and ease of surveillance operations, improving overall performance. Notably, while AI provides substantial support, it is not intended to replace the human presence but rather to complement and optimize human efforts in surveillance tasks.

Research on the training of CCTV operators is notably lacking, primarily focusing on applying general psychological theories. One notable study suggests that an individual's situational awareness significantly enhances operational performance [25]. Existing literature predominantly aims at improving the efficiency of CCTV operators while concurrently minimizing their cognitive load [26].

IV. Methodology

A. Participants

This study recruited 30 participants, aged between 20 and 49, through flyers that provided detailed information about the research. These flyers were distributed to both Iowa State University students and residents. All participants gave informed consent prior to their involvement in the study. The research procedures adhered to ethical guidelines and were approved by the Human Institutional Review Board (IRB) at Iowa State University. To qualify for the experiment, participants needed to be physically and mentally capable of meeting the study's demands. It included being physically present for the entire duration of the study sessions and being able to handle the physical requirements without experiencing excessive fatigue or discomfort. Participants had to be mentally prepared to manage any potential stress or discomfort associated with the study. Furthermore, normal visual acuity was a prerequisite for participation.

B. Experimental Design

The main objective of this study was to evaluate and enhance the performance, visual discriminations, cognitive load, trust, and confidence for both trained and non-trained groups. The design was adopted to evaluate two independent variables: personnel and system. One-way ANOVA and T-test were performed, and all participants were distributed randomly between two groups. The study spanned 18 days, with participants returning for a second visit four days after the first visit and a third visit two weeks after the second visit to evaluate their performance. Each observation session lasted 20 minutes. The first group had a training session, and the second group had no training, but both groups were tested with an assisted and no assisted system.

In this study, we employed an experimental design incorporating two independent variables. The first independent variable is the level of personnel training, which is categorized into two groups: trained personnel and untrained personnel. Personnel variables are essential in assessing the impact of professional training on the study's outcomes. The second independent variable is system. The system variable is similarly divided into two levels: the non-assisted system and the assisted system. Both independent variables are critical to our investigation, enabling a comprehensive analysis of the interplay between human training and technological support see Fig. 1.

			System	
			Non-assisted system	Assisted system
		Levels	**1**	**2**
Personnel	Non-trained personnel	**1**	Y11	Y12
	Trained personnel	**2**	Y21	Y22

Fig. 1. Factors, levels, and treatment combination yields.

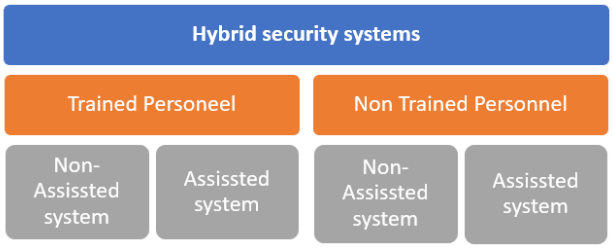

Fig. 3. Hybrid security systems breakdown.

C. Measurements

This study encompasses five dependent variables: performance, cognitive load, visual discrimination, trust, and confidence. Performance is assessed through response time, the number of hits, errors, and misses. Cognitive load is quantified using the NASA-TLX scale [27], [28]. Visual discrimination is evaluated by categorizing participants' responses on a point scale, awarding one point for each correctly identified abnormal behavior and zero points for failures to recognize abnormal behaviors. Trust is scaled on a continuum from 0 (no trust in the system) to 100 (complete trust in the system). Confidence is measured on a similar scale, ranging from 0 (no confidence in decisions) to 100 (complete confidence in decisions) see Fig. 2.

Dependent Variable	Metric	Components	Measurement Frequency
Performance	Hits	Number of catch	During the trial
Performance	Error	False Alarm (FA) Missed Hit (H)	During the trial
Performance	Response time	The time between when a target appears on the monitor and the decision taken	During the trial
Cognitive load	NASA-TLX	Scale	After trial
Visual discrimination	Scale 0-1	1 = identify behaviors 0 = could not identify behaviors	During the trial
Trust	Trust Survey	Likert scale	After trial
Confidence	Confidence Survey	Likert scale	After trial

Fig. 2. Description of dependent variable metrics, units, and frequencies.

D. Procedure

Detailed information about the research objectives and study procedures was presented to participants to ensure clarity and understanding and eliminate any potential bias before commencing the study. The study lasted 18 days weeks to determine factors that can influence the performance of the participants. The protocol will include having the participant randomly assigned to one of two groups of getting security training or not (between subject). Then each participant in each group (trained or not trained) will be tested under two conditions of system monitoring (AI Assistant system and no AI Assistant system) (within subjects) see Fig. 3.

1) Visit 1: Participants' visual acuity was assessed using the Snellen eye chart test, while their dominant eye was determined through the Porta Test, a sighting test designed for this purpose [29]. Additionally, their color vision was evaluated using a series of plates, each featuring a circle filled with numerous small colored dots that form numbers. Those in the trained group experienced a dedicated training session to enhance their performance. Upon completion of these examinations and training, participants completed a survey see Fig. 4.

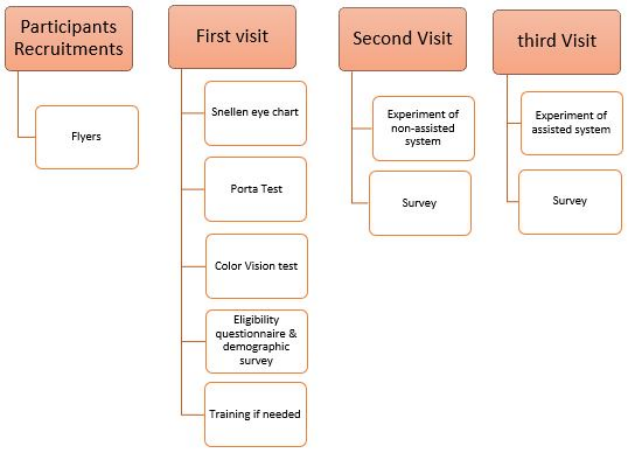

Fig. 4. Study procedure.

2) Visit 2: To mitigate immediate recall bias, participants returned to perform the experiment four days after the initial visit. Participants from both groups were asked to watch the 20 minutes long video and indicate any observed abnormal behaviors within the crowd by moving the cursor to the target and providing detailed explanations of their observations. Additionally, participants were instructed to describe any abnormal behaviors that could disturb the walking crowd verbally. The collected verbal protocol data was utilized for analysis. Following the video task, participants completed a survey to assess their overall experience and the effectiveness of the experiment see Fig. 4.

3) Visit 3: We applied our algorithm to the video to detect abnormal behaviors. Participants from both groups watched the 20 minutes long video, identified abnormal behaviors that could disturb the walking crowd by moving the cursor to the target, and described the specifics of their observations. Verbal protocols were used. Following the video task, participants completed a survey to assess their overall experience and the experiment's effectiveness see Fig. 4.

V. Result

All data were analyzed using SPSS 28. We used One-way ANOVA to assess mean differences between trained and non-trained groups with and with no AI assistants in monitoring abnormal behavior in terms of performance (response time, miss, hit, and mistake) and measuring the trust, confidence, cognitive load, and visual discrimination level. Also, a Paired-Sample T-test was needed to discover the efficiency of AI assistance compared to no AI assistance for each group separately to distinguish the individual differences in performance (response time, miss, hit, and mistake) and measure the trust, confidence, cognitive load, and visual discrimination level.

A. Trained and Non-Trained Groups with AI Assistant (Between-subject)

1) Performance: H1: While monitoring abnormal behaviors through CCTV, there will be a significant difference in performance (response time, miss, hit, and mistake) between trained and no trained groups with AI assistants.

The ANOVA test comparing the performance of trained and non-trained groups with AI assists in response time, misses, hits, and mistakes. The trained group did not show a significant difference in the response time (F (1, 28) = 0.059, p = 0.810) compared to the non-trained group. On the other hand, the trained group significantly had fewer missed incidents (mean difference = 29.34, F (1, 28) = 30.838, p ¡ 0.001) compared to the non-trained group. Also, the trained group significantly had fewer mistakes in catching incidents (mean difference 19.47, F (1, 28) = 48.532, p ¡ 0.001) compared to the non-trained group. Moreover, the trained group had significantly more incident hits (mean difference = 29.34, F (1, 28) = 30.838, p ¡ 0.001) compared to the non-trained group. Therefore, these results suggest that training significantly improves performance by reducing misses and mistakes and increasing hits, as shown in Fig. 5.

Fig. 5. The average of the trained and non-trained groups with AI assistants in a number of mistakes misses, and hits.

2) Trust, Confidence, Cognitive Load, and Visual Discrimination: H2: While monitoring abnormal behaviors through CCTV, trust level will be significantly different between trained and no trained groups with AI assistance.

H3: While monitoring abnormal behaviors through CCTV, confidence levels will significantly differ between trained and non-trained groups with AI assistance.

H4: While monitoring abnormal behaviors through CCTV, there will be a significant difference in cognitive load between trained and non-trained groups with AI assistance.

H5: While monitoring abnormal behaviors through CCTV, there will be a significant reduction in visual discrimination between trained and non-trained groups with AI assistance.

The ANOVA test compares the performance of trained and non-trained groups with AI assistants in terms of the trust, confidence, cognitive load, and visual discrimination. The trust score between non-trained and trained groups is not statistically significant (p = 0.445). Also, there was no significant difference between trained and non-trained in the level of confidence (p = 0.125). Cognitive load, the non-trained group also shows no significant difference compared to the trained group (p = 0.30). However, visual discrimination shows a near-significant difference; the trained had a high (mean difference of 0.025 F (1, 28) = 3.758, p = 0.063) compared to the non-trained group. These findings suggest that training does not significantly impact trust, confidence, or cognitive load but may have a marginal effect on improving visual discrimination.

B. Trained and Non-Trained Group with no AI Assistant (Between-Subject)

1) Performance: H6: While monitoring abnormal behaviors through CCTV, there will be a significant difference in performance (response time, miss, hit, and mistake) between trained and non-trained groups with no AI assistant.

The ANOVA test compares the performance of trained and non-trained groups with no AI assistant regarding response time, misses, hits, and mistakes. There is no significant difference in response time between the non-trained and trained with no AI assistant (p = 0.512). However, the trained group had significantly fewer missed incidents (mean difference = 25.53, F (1, 28) = 19.735, p ¡ 0.001) compared to the non-trained group. Moreover, the trained group had significantly more hit incidents (mean difference: 25.53, F (1, 28) = 19.735, p ¡ 0.001) compared to the non-trained group. Also, the trained group had significantly fewer mistakes in catching incidents (mean difference = 12.47, F (1, 28) = 22.783, p ¡ 0.001) compared to the non-trained group. Therefore, these results suggest that training significantly improves performance by reducing misses and mistakes and increasing hits, as shown in Fig. 6.

2) Trust, Confidence, Cognitive Load, and Visual Discrimination: H7: While monitoring abnormal behaviors through CCTV, there will be a significant difference in trust level between trained and non-trained groups with no AI assistance.

H8: While monitoring abnormal behaviors through CCTV, there will be a significant difference in confidence levels between trained and non-trained groups with no AI assistance.

H9: While monitoring abnormal behaviors through CCTV, there will be a significant difference in cognitive load between trained and non-trained groups with no AI assistance.

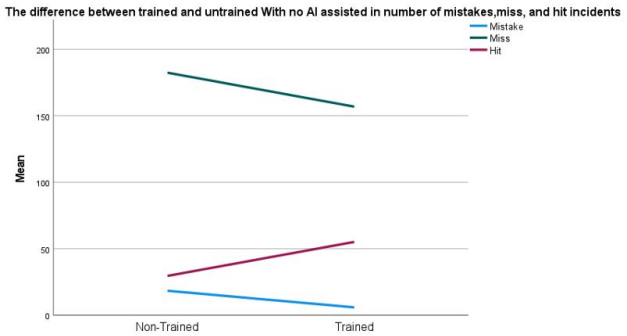

Fig. 6. The average of the trained and non-trained groups with no AI assistant in a number of mistakes, misses, and hits.

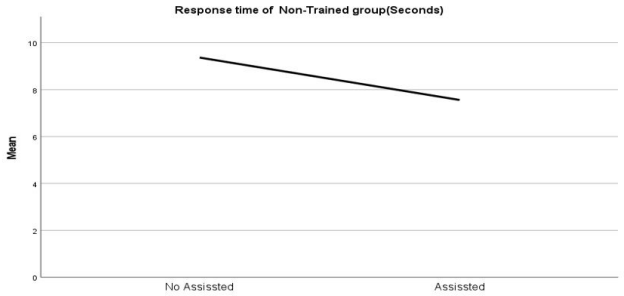

Fig. 7. The response time average of the non-trained group with AI and no AI assistants.

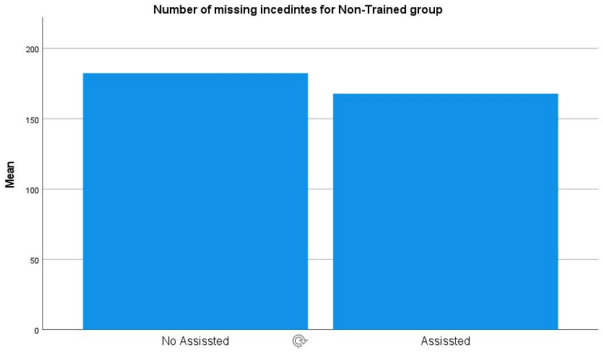

Fig. 8. The missing incidents average of the non-trained group with AI and with no AI assistant.

H10: While monitoring abnormal behaviors through CCTV, there will be a significant reduction in visual discrimination between trained and non-trained groups with no AI assistance.

The ANOVA test compares the performance of trained and non-trained groups with no AI assistant regarding trust, confidence, cognitive load, and visual discrimination. The trust score between non-trained and trained groups is not statistically significant (p = 0.397). Also, the confidence level was not significantly different between trained and non-trained (p = 0.320). However, cognitive load shows a near-significant difference, with the trained group experiencing a higher cognitive load than the non-trained group (mean difference: 10.89, F (1, 28) = 3.218, p = 0.084). Finally, visual discrimination did not show a statistically significant difference between the groups (mean difference (p = 0.303). Therefore, these results did not significantly impact trust, confidence, or visual discrimination but may increase cognitive load.

C. Non-Trained Group (Within-Subject)

1) Performance: The results of the paired t-tests reveal significant reduction in response times for AI (mean difference = 1.81 seconds, t = 2.409, p = .030) compared to no AI assistant, see Fig. 7, and number of misses (mean difference = 14.53, t = 6.200, p ¡.001). See Fig. 8, indicating that AI assistance is significantly enhance the user's performance in both time consuming to catch incidents and number of missed incidents compared to no AI assistance. Also, the number of hits of AI assistants is significantly increased (mean difference = -14.53, t= -6.200, p ¡ .001) compared to no AI assistants, see Fig. 9. However, this improvement in hits is accompanied by a significant increase in the number of mistakes (mean difference = -8.60, t= -5.644, p¡.001) see Fig. 10. Therefore, these results suggest that AI assistance boosts performance speed and hit accuracy, but it leads to a higher mistake rate.

2) Trust, Confidence, Cognitive Load, and Visual discrimination: The results of the paired t-tests reveal significant differences between the non-trained group with AI assistance and those with no AI assistance across four aspects: trust, confidence, cognitive load, and visual discrimination. The user's trust level with AI assistance is significantly higher (mean difference = -15.667, t = -3.063, p = .008) than with no AI assistance. Also, the user's confidence level is highly significant with AI assistance (mean difference = -22, t = -4.069, p = .001). However, cognitive load results were not statistically significant (p = .100), and AI assistance did not significantly affect visual discrimination (p = .718) compared to no AI assistance see Fig. 11.

Trained Group (Within- subject)

3) Performance: The paired t-test results revealed significant differences in the performance of the trained group with an AI assistant and no AI assistant. Participants with AI assistants spent significantly less time catching incidents (Mean difference 1.12 seconds, t = 2.221, p = .043) than no AI assistant see Fig. 12. Also, participants with AI assistants had significantly fewer miss incidents (mean difference 18.33, t = 6.510, p ¡ .001) compared to the no AI assistant see Fig. 13. Moreover, participants with AI assistant significantly had higher hits incidents (mean difference = 18.33 hits, t = -6.510, p ¡ .001) compared to the no AI assistant see Fig. 14. However, there is no significant difference in the number of mistakes in catching incidents between AI assistance and no AI assistance (p = .164) see Fig. 15.

4) Trust, Confidence, Cognitive Load, and Visual discrimination: The paired t-test results revealed a significant difference between AI and no AI assistant in cognitive load; however, there was no significant difference between the others. The participants with AI assistance had a significantly less cognitive load (mean difference = 15.94, t = 2.151, p = .049) than the no AI assistant see Fig. 16. However, there was no significant difference in trust level (p = .946). Similarly, there was no significant difference in confidence level (p= 1.00). Lastly, visual discrimination also showed no significant difference (p = .455).

Fig. 9. The hits incidents average of the non-trained group with AI and with no AI assistant.

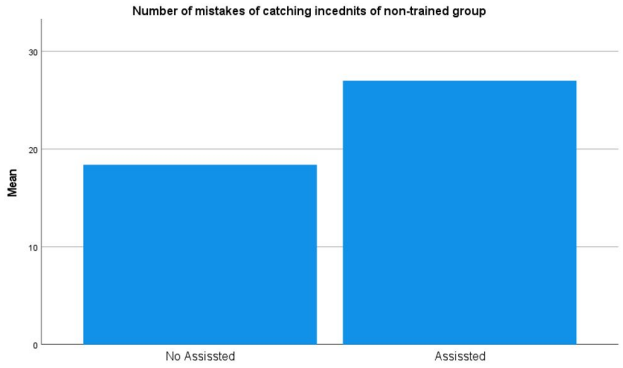

Fig. 10. The hits accompanied by a significant increase in the number of mistakes.

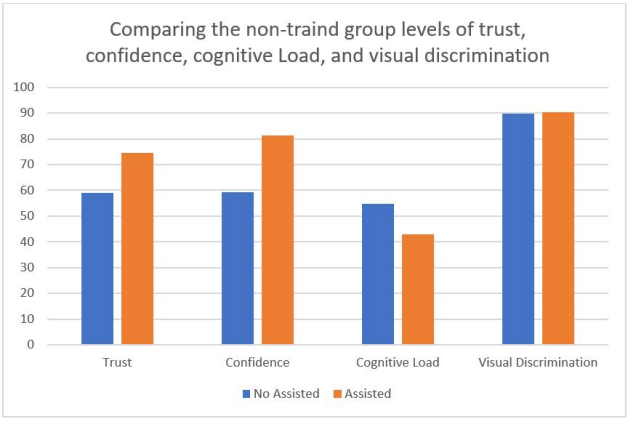

Fig. 11. The average of trust, confidence, cognitive load, and visual discrimination for a non-trained group with and with no AI assistance.

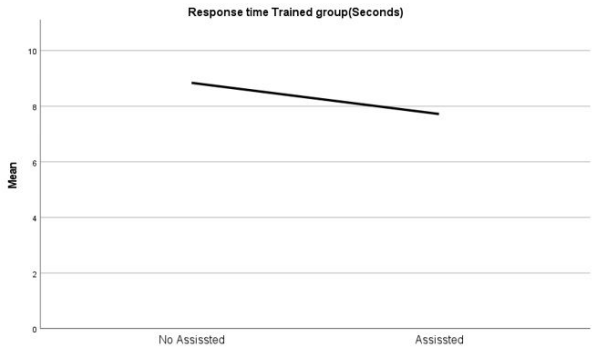

Fig. 12. The response time average of the trained group with AI and no AI assistants.

VI. Modeling

A. Data Preprocessing

Data preprocessing was completed. it transformed the data into a meaningful, efficient format, ready for machine learning models. This study focuses on two categorical variables: personal and system. These categorical variables were transformed using the one-hot encoding method. One-hot encoding involves converting each unique category within a categorical variable into a separate binary feature in a new column. Consequently, for each observation, a binary indicator of 1 is assigned to the feature corresponding to its original category, while all other features receive a binary value of 0. This method generates a new binary feature for each possible category, improving the model accuracy and predictive analysis.

B. Multicollinearity

The Variance Inflation Factor (VIF) method was utilized to quantify the degree of multicollinearity among the regression variables. Multicollinearity occurs when two or more predictors exhibit a high degree of correlation simultaneously, potentially reducing the statistical significance of individual independent variables [30]. The Variance Inflation Factor (VIF) values will uniformly be 1 in procedures with no correlated predictors. VIF values exceeding five indicate multicollinearity and may consider further investigation or removal from the model (see Table I). The VIF is calculated using the following formula:

$$\mathrm{VIF}_j = \frac{1}{1 - R_J^2} \tag{1}$$

TABLE I. Variance Inflation Factors for Features

Feature	VIF
Trust	2.458356
Confidence	2.088327
Mental	2.927619
Physical	1.284067
Temporal	2.227160
Performance	1.855252
Effort	3.794435
Frustration	2.163868
Training-Non-trained	∞
AIassisstance_AI_Assisstance	∞
AIassisstance_No_AI_Assisstance	∞

C. Model Development

This study employed linear regression to predict performance. Linear regression is a statistical approach used for modeling the association between a dependent variable and independent variables by providing a linear equation to observed data. The purpose of linear regression

Fig. 13. The missing incidents average of the trained group with AI and with no AI assistant.

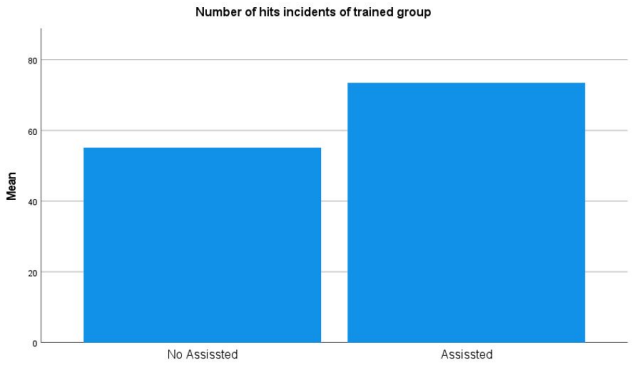

Fig. 14. The hits incidents average of the trained group with AI and with no AI assistant.

is to predict the dependent variable based on the values of the independent variables [31]. This method is favored for its simplicity, interpretability, and efficiency in modeling linear relationships, making it widely applicable in various fields such as predictive study, elucidating variable relationships, and data sciences. Linear regression is particularly valued for its ability to provide clear insights into the strength and direction of relationships between variables and for its utility in predictive analytics.

D. Performance Metrics

In our study, we assessed the performance of our predictive models using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). They are widely utilized metrics for assessing models [32]. Both metrics are essential for evaluating the accuracy of our models and quantifying the deviation from the actual values. MAE is the mean of the gap between the anticipated values and the actual values of the target variable (see Eq. 2). RMSE, on the other hand, is calculated as the square root of the average of the squared errors (see Eq. 3). These metrics enabled us to rigorously evaluate the precision of the models and facilitate comparative performance analysis. MAE and RMSE deliver a comprehensive study of model error.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

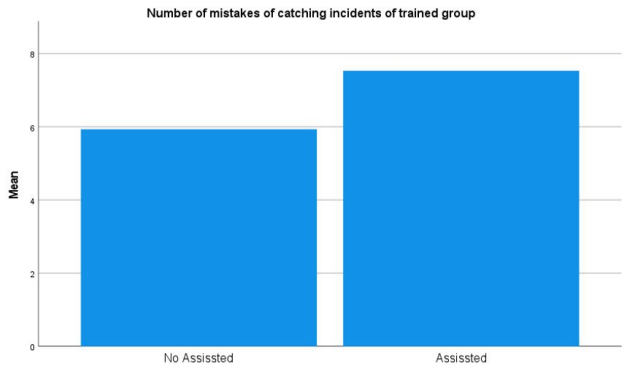

Fig. 15. No significant difference in the number of mistakes in catching incidents between AI and no AI assistance.

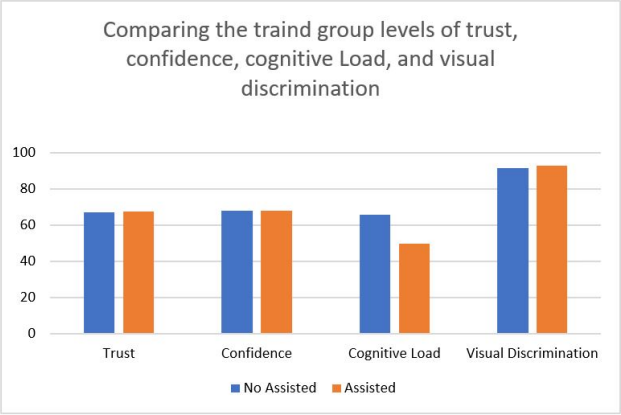

Fig. 16. The average of trust, confidence, cognitive load, and visual discrimination for a trained group with and with no AI assistance.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3}$$

E. Prediction Results

TABLE II. PERFORMANCE ANALYSIS FOR DIFFERENT MODELS ON TEST DATA

Metrics	LR
RMSE	17.247 (3.05)
MAE	14.031 (2.67)

The performance was assessed using the test metrics MAE and RMSE (Table II). Fig. 17 compares actual versus predicted values using scatter plots for the Linear Regression model. The scatter plot for the LR model shows a wide dispersion of data points, indicating the prediction error of the model. This analysis demonstrates that the Linear Regression model provides a reasonably accurate prediction with a lower MAE and RMSE.

VII. DISCUSSION

This study explored the effectiveness of hybrid security systems that integrate human oversight with automated

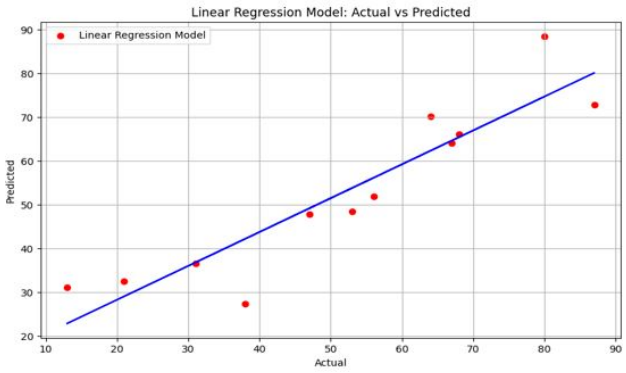

Fig. 17. Comparison of actual vs. predicted values for linear regression.

surveillance powered by advanced AI technologies. The findings emphasize the critical interplay between human operators and AI systems in enhancing the overall performance of surveillance operations. Our results demonstrate that AI-assisted surveillance systems significantly improve the detection of abnormal behaviors compared to systems solely reliant on human operators. These automated capabilities allow for proactive monitoring, reducing the cognitive load on human operators and enabling them to focus on critical incidents requiring human judgment and expertise. The study revealed that trained personnel significantly outperformed untrained personnel in identifying incidents, both with and without AI assistance. Specifically, the trained group exhibited fewer missed incidents and mistakes and more behavior identification incidents. In addition, it highlights the importance of comprehensive training for CCTV operators, ensuring they can effectively collaborate with AI systems to enhance surveillance efficacy. The performance metrics assessed included response time, number of hits, misses, and mistakes. AI assistance notably reduced response times and increased the number of hits for both trained and untrained groups. However, it also led to increased mistakes among the non-trained group, suggesting that while AI enhances performance, it requires the human operator's expertise to mitigate errors effectively [33]. The cognitive load, measured using the NASA-TLX scale, showed mixed results. For the non-trained group, AI assistance did not significantly impact cognitive load, while on the contrary, for the trained group, AI assistance resulted in a significantly lower cognitive load. The result indicates that trained personnel can better leverage AI capabilities to reduce mental strain, enhancing their performance and efficiency. The study also examined trust, confidence, and visual discrimination. While AI assistance did not significantly impact trust and confidence levels for either group, it could marginally improve visual discrimination among trained personnel. Also, operators can benefit from AI assistance to enhance their ability to discern subtle differences and abnormalities in monitoring footage when adequately trained. Our modeling efforts involved linear regression to predict performance metrics based on various factors. The feature importance analysis revealed that factors such as training level and AI assistance were significant predictors of surveillance efficacy. These findings emphasize surveillance performance's multifaceted nature, where human and technological factors interplay to determine overall effectiveness.

The integration of AI technologies in surveillance systems significantly enhances their effectiveness, particularly when complemented by well-trained human operators. The findings underscore the necessity for continuous training and support for CCTV operators, ensuring they can leverage AI capabilities to their full potential. Furthermore, the hybrid approach of combining AI precision with human contextual understanding offers a balanced solution that maximizes the strengths of both elements.

VIII. Future Work

While the study highlights the benefits of integrating AI with human oversight in surveillance systems, future research can explore the gamification of these systems to boost participant motivation and interaction [34].

Gamification uses game-design elements in non-game settings to improve engagement and motivation [35]. In hybrid surveillance systems, gamification can:

- Enhance Training: Gamified training sessions with points, badges, and leaderboards can make learning enjoyable and effective.
- Provide Real-time Feedback: Scoring systems and instant rewards can reinforce positive behaviors and enhance attentiveness.
- Boost Cognitive Engagement: Challenges and missions can reduce monotony and cognitive fatigue, making tasks more engaging.
- Foster Collaboration: Team-based challenges can improve teamwork and collective performance in large operations.

Gamifying hybrid surveillance systems can enhance operator engagement and performance, leveraging the full potential of both human and AI capabilities for more effective surveillance operations [36].

Another area for exploration is the inclusion of more advanced machine learning models such as Random Forest, which has shown promise in predicting surveillance performance metrics. Random Forest, known for its robustness and versatility, can handle complex interactions between features and offer unique insights. However, Random Forest models might be promising but require larger studies to assess their performance adequately. Hyperparameter tuning, such as optimizing the number of trees, maximum depth, and other parameters, can improve the model's performance. Investigating the importance of different features in the Random Forest model can also provide deeper insights into the factors that significantly impact surveillance performance.

IX. Conclusion

The study provides evidence that hybrid surveillance systems, which integrate AI with human oversight, enhance detection capabilities, reduce cognitive load, and improve overall performance. Future research should focus on strategies to enhance operator trust in AI-assisted systems and confidence, further optimizing the collaborative potential of hybrid surveillance frameworks. This approach will ensure safer and more secure environments in increasingly urbanized and densely populated areas.

REFERENCES

[1] A. Mumani and R. Stone, "State of the art of user packaging interaction (upi)," *Packaging Technology and Science*, vol. 31, no. 6, pp. 401–419, 2018.

[2] B. Westby, "Racial, socioeconomic, and regional effects on perception of law enforcement uniforms," in *Iowa State Conference on Race and Ethnicity*, vol. 23, no. 1. ISCORE, 2022.

[3] R. Stone, J. Kim, C. Xu, F. Mgaedeh, C. Fales, and B. Westby, "Effects of semi-automatic pistol slide pull device on law-enforcement racking process," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 66, no. 1. SAGE Publications Sage CA: Los Angeles, CA, 2022, pp. 903–907.

[4] R. Stone, M. Vasan, F. Mgaedeh, Z. Wang, and B. Westby, "Evaluation of latest computer workstation standards," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 66, no. 1. SAGE Publications Sage CA: Los Angeles, CA, 2022, pp. 853–857.

[5] R. T. Stone, S. Pujari, A. Mumani, C. Fales, and M. Ameen, "Cobot and robot risk assessment (carra) method: an automation level-based safety assessment tool to improve fluency in safe human cobot/robot interaction," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 65, no. 1. SAGE Publications Sage CA: Los Angeles, CA, 2021, pp. 737–741.

[6] M. Medwecki, "You can run but you can't hide… leveraging cctv coverage," *Business information review*, vol. 26, no. 4, pp. 244–247, 2009.

[7] T. Nagalakshmi, "A study on usage of cctv surveillance system with special reference to business outlets in hyderabad," 2012.

[8] C. A. Williams, "Police surveillance and the emergence of cctv in the 1960s," *Crime Prevention and Community Safety*, vol. 5, pp. 27–37, 2003.

[9] 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.

[10] N. Dadashi, A. W. Stedmon, and T. P. 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, no. 5, pp. 730–738, 2013.

[11] M. Ameen and R. Stone, "Operator machine augmentation resource framework art," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 6, 2024.

[12] R. T. Stone, A. M. Bisantz, J. Llinas, and V. Paquet, "Augmented multisensory interface design (amid): A human-centric approach to unisensory and multisensory augmented reality design," *Journal of Cognitive Engineering and Decision Making*, vol. 3, no. 4, pp. 362–388, 2009.

[13] T. M. Schnieders, A. A. Mumani, R. T. Stone, and B. Westby, "An analytic network process model for ranking exoskeleton evaluation criteria," *Theoretical Issues in Ergonomics Science*, pp. 1–11, 2024.

[14] R. T. Stone, K. P. Watts, P. Zhong, and C.-S. Wei, "Physical and cognitive effects of virtual reality integrated training," *Human factors*, vol. 53, no. 5, pp. 558–572, 2011.

[15] A. M. Bisantz, J. Pfautz, R. Stone, E. M. Roth, G. Thomas-Meyers, and A. Fouse, "Assessment of display attributes for displaying metainformation on maps," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 3. SAGE Publications Sage CA: Los Angeles, CA, 2006, pp. 289–293.

[16] R. Stone, K. Watts, and P. Zhong, "Virtual reality integrated welder training," 2011.

[17] J. Lee and S.-J. Shin, "A study of video-based abnormal behavior recognition model using deep learning," *International journal of advanced smart convergence*, vol. 9, no. 4, pp. 115–119, 2020.

[18] T. Wiangwiset, C. Surawanitkun, W. Wongsinlatam, T. Remsungnen, A. Siritaratiwat, C. Srichan, P. Thepparat, W. Bunsuk, A. Kaewchan, and A. Namvong, "Design and implementation of a real-time crowd monitoring system based on public wi-fi infrastructure: A case study on the sri chiang mai smart city," *Smart Cities*, vol. 6, no. 2, pp. 987–1008, 2023.

[19] R. Mubashar, M. A. B. Siddique, A. U. Rehman, A. Asad, and A. Rasool, "Comparative performance analysis of short-range wireless protocols for wireless personal area network," *Iran Journal of Computer Science*, vol. 4, pp. 201–210, 2021.

[20] W. Raad, A. Hussein, M. Mohandes, B. Liu, and A. Al-Shaikhi, "Crowd anomaly detection systems using rfid and wsn review," in *2021 4th International Symposium on Advanced Electrical and Communication Technologies (ISAECT)*. IEEE, 2021, pp. 1–5.

[21] E. Baena, S. Fortes, Ö. Alay, M. Xie, H. Lønsethagen, and R. Barco, "Cellular network radio monitoring and management through virtual ue probes: A study case based on crowded events," *Sensors*, vol. 21, no. 10, p. 3404, 2021.

[22] 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.

[23] T. Alafif, B. Alzahrani, Y. Cao, R. Alotaibi, A. Barnawi, and M. Chen, "Generative adversarial network based abnormal behavior detection in massive crowd videos: a hajj case study," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, no. 8, pp. 4077–4088, 2022.

[24] T. Alafif, A. Hadi, M. Allahyani, B. Alzahrani, A. Alhothali, R. Alotaibi, and A. Barnawi, "Hybrid classifiers for spatio-temporal real-time abnormal behaviors detection, tracking, and recognition in massive hajj crowds," *arXiv preprint arXiv:2207.11931*, 2022.

[25] 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.

[26] F. M. Donald, "A model of cctv surveillance operator performance," *Ergonomics SA: Journal of the Ergonomics Society of South Africa*, vol. 22, no. 2, pp. 2–13, 2010.

[27] J. C. Byers, A. Bittner, and S. G. Hill, "Traditional and raw task load index (tlx) correlations: Are paired comparisons necessary," *Advances in industrial ergonomics and safety*, vol. 1, pp. 481–485, 1989.

[28] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Advances in psychology*. Elsevier, 1988, vol. 52, pp. 139–183.

[29] H. L. Roth, A. N. Lora, and K. M. Heilman, "Effects of monocular viewing and eye dominance on spatial attention," *Brain*, vol. 125, no. 9, pp. 2023–2035, 2002.

[30] K. P. Vatcheva, M. Lee, J. B. McCormick, and M. H. Rahbar, "Multicollinearity in regression analyses conducted in epidemiologic studies," *Epidemiology (Sunnyvale, Calif.)*, vol. 6, no. 2, 2016.

[31] K. Kumari and S. Yadav, "Linear regression analysis study," *Journal of the practice of Cardiovascular Sciences*, vol. 4, no. 1, pp. 33–36, 2018.

[32] T. Hodson, "Root-mean-square error (rmse) or mean absolute error (mae): when to use them or not," *Geoscientific Model Development*, 2022.

[33] E. S. Abdelall, Z. Eagle, T. Finseth, A. A. Mumani, Z. Wang, M. C. Dorneich, and R. T. Stone, "The interaction between physical and psychosocial stressors," *Frontiers in behavioral neuroscience*, vol. 14, p. 63, 2020.

[34] M. Hariri and R. Stone, "Trigger screen restriction framework, ios use case towards building a gamified physical intervention," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 5, 2024. [Online]. Available: http://dx.doi.org/10.14569/IJACSA.2024.0150502

[35] ——, "Gamification in physical activity: State-of-the-art," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 10, 2023. [Online]. Available: http://dx.doi.org/10.14569/IJACSA.2023.01410105

[36] ——, "Triggered screen restriction: Gamification framework," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 11, 2023. [Online]. Available: http://dx.doi.org/10.14569/IJACSA.2023.01411130